Partial Least Squares-discriminant Analysis for the Prediction of Hemodynamic Changes Using Near Infrared Spectroscopy

Youngwook Seo[1], Seungduk Lee[2], Dalkwon Koh[1], and Beop-Min Kim[1]*

[1]*Department of Biomedical Engineering, Korea University, Seoul 163-103, Korea*
[2]*Department of Biomicrosystem Engineering, Korea University, Seoul 163-103, Korea*

Using continuous wave near-infrared spectroscopy, we measured time-resolved concentration changes of oxy-hemoglobin and deoxy-hemoglobin from the primary motor cortex following finger tapping tasks. These data were processed using partial least squares-discriminant analysis (PLS-DA) to develop a prediction model for a brain-computer interface. The tasks were composed of a series of finger tapping for 15 sec and relaxation for 45 sec. The location of the motor cortex was confirmed by the anti-phasic behavior of the oxy- and deoxy-hemoglobin changes. The results were compared with those obtained using the hidden Markov model (HMM) which has been known to produce the best prediction model. Our data imply that PLS-DA makes better judgments in determining the onset of the events than HMM.

Keywords : Near-infrared spectroscopy, Brain-computer interface, Partial least squares-discriminant analysis, Hidden Markov model
OCIS codes : (170.0170) Medical optics and biotechnology; (170.2655) Functional monitoring and imaging; (170.6510) Spectroscopy, tissue diagnostics

I. INTRODUCTION

Near-infrared spectroscopy (NIRS) measures temporal and spatial fluctuations of oxy- and deoxy-hemoglobin concentrations along with total blood volume from the cortical area [1, 2]. It provides a unique tool in neuroscience, since brain activities can be detected noninvasively and potentially in mobile conditions [3, 4]. Applications of NIRS extend from a simple measurement of tissue oxygenation to large-scale imaging correlated with cognitive and motor cortex responses [5, 6].

Brain-computer interfaces (BCIs) stem from communication and control technology with which a user can facilitate interactions with the outside world using specific signals originating from the brain [7, 8]. Several modalities have been tested for potential BCI applications, including electroencephalography (EEG) [9], electrocorticography (ECoG) [10], functional magnetic resonance imaging (fMRI) [11, 12] and NIRS [13-15]. These modalities can be categorized as invasive (ECoG) and non-invasive (EEG, fMRI and NIRS) methods. Each of these modalities has advantages, such as high temporal resolution (EEG, ECoG, and NIRS) and the

ability to perform imaging of the deep brain areas (fMRI). However, these cannot be easily applied to BCI due to shortcomings, such as low spatial resolution (EEG), invasiveness (ECoG), and large size and low temporal resolution (fMRI). NIRS has a number of merits to be used for BCI since it provides high temporal resolution, real-time detection, and mobility. Also, since optical signals do not interfere with electrical signals, it can be easily combined with other systems such as EEG [16].

Regardless of which modality is used, a prediction model is needed so that the brain signal can be interpreted. So far, a number of algorithms have been developed to build a prediction model for NIRS-based BCI such as linear discriminant analysis (LDA) [6], the support vector machine (SVM) [13], and the hidden Markov model (HMM) [13, 17].

The HMM was modified from the original Markov chain theory by Baum in the 1960s [18], and it has been used for a number of applications, including speech recognition studies [19]. It has also been successfully used as a prediction model for NIRS-based BCI. Researchers compared the SVM and the HMM to evaluate an NIRS prediction model during finger tapping and motor imagery studies and concluded

---

*Corresponding author: bmk515@korea.ac.kr

Color versions of one or more of the figures in this paper are available online.

that HMM performs better than SVM [13]. HMM has also been used to classify mental arithmetic and music imagery signals in a separate study [20]. PLS-DA is a combined version of the partial least square (PLS) regression method and the LDA [6, 21]. Recently, applications of the PLS-DA have been expanded to behavioral analysis for positron emission tomography (PET) [22], the spatiotemporal study of auditory and visual perception [23], and structure seeking [17] using fMRI. The LDA was used to classify motor imagery signals using hemodynamic responses obtained by NIRS [24].

In this study, we measured the hemodynamic responses for finger tapping in the left hemisphere motor cortex using NIRS and developed an analytic tool for classification between the finger tapping task and relaxation. Using the PLS-DA method, we developed the prediction model to discriminate TASK and REST states and compared our results with those obtained using the HMM.

