

Toward more intuitive brain– computer interfacing: classification of binary covert intentions using functional near-infrared spectroscopy

Han-Jeong Hwang
Han Choi
Jeong-Youn Kim
Won-Du Chang
Do-Won Kim
Kiwoong Kim
Sungho Jo
Chang-Hwan Im

Toward more intuitive brain–computer interfacing: classification of binary covert intentions using functional near-infrared spectroscopy

Han-Jeong Hwang,^a Han Choi,^b Jeong-Youn Kim,^b Won-Du Chang,^b Do-Won Kim,^{b,c} Kiwoong Kim,^d Sungho Jo,^e and Chang-Hwan Im^{b,*}

^aKumoh National Institute of Technology, Department of Medical IT Convergence Engineering, 61 Daehak-ro, Gumi, Gyeongbuk 730-701, Republic of Korea

^bHanyang University, Department of Biomedical Engineering, 222 Wangsimni-ro, Seongdong-gu, Seoul 04763, Republic of Korea

^cBerlin Institute of Technology, Machine Learning Group, Marchstraße 23, 10587 Berlin, Germany

^dKorea Research Institute of Standard and Science, 267 Gajeong-ro, Yuseong-gu, Daejeon 34113, Republic of Korea

^eKorea Advanced Institute of Science and Technology, Department of Computer Science, Daehak-ro, Yuseong-gu, Daejeon 34141, Republic of Korea

Abstract. In traditional brain–computer interface (BCI) studies, binary communication systems have generally been implemented using two mental tasks arbitrarily assigned to “yes” or “no” intentions (e.g., mental arithmetic calculation for “yes”). A recent pilot study performed with one paralyzed patient showed the possibility of a more intuitive paradigm for binary BCI communications, in which the patient’s internal yes/no intentions were directly decoded from functional near-infrared spectroscopy (fNIRS). We investigated whether such an “fNIRS-based direct intention decoding” paradigm can be reliably used for practical BCI communications. Eight healthy subjects participated in this study, and each participant was administered 70 disjunctive questions. Brain hemodynamic responses were recorded using a multichannel fNIRS device, while the participants were internally expressing “yes” or “no” intentions to each question. Different feature types, feature numbers, and time window sizes were tested to investigate optimal conditions for classifying the internal binary intentions. About 75% of the answers were correctly classified when the individual best feature set was employed ($75.89\% \pm 1.39$ and $74.08\% \pm 2.87$ for oxygenated and deoxygenated hemoglobin responses, respectively), which was significantly higher than a random chance level (68.57% for $p < 0.001$). The kurtosis feature showed the highest mean classification accuracy among all feature types. The grand-averaged hemodynamic responses showed that wide brain regions are associated with the processing of binary implicit intentions. Our experimental results demonstrated that direct decoding of internal binary intention has the potential to be used for implementing more intuitive and user-friendly communication systems for patients with motor disabilities. © 2016 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.JBO.21.9.091303]

Keywords: brain–computer interface; functional near-infrared spectroscopy; binary communication; intuitive paradigm; covert intentions; neurological diseases.

Paper 150751SSRR received Nov. 5, 2015; accepted for publication Mar. 7, 2016; published online Apr. 5, 2016.

1 Introduction

There are a number of individuals suffering from neurological disorders such as amyotrophic lateral sclerosis, multiple sclerosis, and brainstem stroke. They progressively lose their ability to regulate their neuromuscular systems, and consequently have difficulties in expressing their thoughts or intentions through the brain’s normal output pathways. To provide those who are severely paralyzed with an alternative communication channel, brain–computer interfaces (BCIs) have been intensively studied for the last few decades. BCI is a nonmuscular communication method that translates brain signals into predefined commands for controlling external devices or communicating with others. To implement BCI systems, various neuroimaging modalities have been employed, such as electroencephalography (EEG), functional near-infrared spectroscopy (fNIRS), magnetoencephalography, and functional magnetic resonance imaging (fMRI), among which EEG has been most frequently used due to its reasonable price and portability.¹

Recently, fNIRS has attracted growing interest from BCI researchers, as it is less susceptible to environmental noises and electrophysiological artifacts (e.g., electro-oculogram and muscle activity) compared to EEG.² A variety of mental tasks have been used for implementing fNIRS-based BCI systems, e.g., motor imagery, mental arithmetic, music imagery, mental counting, and object rotation.² So far, a large number of fNIRS-based BCI studies^{3–10} have demonstrated that different mental tasks can be classified with accuracies high enough to be used for practical BCI communication systems (at least 70% accuracy is necessary for practical binary communications).¹¹

A representative BCI application based on the mental-imagery–based BCI paradigms is a binary yes/no communication system, which is simple, but very useful for patients with severe locked-in syndrome (LIS). In general scenarios of fNIRS-based binary communication, two different mental tasks that elicit distinct brain hemodynamic responses (e.g., left- and right-hand motor imagery) are selected. While the user of the BCI system performs one of two mental tasks, the pretrained

*Address all correspondence to: Chang-Hwan Im, E-mail: ich@hanyang.ac.kr

classifier decides which mental task the user is currently performing. Since each of the two mental tasks is assigned to either a “yes” or “no” intention, the user can express his or her binary intentions by simply performing one of the two designated mental imagery tasks.

Several previous studies showed the possibility of applying the mental imagery paradigm to the development of binary BCI systems,^{12–14} but the arbitrary task-to-intention mapping has inherent limitations. It generally requires short-term working memory resources to imagine given mental tasks.¹⁵ In addition, user training periods are necessary to generate reliable task-specific brain patterns.^{16–18} Most importantly, it has been frequently reported from EEG-based BCI studies that some subjects called “BCI illiterate” cannot elicit consistent brain activity patterns for specific mental tasks even after extensive training (the portion is about 15% to 30%, depending on the tasks).^{19,20}

On the other hand, a recent pilot study showed that it might be possible to decode internal yes/no responses to disjunctive questions from fNIRS signals.²¹ The “direct intention decoding” paradigm was applied to a patient with complete LIS, and a meaningful classification accuracy (>70%) was reported. This paradigm could be used as a promising alternative to the conventional paradigms for developing a more intuitive and user-friendly binary BCI system; however, since the direct intention decoding paradigm was tested with only one subject in the previous study, additional validation studies are needed to further confirm its feasibility and reliability.

In this study, motivated by the previous pilot study that decoded internal yes/no intentions from brain hemodynamic responses,²¹ we investigated whether this direct intention decoding paradigm can be reliably used for practical BCI communications. To this end, we replicated an experimental paradigm similar to the previous study with eight healthy subjects. In the experiments, the participants were given 70 disjunctive questions through a set of speakers, and asked to internally answer “yes” or “no,” during which the hemodynamic responses of the brain were measured using a multichannel fNIRS device. The recorded hemodynamic responses were used for discriminating “yes” and “no” covert intentions, where different feature types, numbers of features, and analysis time periods were tested to investigate optimal classification conditions.

2 Method

2.1 Participants

Eight participants (all males and right handed, 23 to 30 years of age) took part in this study. None of them had a previous history of neurological or psychiatric disorders that might affect the experimental results. The general concept of BCI technology, our research goal, and experimental procedures were sequentially explained to the participants before the experiments, and they provided written informed consent. They were given monetary reimbursement for their participation in the experiments. Ethical approval was obtained from the Institutional Review Board of Hanyang University.