II. METHODS AND EXPERIMENTS

2.1. Subjects

Seven volunteers, 3 men and 4 women, with a mean age of 24 (± 5.5 years) participated in this study. All participants were right-handed and had black hair. They had no history of any neuropathologic conditions, nor were they taking any specific medicines that might interfere with the study.

2.2. Experimental Procedure

NIRS activities were recorded from the left hemisphere primary motor cortex located in the C3 area [25]. The C3 position in the brain is shown in Fig. 1. Fig. 1(a) shows the top view of the brain, and Fig. 1(b) indicates the left view of the face. A custom-made headset accommodated eight light source pairs (785 and 830 nm) and two detectors. The source and detector distances were maintained at about 3.5-4 cm, which corresponds to a penetration depth of 1 to 2 cm in the adult head [26]. The light sources are delivered using optical fibers with diameters of 400 μm, and the reflected lights are coupled with two detectors using 3 mm-diameter fiber bundles. The optical fibers are arranged
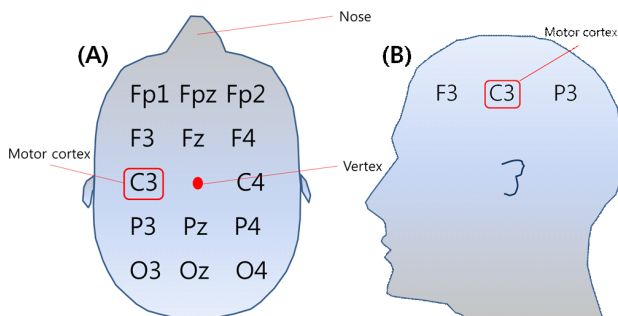

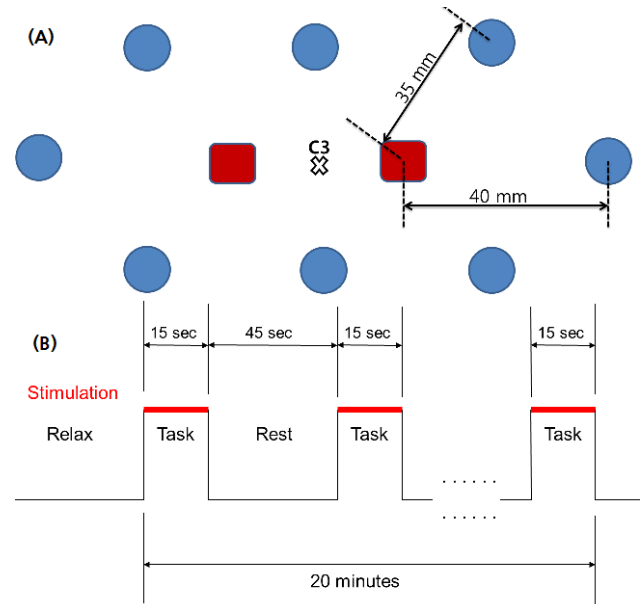FIG. 1. Location of the motor cortex in the human brain.

FIG. 2. Arrangements of the light sources (blue circles) and the detectors (red rectangles) (A) and the experimental protocol for Task/Rest sessions lasting for 20 minutes (B).

as shown in Fig. 2(a). With eight sources and two detectors, the total number of channels is 16, and the channels with large signal fluctuations were chosen for analysis.

Before the experiment, every subject was asked to sit on a chair with hands resting upon each knee. The experimental protocol is shown in Fig. 2(b), which was composed of a series of TASK and REST sessions. The TASK session was composed of 15 second-long finger tapping periods using all five fingers with frequency of 3-4 Hz, after which the subject was totally relaxed for 45 sec (REST session). These series of sessions lasted for 20 minutes. During the REST session, the subjects are asked to remain motionless.

2.3. Calculation of Hemoglobin Concentration

To remove excessive high frequency noise and low frequency drift in hemodynamic responses, we adopted a moving average filter with every 3000 points divided from raw ac intensity data. The signal was then low- and high-pass filtered with cutoff frequencies of 0.5 and 0.01 Hz.

A modified Beer-Lambert law was used to calculate the oxy-hemoglobin ($\triangle$HbO$_2$) and the deoxy-hemoglobin ($\triangle$Hbr) concentrations as in Eq. 1.

$$\Delta OD^\lambda = -\ln \frac{I_{Final}}{I_{Initial}} = (\varepsilon_{HbO_2}^\lambda \Delta[HbO_2] + \varepsilon_{Hbr}^\lambda \Delta[Hbr])B^\lambda L, \quad (1)$$

Where OD is the optical density, I_{Final} is the measured light density, $I_{Initial}$ is the initial light intensity, ε represents the extinction coefficients of the oxy and deoxy-hemoglobin, B is the differential path-length factor, and L is the source-detector separation length at the given wavelength (785 and

830 nm). The change in concentration of oxy-hemoglobin ($\triangle HbO_2$) and deoxy-hemoglobin (Hbr) were calculated using the following equations:

$$\Delta[HbO_2] = \frac{\varepsilon_{Hbr}^{\lambda_1} \dfrac{\Delta OD^{\lambda_2}}{B^{\lambda_2}} - \varepsilon_{Hbr}^{\lambda_2} \dfrac{\Delta OD^{\lambda_1}}{B^{\lambda_1}}}{(\varepsilon_{Hbr}^{\lambda_1} \varepsilon_{HbO_2}^{\lambda_2} - \varepsilon_{Hbr}^{\lambda_2} \varepsilon_{HbO_2}^{\lambda_1})L}, \tag{2}$$

and

$$\Delta[Hbr] = \frac{\varepsilon_{HbO_2}^{\lambda_2} \dfrac{\Delta OD^{\lambda_1}}{B^{\lambda_1}} - \varepsilon_{HbO_2}^{\lambda_1} \dfrac{\Delta OD^{\lambda_2}}{B^{\lambda_2}}}{(\varepsilon_{Hbr}^{\lambda_1} \varepsilon_{HbO_2}^{\lambda_2} - \varepsilon_{Hbr}^{\lambda_2} \varepsilon_{HbO_2}^{\lambda_1})L}. \tag{3}$$

2.4. Hidden Markov Model

The hidden Markov model (HMM) [18, 19] is a well-known statistical tool for modeling time series data. It contains discrete states, S, and emits an observation vector, o_t, at every time t. It uses two sets of probability vectors, including the state transition probability distribution $\mathbf{A} = \{a_{ij}\}$ and the observation probability distribution, $\mathbf{B} = \{b_j(o_t)\}$ at each state $j = 1, \cdots, S$. In general, each observation probability distribution is modeled by Gaussian mixtures [13]. To determine these parameters, we adopted the classification algorithm defined in a previous study [17]. So-called "train and test" was performed using the hemodynamic responses data from each participant; the whole data set was divided into two sub-sets: one training set and one test set. During the training process, the state transition probability distribution ($\mathbf{A}$), observation probability distribution ($\mathbf{B}$), and initial state probability distribution (π) were optimized using the expectation-maximization algorithm [13]. The log-likelihood (LL) was determined both from the TASK and REST session data using the forward-backward algorithm [13]. In HMM, the higher LL value indicates the test data lean towards the specific model (TASK or REST). We performed cross-validation to determine the classification accuracy between the training and test data set. The actual calculation was done using a public toolbox written in Matlab.

2.5. Partial Least Squares-Discriminant Analysis (PLS-DA)

PLS-DA is one of the classification methods and is based on the PLS regression algorithm [27]. PLS regression is a method for modeling a linear relationship between a set of output variables (chemical concentration), $\mathbf{y}$, and a set of input variables (regressors), $\mathbf{X}$, which can be expressed as $\mathbf{y} = \mathbf{Xb} + b_0$. Here $\mathbf{X}$ is an $n \times p$ matrix, containing the values of the predictor values at n data points, $\mathbf{b} = (b_1, \cdots, b_p)^T$ is a p-dimensional column vector which contains regression coefficients; the T denotes the transpose of a vector or matrix. The regression coefficients, b, was defined as

$$\mathbf{b} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y} \tag{1}$$

where

$$\mathbf{b} = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_p \end{bmatrix}, \ \mathbf{X} = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1p} \\ 1 & x_{21} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & \cdots & x_{np} \end{bmatrix}, \ \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \tag{2}$$

In this study, we noted y as TASK (1) and REST (-1), determined the regression model using principal components, and computed classification accuracy using cross-validation. We used the PLS_toolbox (Eigenvector Research, Inc., WA, USA) for analysis.