2.2 Disjunctive Questions

In the previous reference study by Gallegos-Ayala et al.,²¹ two types of questions with known and unknown answers were used, where mostly questions with known answers were tested. In this study, we only considered questions with unknown answers to

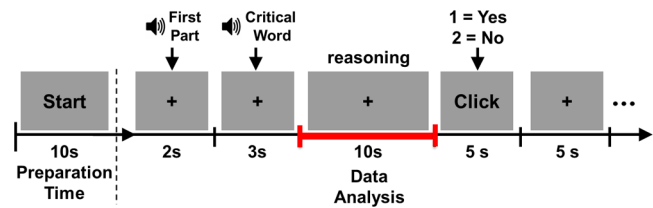


Fig. 1 Experimental paradigm used for recording task-related hemodynamic responses.

reflect realistic communication situations in daily life. Seventy disjunctive questions that were grouped into five different classes were prepared for this study (see Appendix). All of the questions asked about the participant’s current emotional state and/or personal tastes in various objects, e.g., “Are you thirsty?,” “Do you like cats?,” “Do you want to go to the Amazon?,” and “Do you want to get a fountain pen?.” Each question was divided into two parts, and the second part included a critical word (e.g., “Are you” + “thirsty?,” “Do you want to go to” + “Amazon”) that determines the answer. All of the questions were translated into Korean to help the native Korean participants understand the meaning of the questions without a delay.

2.3 Experimental Paradigm

Before the experiments, the participants were asked to sit in a comfortable armchair facing a 17-in. LCD monitor placed at a 1-m distance from the participants. Figure 1 shows the experimental paradigm used in this study. In the beginning of the experiment, a preparation time of 10 s was given to the participants, when they were instructed not to move their body and concentrate only on the upcoming question. Following this, the first part of a disjunctive question was presented for 2 s (e.g., “Are you”) and the critical word was presented for 3 s (e.g., “hungry?”) through two loud speakers placed at both sides of the monitor. The durations of the time periods (2 and 3 s) were determined considering the general lengths of the questions translated into Korean. The participants were asked to internally answer either “yes” or “no” for 10 s right after listening to a question, while they were asked to stare at a black fixation cross located at the center of the monitor. Each participant was asked to press either the “yes” button or the “no” button on a response pad when the word “Click” appeared on the monitor. All of the participants generally responded within 1 to 2 s. After a 5-s break, the same procedure was repeated until all 70 questions were presented. The numbers of “yes” and “no” responses were different among the subjects, as they responded based on their emotional state and personal tastes, but the mean ratio of “yes” and “no” responses was almost identical (0.502 and 0.498, respectively), as shown in Table 1.

2.4 fNIRS Data Acquisition

During the experiments, task-related cerebral hemodynamic responses were recorded using a multichannel near-infrared spectroscopy (NIRS) imaging system (FOIRE-3000, Shimadzu Co. Ltd., Kyoto, Japan) in a dimly lit, soundproof room. Figure 2 shows the NIRS optode distribution and the channel location. For data recording, 31 optical fiber probes, consisting of 16 sources and 15 detectors, were attached to the subject’s scalp. The light sources emitted three different wavelengths of 780, 805, and 830 nm. The central optode was placed on

Table 1 Numbers of yes and no responses, and their respective ratios for each subject.

Subject	Num. Yes:No	Ratio Yes:No
S1	40:30	0.57:0.43
S2	41:29	0.59:0.41
S3	39:31	0.56:0.44
S4	30:40	0.43:0.57
S5	31:39	0.44:0.56
S6	23:47	0.33:0.67
S7	31:39	0.44:0.56
S8	46:24	0.66:0.34
Mean	35.125:34.875	0.502:0.498

Cz according to the international 10-20 EEG electrode system, as shown in Fig. 2. The distance between adjacent optodes was set to 3 cm because it has been shown in previous fNIRS studies that this interoptode distance is ideal for measuring cortical hemodynamic response changes.²² Since all subjects had black hair that might have decreased the signal-to-noise ratio of the NIRS signals, we carefully adjusted their hair during NIRS preparation until the measurement intensity level became

acceptable. We used the Automatic Adjustment Function in the measurement software provided by the company. In particular, the negative high-voltage values of the photomultiplier were appropriately set between 100 and 1000 V, and measurement intensity was thereby set between 0.05 and 4 V for all three wavelengths when input intensity was between 0 and 10 V. The task-related light intensity changes were recorded from 50 scalp locations with a sampling rate of 10 Hz. The recorded light intensities were then converted into the concentration changes of oxygenated and deoxygenated hemoglobin ([oxy-Hb] and [deoxy-Hb], respectively) using the modified Beer–Lambert law before data analysis (extinction coefficients of 780-, 805-, and 830-nm wavelengths: 0.7360, 0.8973, and 1.0507 for [oxy-Hb], and 1.1050, 0.8146, and 0.7804 for [deoxy-Hb]; differential pathlength factor: 1 for all three wavelengths).²³ The NIRS optode configuration used in this study fully covered the sensorimotor area and partly covered the temporal area, which are the brain regions employed in the previous study that first attempted to classify internal yes/no responses using fNIRS signals.²¹ (Pre)frontal and occipital areas were also partly covered by our optode configuration. This wider coverage of brain regions can allow for investigating brain areas relevant to the processing of internal yes/no intentions.

2.5 fNIRS Data Analysis

Figure 3 shows the pipeline of the data processing, consisting of preprocessing, feature extraction, feature selection, and classification. The following subsections provide detailed descriptions of each analysis step.

2.5.1 Preprocessing

A common average reference (CAR) spatial filter was first applied to the [oxy-Hb] and [deoxy-Hb] responses to remove common noise components (e.g., effect of scalp blood flow changes), for which the mean hemoglobin values of all channels were calculated and then subtracted from the hemoglobin values of each channel for every time point. The CAR method, commonly used in EEG-based BCIs, has shown good performance for reducing common noise components,^{24,25} and has also been successfully applied to fNIRS data analyses.^{26–28} After the application of the CAR filter, each hemoglobin response was filtered using a bandpass filter with cutoff frequencies of 0.01 and 0.09 Hz using a fourth-order Butterworth zero-phase filter in order to filter out spontaneously generated physiological components known to originate from Mayer waves (approximately 0.1 Hz), heart rate (1.6 to 1.8 Hz), and breathing (0.2 to 0.3 Hz). Also, the bandpass filtering removed linear trend of NIRS signals by filtering out NIRS signals below 0.01 Hz. The frequency band of 0.01 to 0.09 Hz was selected based on the previous fNIRS-based BCI studies.^{6,29}

2.5.2 Feature extraction

For the feature extraction, we used 10-s epochs measured while the participants were covertly answering a given question (see Fig. 1). Since task-related hemodynamic responses generally appear with a varying delay of 3 to 8 s,^{4,5,9,30} different time windows were used to extract candidate features. Each time window was defined by a start time ranging from 1 to 6 s and an end time spanning from 5 to 10 s. The increment time was set to 1 s for both start and end times. All possible combinations of start