III. RESULTS & DISCUSSION

In one set of experiments, the TASK session was continued for a whole minute to observe hemodynamic changes (Fig. 3). The maximum peak was observed at the middle of duration (around 30-40 sec). The low p-value (< 0.0001) between oxy-HB and deoxy-HB in the TASK session indicates strong anti-correlation.

Out of 16 channels with 8 sources and two detectors, only the 8 channels connecting the closest source-detector combinations were used for our experiment. Only statistically significant signals were counted for our BCI study. A typical signal fluctuation is shown in Fig. 4. The signal indicates that a rapid reaction from the finger tapping action is reported after 2~5 seconds latency. Also, oxy-Hb started to decrease immediately when the finger tapping ceased.

In Fig. 5, an example of the 8-channel signals is shown. The onset of the TASK session was initiated at time zero in each graph. Only the signals from channels 4 and 5 were consistent with expectations. Signals from other channels
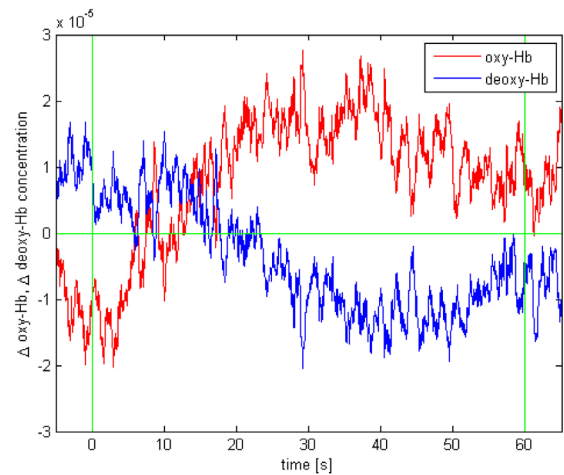

FIG. 3. The anti-phasic nature of oxy- and deoxy-Hb fluctuation during the 1-minute Task session. The red line is the oxy-Hb, and the blue one is the deoxy-Hb. Vertical green lines indicate the start and stop time of the finger tapping.

do not seem to reflect the expected hemodynamic changes. There are two possible reasons for this low success rate; the diffusely reflected light from the cortical area does not pass through the motor cortex of the participant, or a sufficient amount of light was not received on the detector to create significant signals. For many participants in general, the light transmitted through the scalp, skull, and cortex area was
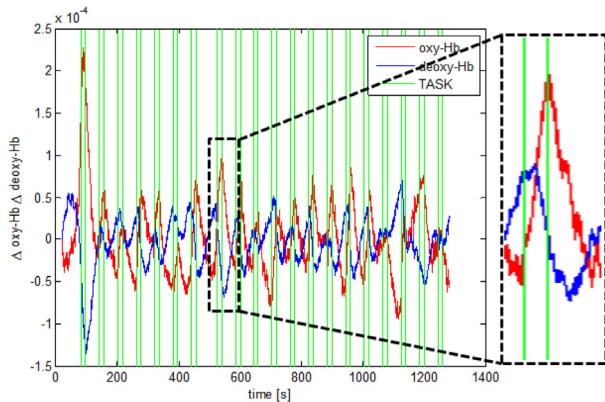

FIG. 4. A typical example of the oxy-Hb and deoxy-Hb fluctuation during a series of TASK and REST sessions.

too weak to generate a signal on the detector. However, approximately 3~5 cm separation between a source and a detector is known to be sufficient to cover the motor cortex region [28, 29]. A prediction model based on HMM and PLS-DA was developed to classify TASK and REST signals from raw oxy-Hb data. The developed model was used to predict the intention of the participants when the oxy-Hb signal with different time segments (15, 10, and 5 sec) was used as input data. The results are shown in Tables 1 and 2 and visualized in Fig 6.

As anticipated, in most cases, the prediction result scored the highest when longer time segments were used. More interestingly, the PLS-DA and HMM-based models produced similar predictions, while PLS-DA created a slightly better result when longer time segments were used. When the 15 sec data sets from the prediction model were compared, the success rate was 90.21% for PLS-DA and 85.78% for HMM. Additionally, Tables 1 and 2 imply that longer time segments do not always guarantee a higher success rate. For the case of HMM results, the success rate for the 10 sec segment (79.76%) was slightly worse than the 5 sec segment success rate (80.29%).

At this stage, it is not fair to conclude that PLS-DA is superior to HMM in constructing the prediction model using
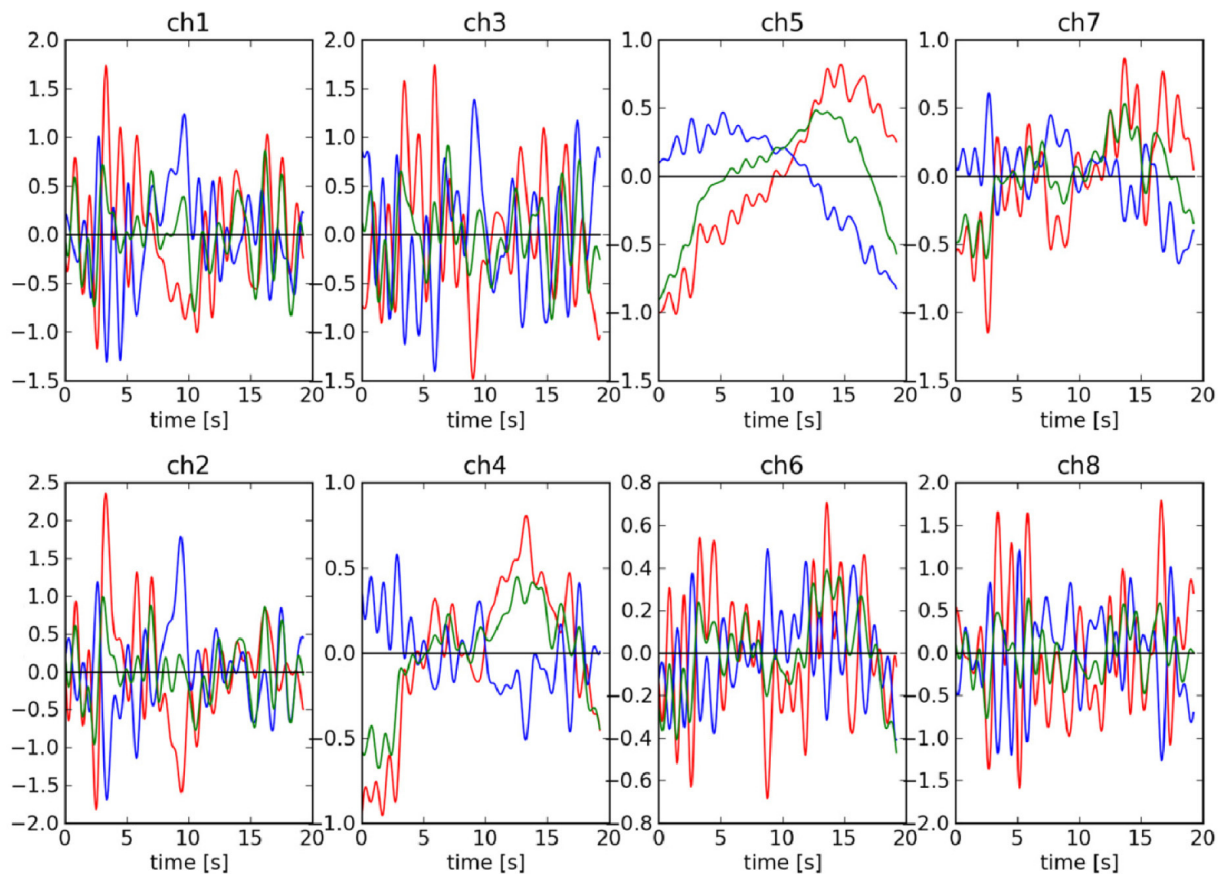

FIG. 5. Hemodynamic signals from all eight channels. Oxy-Hb (red), deoxy-Hb (blue) and total-Hb (green). The units of the y-axis is arbitrary.