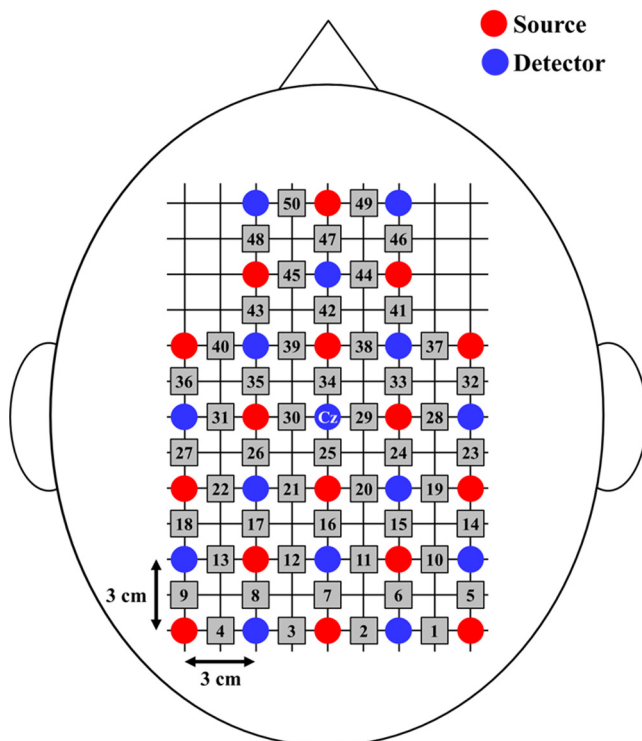


Fig. 2 Arrangement of 31 optical fiber probes and 50 NIR channels. The red and blue circles represent the 16 sources and the 15 detectors, respectively. The gray numbered squares indicate the 50 NIR channels. The distance between adjacent optodes is 3 cm. The Cz position based on the 10-20 international system is set as the center of this configuration.

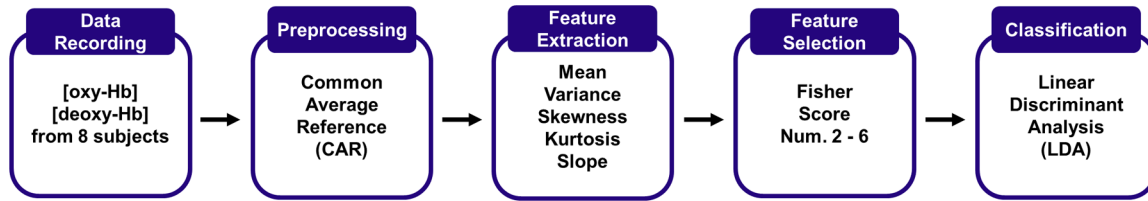


Fig. 3 Flowchart showing the fNIRS data acquisition and four data analysis steps consisting of preprocessing, feature extraction, feature selection, and classification.

and end times were then used to construct different time windows.^{5,9,31} Based on the previous fNIRS-based BCI studies,^{5,9,31,32} the following five different feature types were extracted from each time window's data to investigate the influence of feature types on the classification performance:

1. Mean: Averaged signal amplitude over the time window

$$\bar{x} = \frac{1}{b-a+1} \sum_{t=a}^b x_t. \quad (1)$$

2. Variance: A measure of signal spread

$$\sigma^2 = \frac{1}{b-a+1} \sum_{t=a}^b (x_t - \bar{x})^2. \quad (2)$$

3. Skewness: A measure of the asymmetry of the probability distribution of signal values relative to its mean

$$s_1 = \frac{\frac{1}{b-a+1} \sum_{t=a}^b (x_t - \bar{x})^3}{\left(\sqrt{\frac{1}{b-a+1} \sum_{t=a}^b (x_t - \bar{x})^2} \right)^3}. \quad (3)$$

4. Kurtosis: A measure of the peakedness of the probability distribution of signal values relative to its mean

$$k = \frac{\frac{1}{b-a+1} \sum_{t=a}^b (x_t - \bar{x})^4}{\left(\sqrt{\frac{1}{b-a+1} \sum_{t=a}^b (x_t - \bar{x})^2} \right)^2}. \quad (4)$$

5. Slope: A measure of the steepness of signal values

$$s_2 = \frac{x_b - x_a}{b-a}, \quad (5)$$

where x_t is a hemodynamic response signal value at a time point t , and a and b are the first and last time points in each time window, respectively.

The candidate features were independently evaluated for all possible combinations of the time windows, 50 channels, and the five feature types for the [oxy-Hb] and [deoxy-Hb] responses recorded during 70 trials (70 different questions).

2.5.3 Feature selection

A large number of features might cause overfitting of a classifier constructed from the training data, thus feature selection is an

essential step in pattern classification problems based on machine learning techniques. In this study, we used the Fisher score, which has been frequently used in fNIRS-based BCI literature and has shown high performance.^{6,7,9,33} The Fisher scores were calculated for all elements of the constructed feature set

$$FS_i = \frac{(\mu_{\text{yes}} - \mu_{\text{no}})^2}{v_{\text{yes}} + v_{\text{no}}}, \quad (6)$$

where μ and v are the mean and variance, respectively, of the designated class. The subscript i represents the i 'th feature element. Since the Fisher score is defined as the ratio of the difference of the means of features in two classes to the sum of variances of features in the same class, a higher Fisher score guarantees larger separability between different classes. Thus, the best feature subset is generally constructed by selecting the top N features having the highest Fisher scores. As the optimal number of features varies among subjects as well as feature sets, different numbers of features ranging from 2 to 6 were independently tested for each feature set, as in the previous fNIRS-based BCI study,⁹ which showed little change in the classification accuracy with more than six features.

2.5.4 Classification

The leave-one-out cross-validation (LOOCV) method was used to evaluate the performance of classifying covert binary intentions. The LOOCV method leaves one trial as the test data and constructs a classifier using the other trials (training data set). The one trial left is then tested using the constructed classifier, and this procedure is iterated until every trial is used as the test data. The numbers of correct and incorrect predictions are counted during LOOCV, and the classification accuracy is then estimated by dividing the number of correct predictions by the total number of trials. In this study, a classification accuracy was calculated separately for "yes" trials (sensitivity) and "no" trials (specificity), then an adjusted accuracy was evaluated by averaging sensitivity and specificity to accounts for the unbalanced numbers of "yes" and "no" trials.³⁴ The LOOCV method was separately applied to all possible [oxy-Hb] and [deoxy-Hb] feature sets for each subject. To avoid potential biases in estimating classification accuracy, feature selection was independently performed using the Fisher score in every cross-validation step; thereby features selected for constructing a classifier varied among the 70 cross-validations. Linear discriminant analysis, which was used as a classifier, has been successfully employed in many fNIRS-based BCI studies.^{4-6,9,32,33,35}

Table 2 Classification accuracies of each subject attained using the best feature set for [oxy-Hb] and [deoxy-Hb] responses. Selected features (channels) were different in each cross-validation step because training data were continuously changed during LOOCV. The six channels most frequently selected during the whole LOOCV steps over all subjects and their selection ratios are as follows: ch. 14 (20.22%), 16 (18.42%), 6 (18.17%), 50 (18.02%), 38 (17.09%), and 25 (16.81%) for [oxy-Hb] responses, and ch. 14 (20.19%), 4 (19.4%), 33 (18.94%), 32 (18.72%), 15 (18.54%), and 16 (18.27%) for [deoxy-Hb] responses, respectively.