TABLE 1. Test results with HMM in 15, 10 and 5 sec data with 7 participants

Participant	Status	15 sec	10 sec	5 sec
Subj. 1	HMM-T	88.46%	81.98%	75.62%
	HMM-R	74.40%	85.58%	71.50%
Subj. 2	HMM-T	96.08%	98.27%	85.18%
	HMM-R	82.96%	66.39%	69.78%
Subj. 3	HMM-T	81.71%	75.30%	98.22%
	HMM-R	78.18%	71.46%	71.53%
Subj. 4	HMM-T	82.96%	90.46%	94.47%
	HMM-R	94.30%	86.13%	87.43%
Subj. 5	HMM-T	83.01%	83.72%	84.24%
	HMM-R	79.54%	78.44%	78.10%
Subj. 6	HMM-T	91.96%	75.56%	75.53%
	HMM-R	86.65%	79.92%	78.27%
Subj. 7	HMM-T	95.80%	72.45%	82.83%
	HMM-R	84.96%	71.00%	71.39%
Average	Oxy-Hb-T	88.57%	82.53%	85.16%
	Oxy-Hb-R	83.00%	76.99%	76.28%
	Total	85.78%	79.76%	80.29%

TABLE 2. Prediction results with PLS-DA in 15, 10 and 5 sec data with 7 participants

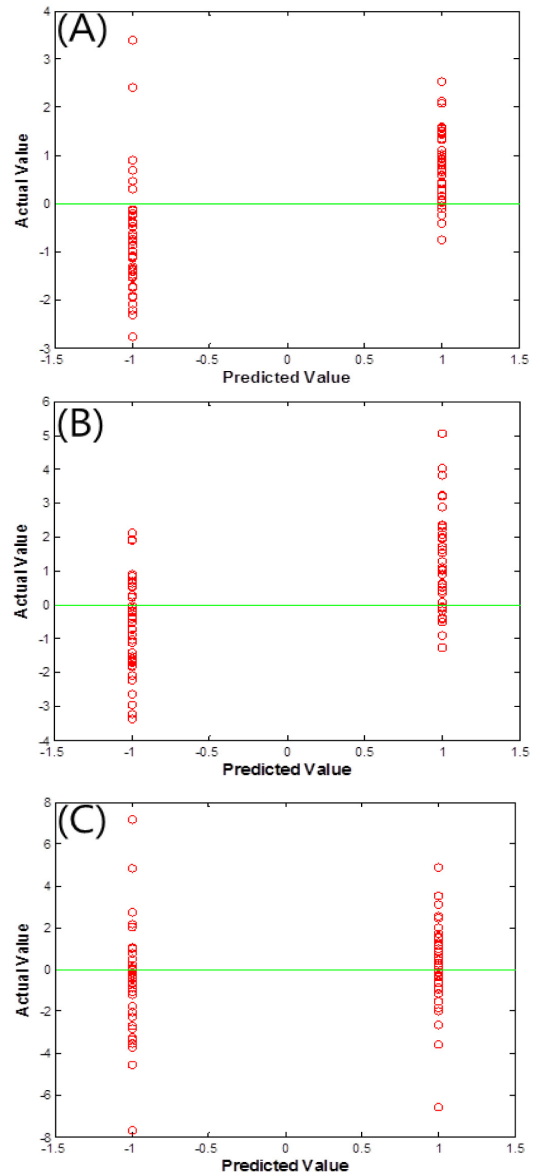| Participant	Status	15 sec	10 sec	5 sec
Subj. 1	Oxy-Hb-T	92.50%	79%	80%
	Oxy-Hb-R	93.70%	90%	73%
Subj. 2	Oxy-Hb-T	94%	82%	79%
	Oxy-Hb-R	90%	88%	71%
Subj. 3	Oxy-Hb-T	84%	95%	75%
	Oxy-Hb-R	95%	64%	72%
Subj. 4	Oxy-Hb-T	82.10%	82%	84.20%
	Oxy-Hb-R	93.70%	84%	71.60%
Subj. 5	Oxy-Hb-T	95%	81%	79%
	Oxy-Hb-R	90%	89%	76%
Subj. 6	Oxy-Hb-T	90%	83%	70%
	Oxy-Hb-R	89%	79%	79%
Subj. 7	Oxy-Hb-T	91%	88%	79%
	Oxy-Hb-R	83%	93%	70%
Average	Oxy-Hb-T	89.80%	84.29%	78.03%
	Oxy-Hb-R	90.63%	83.86%	73.23%
	Total	90.21%	84.07%	75.63%

FIG. 6. Results of the PLS-DA for 15 sec (A, 92.5%), 10 sec (B, 83%), and 5 sec (C, 79%) data. Predicted value 1 indicates the Task and -1 indicates the Rest. The actual values produced using the PLS-DA are classified based on their polarity.

the hemodynamics data. Even if the HMM is frequently used in obtaining time series data pattern recognition, the time segmented data used to construct the model in this study probably was not large enough to develop a robust model. On the other hand, since PLS-DA can obtain the lowest predicted residual sum of squares (PRESS) using iterative methods, such as nonlinear iterative partial least squares (NIPALS) or the simple partial least squares (SIMPLS) algorithm [30], the effect of short time segmentation may not have been as pronounced.

It is well known that the strength of the finger tapping significantly affects the degree of the hemodynamic changes [31]. Therefore, each model used for prediction of the participant's intention using NIRS has to be individually calibrated. In our study, the data from a given participant were used to construct the prediction model for that individual.

Recent BCI studies using NIRS have shown great potential in a number of neuroscience studies, such as a mirror study in the motor cortex [5], a beverage preference test in the prefrontal lobe [20], and a word speller application [22]. Since the hemodynamic signal fluctuation is caused not only by certain specific stimuli but by other natural causes, a robust model based on PLS-DA may give a higher success rate for BCI applications using NIRS.

ACKNOWLEDGMENT

REFERENCES

1. D. A. Boas, T. Gaudette, G. Strangman, X. Cheng, J. J. Marota, and J. B. Mandeville, "The accuracy of near infrared spectroscopy and imaging during focal changes in cerebral hemodynamics," Neuroimage **13**, 76-90 (2001).

2. F. F. Jobsis, "Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters," Science **198**, 1264-1267 (1977).

3. T. Muehlemann, D. Haensse, and M. Wolf, "Wireless miniaturized in-vivo near infrared imaging," Opt. Express **16**, 10323-10330 (2008).

4. T. Muehlemann, N. Reefmann, B. Wechsler, M. Wolf, and L. Gygax, "In vivo functional near-infrared spectroscopy measures mood-modulated cerebral responses to a positive emotional stimulus in sheep," Neuroimage **54**, 1625-1633 (2011).

5. S. Shimada, "Modulation of motor area activity by the outcome for a player during observation of a baseball game," PLoS One **4**, e8034 (2009).

6. K. Tai and T. Chau, "Single-trial classification of NIRS signals during emotional induction tasks: towards a corporeal machine interface," J. Neuroeng. Rehabil. **6**, 39 (2009).

7. J. R. Wolpaw, N. Birbaumer, W. J. Heetderks, D. J. McFarland, P. H. Peckham, G. Schalk, E. Donchin, L. A. Quatrano, C. J. Robinson, and T. M. Vaughan, "Brain-computer interface technology: a review of the first international meeting," IEEE Trans. Rehabil. Eng. **8**, 164-173 (2000).

8. J. J. Daly and J. R. Wolpaw, "Brain-computer interfaces in neurological rehabilitation," Lancet Neurol. **7**, 1032-1043 (2008).

9. F. Lotte, M. Congedo, A. Lecuyer, F. Lamarche, and B. Arnaldi, "A review of classification algorithms for EEG-based brain-computer interfaces," J. Neural Eng. **4**, R1-R13 (2007).

10. P. Brunner, A. L. Ritaccio, J. F. Emrich, H. Bischof, and G. Schalk, "A brain-computer interface using event-related potentials (Erps) and electrocorticographic signals (Ecog) in humans," Epilepsia **50**, 389-389 (2009).

11. N. Weiskopf, K. Mathiak, S. W. Bock, F. Scharnowski, R. Veit, W. Grodd, R. Goebel, and N. Birbaumer, "Principles of a brain-computer interface (BCI) based on real-time functional magnetic resonance imaging (fMRI)," IEEE Trans. Biomed. Eng. **51**, 966-970 (2004).