Subject	Oxy-Hb						Deoxy-Hb					
	Number of selected features	Time period (s)	Feature type	Sensitivity (%)	Specificity (%)	Adjusted accuracy (%)	Number of selected features	Time period (s)	Feature type	Sensitivity (%)	Specificity (%)	Adjusted accuracy (%)
S1	3	2 to 10	Slope	75	80	77.50	4	0 to 5	Slope	80	73.33	76.67
S2	3	4 to 10	Kurtosis	72.86	74.74	73.80	2	3 to 10	Skewness	78.5	75.41	76.95
S3	2	1 to 10	Kurtosis	58.97	90.33	74.64	6	5 to 10	Kurtosis	79.49	67.74	73.61
S4	5	1 to 9	Kurtosis	66.67	82.50	74.58	6	1 to 6	Kurtosis	73.33	80	76.67
S5	4	3 to 8	Skewness	80.65	71.79	76.22	6	4 to 9	Skewness	70.97	71.79	71.38
S6	4	1 to 7	Kurtosis	73.91	78.73	76.31	6	0 to 6	Kurtosis	60.87	76.60	68.73
S7	2	2 to 10	Kurtosis	70.97	82.05	76.51	2	2 to 9	Kurtosis	70.97	76.92	73.94
S8	5	0 to 5	Kurtosis	71.74	83.33	77.53	2	0 to 7	Variance	60.87	87.50	74.18
Mean	3.63	1.8 to 8.6		71.34% ±6.37	80.43% ±5.65	75.89% ±1.39	4.25	1.9 to 7.8		71.87% ±7.69	76.16% ±5.88	74.08% ±2.87

3 Results

Table 2 shows the best classification accuracies of each subject for [oxy-Hb] and [deoxy-Hb] responses, and their corresponding optimal feature sets. For example, the best classification accuracy (77.50%) of subject S1 was obtained from the feature set constructed using slope features in a time period of 2 to 10 s, when the number of selected features was three. All subjects showed significantly higher classification accuracies than a random chance level (68.57% for $p < 0.001$), which are also higher than the marginal classification accuracy of 70% that determines whether a binary BCI system can be used for practical communication.¹¹ The mean classification accuracies of the [oxy-Hb] and [deoxy-Hb] features were almost identical: 75.89% and 74.08%, respectively. The true positive (sensitivity)

and true negative (specificity) prediction rates were 71.34% and 80.43% for the [oxy-Hb] features and 71.87% and 76.16% for the [deoxy-Hb] features, respectively.

Figure 4 shows the best classification accuracies of each subject for the [oxy-Hb] and [deoxy-Hb] responses with respect to the number of selected features. The best mean accuracies of the [oxy-Hb] and [deoxy-Hb] responses were obtained when five and two features were selected, respectively, but the number of selected features did not significantly affect the classification accuracy [Friedman: $\chi^2(4) = 1.48$, $p = 0.83$ for the [oxy-Hb] responses; Friedman: $\chi^2(4) = 3.3$, $p = 0.51$ for the [deoxy-Hb] responses].

Figure 5 shows the optimal analysis time periods of each subject in terms of the classification performance. Each horizontal

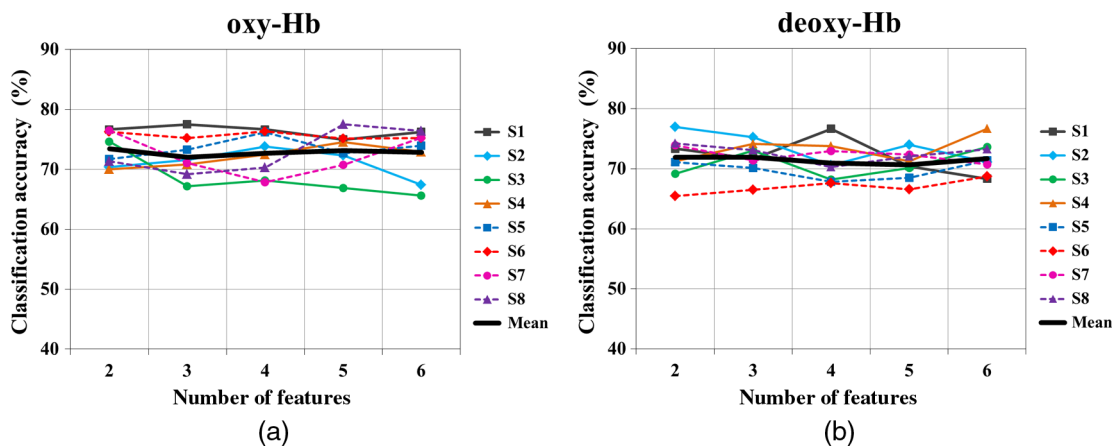


Fig. 4 Best classification accuracies of each subject with respect to the number of features for (a) [oxy-Hb] responses and (b) [deoxy-Hb] responses.

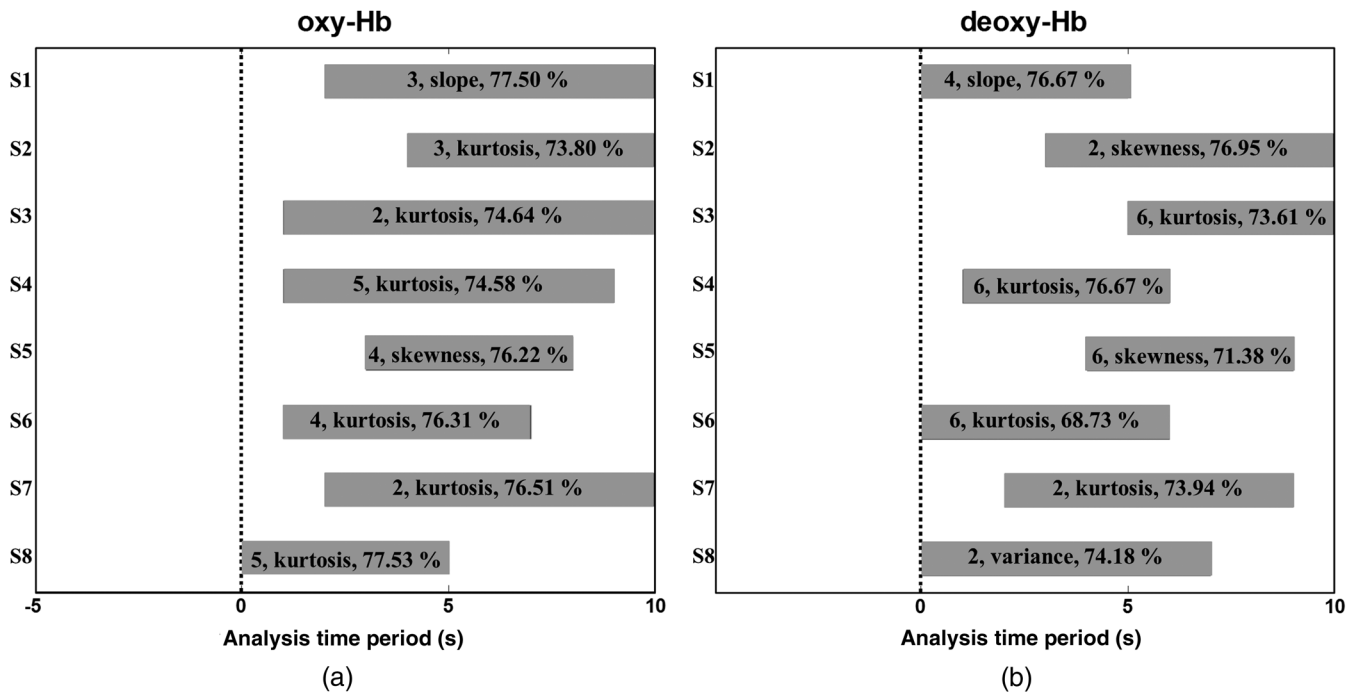


Fig. 5 Optimal analysis time periods of each subject for (a) [oxy-Hb] responses and (b) [deoxy-Hb] responses. A binary question is presented between $t = -5$ and $t = 0$, and the subject starts an internal answer at $t = 0$ for 10 s.

bar represents the best analysis time window, within which the number of selected features, feature type, and classification accuracy are shown. The best analysis time periods were highly dependent on the subjects, and no general trend could be found.

Figure 6 shows the best classification accuracies of each subject with respect to feature type. The kurtosis feature generally showed a higher mean accuracy than the others for both [oxy-

Hb] and [deoxy-Hb] responses. In particular, the kurtosis feature showed a statistically higher performance than the mean and variance features when the features were extracted from [oxy-Hb] responses [Friedman: $\chi^2(4) = 16.5$, $p = 0.0024$; the Bonferroni post hoc analysis: kurtosis > mean = variance, corrected $p < 0.05$].

Figures 7 and 8 show the grand [oxy-Hb] and [deoxy-Hb] responses averaged over all subjects, respectively. The time

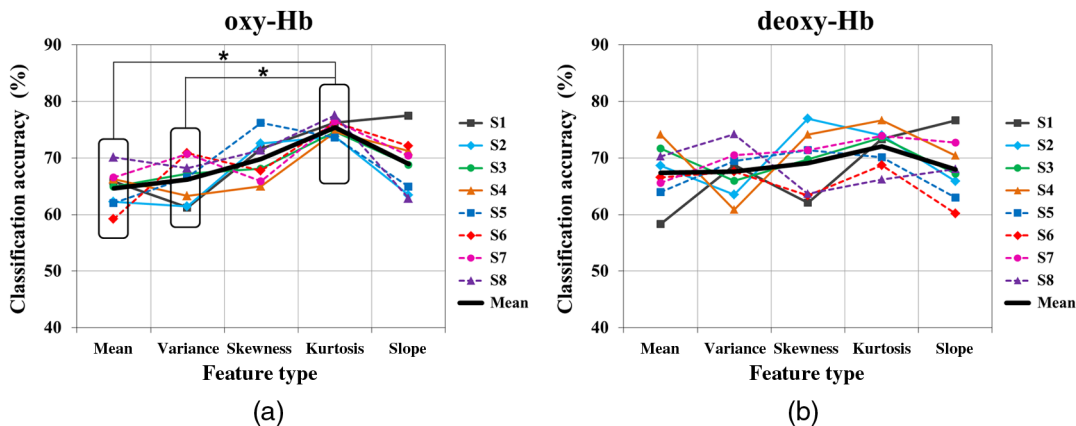


Fig. 6 Best classification accuracies of each subject with respect to the feature type for (a) [oxy-Hb] responses and (b) [deoxy-Hb] responses. The asterisk represents a statistically significant difference between the corresponding pair of feature types in terms of classification accuracy. The kurtosis feature generally showed a higher mean accuracy than the others for both (a) [oxy-Hb] and (b) [deoxy-Hb] responses. Detailed information on kurtosis feature sets used to obtain the best classification accuracies for (a) [oxy-Hb] and (b) [deoxy-Hb] responses are presented in Table 2 for the subjects S2–S4 and S6–S8, and for S3–S4 and S6–S7, respectively. The following is the detailed information on kurtosis feature sets used for the other subjects: (a) [oxy-Hb] responses—S1: six features and a time period of 0–10 s; S5: four features and a time period of 1–9 s; (b) [deoxy-Hb] responses—S1: two features and a time period of 1–8 s; S2: five features and a time period of 1–9 s; S5: three features and a time period of 1–9 s; S8: five features and a time period of 0–10 s.