12. R. Sitaram, A. Caria, R. Veit, T. Gaber, G. Rota, A. Kuebler, and N. Birbaumer, "FMRI brain-computer interface: a tool for neuroscientific research and treatment," Comput. Intell. Neurosci. **2007**, 25487 (2007).

13. R. Sitaram, H. Zhang, C. Guan, M. Thulasidas, Y. Hoshi, A. Ishikawa, K. Shimizu, and N. Birbaumer, "Temporal classification of multichannel near-infrared spectroscopy signals of motor imagery for developing a brain-computer interface,"

14. Neuroimage **34**, 1416-1427 (2007).

15. S. Coyle, T. Ward, C. Markham, and G. McDarby, "On the suitability of near-infrared (NIR) systems for next-generation brain-computer interfaces," Physiol. Meas. **25**, 815-822 (2004).

16. S. M. Coyle, T. E. Ward, and C. M. Markham, "Brain-computer interface using a simplified functional near-infrared spectroscopy system," J. Neural Eng. **4**, 219-226 (2007).

17. N. Roche-Labarbe, B. Zaaimi, P. Berquin, A. Nehlig, R. Grebe, and F. Wallois, "NIRS-measured oxy- and deoxy-hemoglobin changes associated with EEG spike-and-wave discharges in children," Epilepsia **49**, 1871-1880 (2008).

18. S. D. Power, T. H. Falk, and T. Chau, "Classification of prefrontal activity due to mental arithmetic and music imagery using hidden Markov models and frequency domain near-infrared spectroscopy," J. Neural Eng. **7**, 026002 (2010).

19. Z. Ghahramani, "An introduction to hidden Markov models and Bayesian networks," Int. J. Pattern Recogn. **15**, 9-42 (2001).

20. L. R. Rabiner, "A tutorial on hidden Markov-models and selected applications in speech recognition," Proc. IEEE **77**, 257-286 (1989).

21. S. Luu and T. Chau, "Decoding subjective preference from single-trial near-infrared spectroscopy signals," J. Neural Eng. **6**, 016003 (2009).

22. L. Ståhle and S. Wold, "Partial least squares analysis with cross-validation for the two-class problem: a Monte Carlo study," J. Chemometr. **1**, 185-196 (1987).

23. A. R. McIntosh, F. L. Bookstein, J. V. Haxby, and C. L. Grady, "Spatial pattern analysis of functional brain images using partial least squares," Neuroimage **3**, 143-157 (1996).

24. A. R. McIntosh, W. Chau, and A. B. Protzner, "Spatiotemporal analysis of event-related fMRI data using partial least squares," Neuroimage **23**, 764-775 (2004).

25. L. Holper, M. Biallas, and M. Wolf, "Task complexity relates to activation of cortical motor areas during uni- and bimanual performance: a functional NIRS study," Neuroimage **46**, 1105-1113 (2009).

26. R. W. Homan, J. Herman, and P. Purdy, "Cerebral location of international 10-20 system electrode placement," Electroencephalogr. Clin. Neurophysiol. **66**, 376-382 (1987).

27. G. Strangman, D. A. Boas, and J. P. Sutton, "Non-invasive neuroimaging using near-infrared light," Biol. Psychiat. **52**, 679-693 (2002).

28. A. Krishnan, L. J. Williams, A. R. McIntosh, and H. Abdi, "Partial least squares (PLS) methods for neuroimaging: a tutorial and review," Neuroimage **56**, 455-475 (2011).

29. M. A. Franceschini, D. K. Joseph, T. J. Huppert, S. G. Diamond, and D. A. Boas, "Diffuse optical imaging of the whole head," J. Biomed. Opt. **11**, 054007 (2006).

30. M. A. Franceschini and D. A. Boas, "Noninvasive measurement of neuronal activity with near-infrared optical imaging," Neuroimage **21**, 372-386 (2004).

31. S. Wold, M. Sjostrom, and L. Eriksson, "PLS-regression: a basic tool of chemometrics," Chemometr. Intell. Lab. **58**, 109-130 (2001).

32. I. Nambu, R. Osu, M. A. Sato, S. Ando, M. Kawato, and E. Naito, "Single-trial reconstruction of finger-pinch forces from human motor-cortical activation measured by near-infrared spectroscopy (NIRS)," Neuroimage **47**, 628-637 (2009).