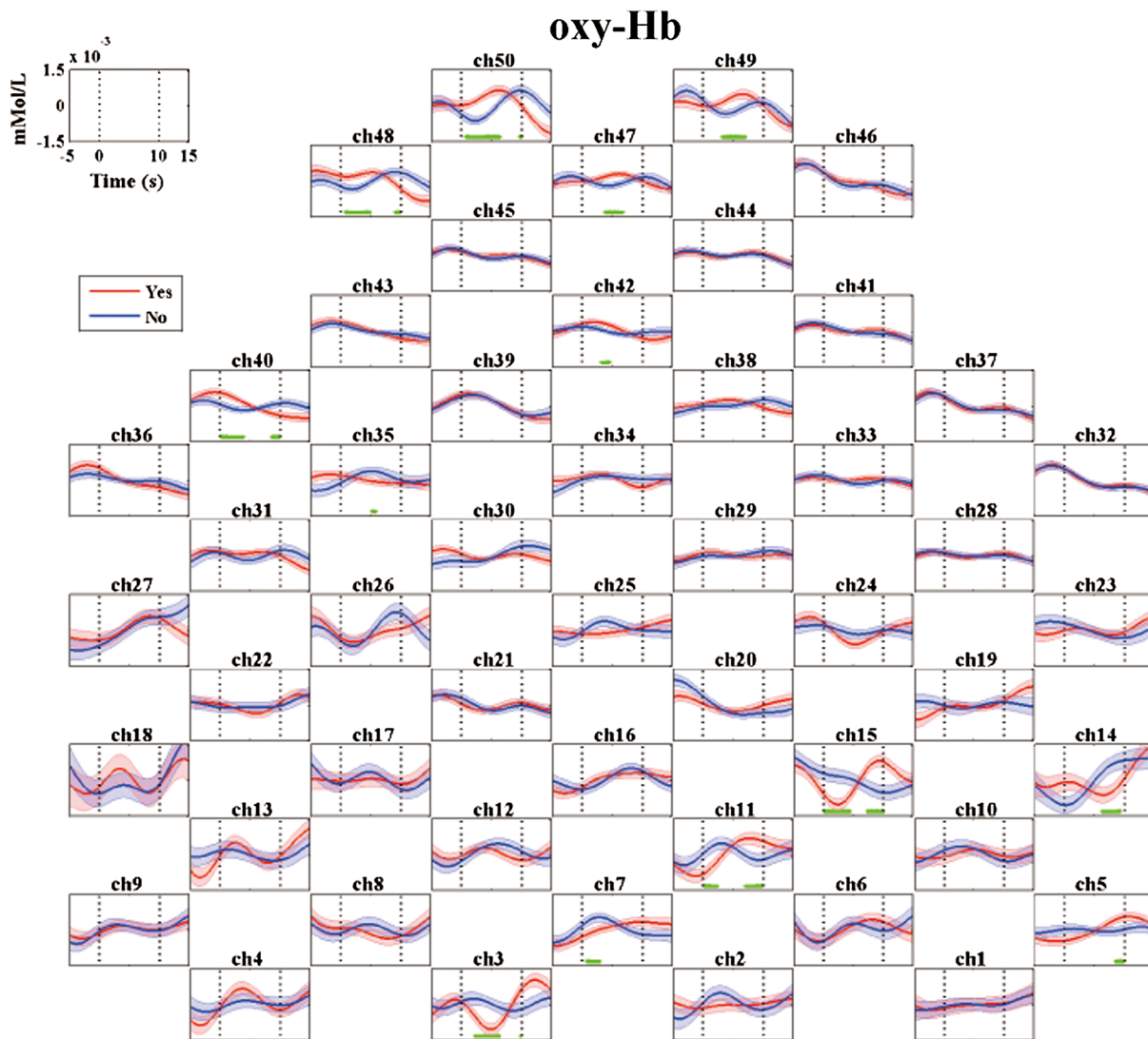


Fig. 7 Grand-averaged [oxy-Hb] responses recorded during covert “yes” and “no” responses for all 50 channels. The shaded regions indicate the standard errors computed across all subjects for each “yes” and “no” condition. The green bars on the bottom of each panel represent time periods showing statistically significant difference of [oxy-Hb] responses between “yes” and “no” conditions (paired *t*-test; $p < 0.05$). A binary question is presented between $t = -5$ and $t = 0$, and the subject starts an internal answer at $t = 0$ for 10 s. The grand-averaged [oxy-Hb] responses were obtained by normalizing preprocessed [oxy-Hb] responses based on single-trial, for which we calculated the mean of each trial, subtracted it from all data points of the corresponding trial, and averaged all the normalized trials across all subjects.

periods showing a statistically significant difference of hemoglobin responses between “yes” and “no” conditions are depicted by green lines on the bottom of each panel. As seen in the figures, the [oxy-Hb] responses generally increase for “yes” responses and decrease for “no” responses, especially in prefrontal areas (ch. 47 to 50), showing many statistically significant time periods. The inverted trend is observed over the parieto-occipital regions (ch. 1 to 26), but both intentions do not generally induce remarkable changes around the frontal cortex (ch. 28 to 45). On the other hand, the [deoxy-Hb] responses are almost reversed compared to the [oxy-Hb] responses, except that the difference between the two conditions is relatively small around the prefrontal areas (ch. 46 to 50).

Figure 9 depicts the distribution of the averaged Fisher scores evaluated for every channel, from which the most important channels can be identified. The channels over the parieto-occipital

areas show high Fisher scores for both [oxy-Hb] and [deoxy-Hb] responses. The Fisher scores for the [oxy-Hb] responses are higher than those for the [deoxy-Hb] responses over the prefrontal regions. Since the Fisher score represents the degree of difference between two conditions, these results coincide well with the grand-averaged hemodynamic responses shown in Figs. 7 and 8.

4 Discussion

In classical fNIRS-based BCI studies, binary communication systems have generally been implemented by assigning two distinct mental imagery tasks to “yes” or “no” answers. For the implementation, so far, more than 10 mental tasks have been adopted.² In this study, following the previous pilot study that tried to directly decode internal yes/no thoughts through fNIRS,²¹ we investigated whether the direct intention decoding

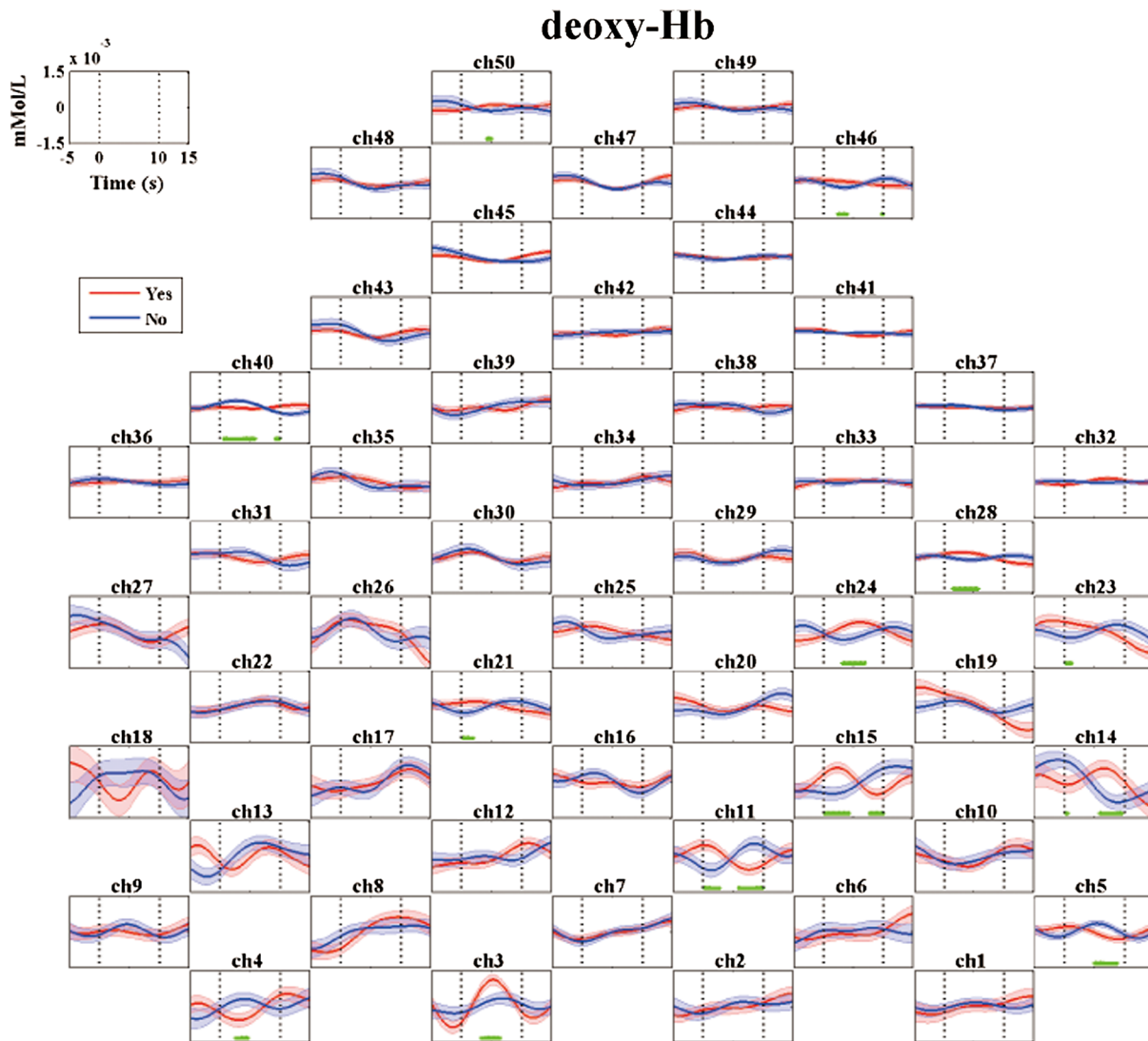


Fig. 8 Grand-averaged [deoxy-Hb] responses recorded during covert “yes” and “no” responses for all 50 channels. The shaded regions indicate the standard errors computed across all subjects for each “yes” and “no” condition. The green bars on the bottom of each panel show statistically significant difference of [deoxy-Hb] responses between “yes” and “no” condition at the corresponding time periods (paired t -test; $p < 0.05$). A binary question is presented between $t = -5$ and $t = 0$, and the subject starts an internal answer at $t = 0$ for 10 s. The grand-averaged [deoxy-Hb] responses were obtained by normalizing pre-processed [deoxy-Hb] responses based on single-trial, for which we calculated the mean of each trial, subtracted it from all data points of the corresponding trial, and averaged all normalized trials across all subjects.

paradigm can be reliably used for binary communications. Our experimental results demonstrated that individual covert binary intentions could be reliably classified using fNIRS signals with classification accuracy high enough to be used for binary communication applications ($>70\%$).¹¹

As discussed in a study by Luu and Chau,³⁵ the traditional task-based BCI paradigm was neither intuitive nor user-friendly for binary communication due to indirect task-to-intention mapping. Compared to the traditional BCI paradigm, the main advantage of the direct intention decoding paradigm is that the user can naturally make yes/no binary decisions based on current internal states, thereby requiring less cognitive load than the conventional one. Thus, it is highly expected that the paradigm based on direct decoding of user intention can contribute to the development of more intuitive binary communication systems. We also expect that this paradigm can be used as an alternative

to the existing mental-task-based BCI paradigms, when patients cannot produce certain task-specific brain activities, generally known as “BCI illiteracy.”^{19,20} Further studies would be needed with more subjects not only to generalize our results, but also to enhance the overall performance (e.g., information transfer rate) for the user-friendly BCI paradigm to be used for practical communication applications. In future studies, development of general classification models that can be constructed using existing training data need to be considered for those who are completely locked-in, because building an individualized classification model might not be always available for some of them.

Interestingly, the kurtosis feature showed statistically higher accuracy than the mean and variance features in [oxy-Hb] results, both of which have been most frequently used in traditional fNIRS-based BCI studies.² Similar trends were also observed in [deoxy-Hb] results, but there was no statistically

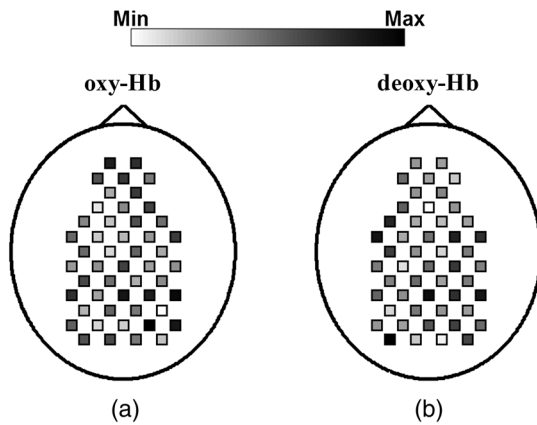


Fig. 9 Grand-averaged gradient maps of the Fisher scores of each channel for (a) [oxy-Hb] responses and (b) [deoxy-Hb] responses. The Fisher scores were calculated by taking into account all possible feature sets used for classification, and normalized by the maximum value of all channels separately for [oxy-Hb] and [deoxy-Hb] responses.

significant difference among different feature types. Note that the previous reference study used only the “mean” feature.²¹

It was observed from Figs. 7 and 8 that “yes” covert intentions increased [oxy-Hb] responses in prefrontal areas and decreased the responses over parieto-occipital areas in general. The inverted [oxy-Hb] responses were mostly seen for “no” intentions over the same brain areas; thereby these brain areas produced significantly different [oxy-Hb] responses between the two conditions (“yes” versus “no”), as shown in Fig. 7. In particular, the significant differences between the two conditions are observed on the prefrontal areas (ch. 47 to 50). As mentioned earlier, [deoxy-Hb] responses are almost reversed compared to the [oxy-Hb] responses, but there was a relatively smaller difference between the two conditions around the prefrontal areas. Thus, remarkable differences in [deoxy-Hb] responses between the two conditions are mostly observed around parieto-occipital areas, especially on right parietal areas (ch. 11, 14, 15, 23, 25, and 28), as shown in Fig. 8. It has been well documented that prefrontal and parietal brain areas are strongly associated with conscious intention, “holding something in mind,”^{21,36–38} which could explain the most significant differences in hemodynamic responses between internal “yes” and “no” conditions over prefrontal and parietal brain areas. In line with these results, the distribution of the Fisher’s score in Fig. 9 also showed that prefrontal and parietal areas would be most closely involved in processing binary covert intentions. In addition, the strong involvement of prefrontal areas observed in the [oxy-Hb] responses indicates that the binary communication system based on the direct intention decoding paradigm could be potentially implemented with portable “headband-type” fNIRS systems that cover only prefrontal areas.³⁹

Appendix: List of Questions Used for the Experiments

Class 1 – Pain

1. Is your waist hurting?
2. Is your head hurting?
3. Is your leg hurting?

4. Is your arm hurting?
5. Is your eye hurting?
6. Is your chest hurting?
7. Is your throat hurting?
8. Is your back hurting?
9. Is your stomach hurting?
10. Is your knee hurting?
11. Is your shoulder hurting?

Class 2 – Do you like ...?

12. Do you like horror movies?
13. Do you like baseball?
14. Do you like Starcraft?
15. Do you like billiard?
16. Do you like cream bread?
17. Do you like IU (Korean singer)?
18. Do you like Dong Won Kang (Korean actor)?
19. Do you like cats?
20. Do you like soju (Korean alcohol)?
21. Do you like broccolis?
22. Do you like soondae (Korean style intestine sausage soup)?
23. Do you like rice cakes?
24. Do you like raw halibut?
25. Do you like coffee?
26. Do you like sapgyulsal (Korean style pork belly)?
27. Do you like Ho Dong Kang (Korean Comedian)?
28. Do you like North Korea?
29. Do you like snow?
30. Do you like the smell of cigarette?
31. Do you like orange?

Class 3 – Do you want to go to ...?

32. Do you want to go to baseball stadium?
33. Do you want to go to North Korea?
34. Do you want to go to New York City?
35. Do you want to go to Japan?
36. Do you want to go to Pyramid?
37. Do you want to go to North Pole?
38. Do you want to go to Amazon?
39. Do you want to go to racetrack?

40. Do you want to go to army base?
41. Do you want to go to Desert Island?
42. Do you want to go to Everland (Korean amusement park)?
43. Do you want to go to Sahara Desert?
44. Do you want to go to Greece?
45. Do you want to go to East Coast?
46. Do you want to go to Mexico?
47. Do you want to go to Mt. Halla (Mountain located in Korea)?
48. Do you want to go to Thailand?
49. Do you want to go to Jeju Island?
50. Do you want to go to Alaska?
51. Do you want to go to Lax Vegas?

Class 4 – Do you like to get ...?

52. Do you like to get soccer shoes?
53. Do you like to get the New iPad?
54. Do you like to get Avante (a car made by Hyundai Motor Company)?
55. Do you like to get a fedora?
56. Do you like to get an audio?
57. Do you like to get a fountain pen?
58. Do you like to get a doll?
59. Do you like to get a flower?

Class 5 – Etc.

60. Are you afraid of height?
61. Are you afraid of cockroach?
62. Is your body stiff?
63. Do you like the experiment?
64. Are you sleepy?
65. Are you hungry?
66. Are you thirsty?
67. Do you feel cold?
68. Do you feel hot?
69. Do you feel disgusted?
70. Do you feel dizzy?

Acknowledgments

This work was supported in part by the National Research Foundation of Korea (NRF) funded by the Korean

Government (MSIP) (NRF-2015M3C7A1031969 and 2014R1A2A1A11051796), and was supported in part by the KRISS-WCL project (Development of Next-Generation Biomagnetic Resonance Technology).

References

1. H.-J. Hwang et al., "EEG-based brain-computer interfaces (BCIs): a thorough literature survey," *Int. J. Hum. Comput. Int.* **29**(12), 814–826 (2013).
2. N. Naseer and K.-S. Hong, "fNIRS-based brain-computer interfaces: a review," *Front. Hum. Neurosci.* **9**, 3 (2015).
3. B. Abibullaev and J. An, "Classification of frontal cortex haemodynamic responses during cognitive tasks using wavelet transforms and machine learning algorithms," *Med. Eng. Phys.* **34**(10), 1394–1410 (2012).
4. G. Bauernfeind et al., "Single-trial classification of antagonistic oxy-hemoglobin responses during mental arithmetic," *Med. Biol. Eng. Comput.* **49**(9), 979–984 (2011).
5. L. Holper and M. Wolf, "Single-trial classification of motor imagery differing in task complexity: a functional near-infrared spectroscopy study," *J. Neuroeng. Rehabil.* **8**, 34 (2011).
6. H.-J. Hwang et al., "Evaluation of various mental task combinations for near-infrared spectroscopy-based brain-computer interfaces," *J. Biomed. Opt.* **19**(7), 077005 (2014).
7. S. D. Power and T. Chau, "Automatic single-trial classification of prefrontal hemodynamic activity in an individual with Duchenne muscular dystrophy," *Dev. Neurorehabil.* **16**(1), 67–72 (2013).
8. 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**(2), 26002 (2010).
9. S. D. Power, A. Kushki, and T. Chau, "Towards a system-paced near-infrared spectroscopy brain-computer interface: differentiating prefrontal activity due to mental arithmetic and mental singing from the no-control state," *J. Neural Eng.* **8**(6), 066004 (2011).
10. R. Sitaram et al., "Temporal classification of multichannel near-infrared spectroscopy signals of motor imagery for developing a brain-computer interface," *NeuroImage* **34**(4), 1416–1427 (2007).
11. J. Perelmouter and N. Birbaumer, "A binary spelling interface with random errors," *IEEE Trans. Rehabil. Eng.* **8**(2), 227–232 (2000).
12. M. Naito et al., "A communication means for totally locked-in ALS patients based on changes in cerebral blood volume measured with near-infrared light," *IEICE Trans. Inf. Syst.* **E90-D**(7), 1028–1037 (2007).
13. N. Naseer, M. J. Hong, and K. S. Hong, "Online binary decision decoding using functional near-infrared spectroscopy for the development of brain-computer interface," *Exp. Brain Res.* **232**(2), 555–564 (2014).
14. N. Naseer et al., "Improving classification accuracy of covert yes/no response decoding using support vector machines: an fNIRS study," in *Proc. Int. Conf. on Robotics and Emerging Allied Technologies in Engineering (iCREATE)*, pp. 6–9, IEEE, Islamabad, Pakistan (2014).
15. J. Decety, "The neurophysiological basis of motor imagery," *Behav. Brain Res.* **77**(1–2), 45–52 (1996).
16. J. A. Pineda et al., "Learning to control brain rhythms: making a brain-computer interface possible," *IEEE Trans. Neural Syst. Rehabil. Eng.* **11**(2), 181–184 (2003).
17. H.-J. Hwang, K. Kwon, and C.-H. Im, "Neurofeedback-based motor imagery training for brain-computer interface (BCI)," *J. Neurosci. Methods* **179**(1), 150–156 (2009).
18. T. Ono, A. Kimura, and J. Ushiba, "Daily training with realistic visual feedback improves reproducibility of event-related desynchronisation following hand motor imagery," *Clin. Neurophysiol.* **124**(9), 1779–1786 (2013).
19. C. Vidaurre and B. Blankertz, "Towards a cure for BCI illiteracy," *Brain Topogr.* **23**(2), 194–198 (2010).
20. T. Dickhaus et al., "Predicting BCI performance to study BCI illiteracy," *BMC Neurosci.* **10**(Suppl. 1), P84 (2009).
21. G. Gallegos-Ayala et al., "Brain communication in a completely locked-in patient using bedside near-infrared spectroscopy," *Neurology* **82**(21), 1930–1932 (2014).
22. T. Yamamoto et al., "Arranging optical fibres for the spatial resolution improvement of topographical images," *Phys. Med. Biol.* **47**(18), 3429–3440 (2002).

23. D. T. Delpy et al., "Estimation of optical pathlength through tissue from direct time of flight measurement," *Phys. Med. Biol.* **33**(12), 1433–1442 (1988).
24. B. Hjorth, "An on-line transformation of EEG scalp potentials into orthogonal source derivations," *Electroencephalogr. Clin. Neurophysiol.* **39**(5), 526–530 (1975).
25. D. J. McFarland et al., "Spatial filter selection for EEG-based communication," *Electroencephalogr. Clin. Neurophysiol.* **103**(3), 386–394 (1997).
26. G. Pfurtscheller et al., "Focal frontal (de)oxyhemoglobin responses during simple arithmetic," *Int. J. Psychophysiol.* **76**(3), 186–192 (2010).
27. A. Kai Keng, Y. Juanhong, and G. Cuntai, "Extracting effective features from high density NIRS-based BCI for assessing numerical cognition," in *Proc. Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2233–2236, IEEE, Kyoto, Japan (2012).
28. Y. Juanhong et al., "A multimodal fNIRS and EEG-based BCI study on motor imagery and passive movement," in *Proc. 6th Int. Conf. on Neural Engineering (NER)*, pp. 5–8, IEEE, San Diego, USA (2013).
29. G. Bauernfeind et al., "Development, set-up and first results for a one-channel near-infrared spectroscopy system," *Biomed. Tech.* **53**(1), 36–43 (2008).
30. D. A. Benaron et al., "Noninvasive functional imaging of human brain using light," *J. Cereb. Blood Flow Metab.* **20**(3), 469–477 (2000).
31. S. D. Power, A. Kushki, and T. Chau, "Intersession consistency of single-trial classification of the prefrontal response to mental arithmetic and the no-control state by NIRS," *PLoS One* **7**(7), e37791 (2012).
32. K. Tai and T. Chau, "Single-trial classification of NIRS signals during emotional induction tasks: towards a corporeal machine interface," *J. Neuroeng. Rehabil.* **6**(1), 32 (2009).
33. M. Saba et al., "Automatic detection of a prefrontal cortical response to emotionally rated music using multi-channel near-infrared spectroscopy," *J. Neural Eng.* **9**(2), 026022 (2012).
34. F. Zeng, R. H. C. Yap, and L. Wong, "Using feature generation and feature selection for accurate prediction of translation initiation sites," *Genome Inform.* **13**, 192–200 (2002).
35. S. Luu and T. Chau, "Decoding subjective preference from single-trial near-infrared spectroscopy signals," *J. Neural Eng.* **6**(1), 016003 (2009).
36. P. Haggard, "Conscious intention and motor cognition," *Trends Cogn. Sci.* **9**(6), 290–295 (2005).
37. C. Frith and R. Dolan, "The role of the prefrontal cortex in higher cognitive functions," *Cognit. Brain Res.* **5**(1–2), 175–181 (1996).
38. J.-D. Haynes et al., "Reading hidden intentions in the human brain," *Curr. Biol.* **17**(4), 323–328 (2007).
39. P. V. Zephaniah and J. G. Kim, "Recent functional near infrared spectroscopy based brain computer interface systems: developments, applications and challenges," *Biomed. Eng. Lett.* **4**(3), 223–230 (2014).

Biographies for the authors are not available.