

# Journal of Biomedical Optics

BiomedicalOptics.SPIEDigitalLibrary.org

## **Multiclass classification of hemodynamic responses for performance improvement of functional near-infrared spectroscopy- based brain–computer interface**

Jaeyoung Shin  
Jichai Jeong

# Multiclass classification of hemodynamic responses for performance improvement of functional near-infrared spectroscopy-based brain–computer interface

Jaeyoung Shin and Jichai Jeong\*

Korea University, Department of Brain and Cognitive Engineering, 145 Anam-Ro, Sungbuk-Ku, Seoul 136-713, Republic of Korea

**Abstract.** We improved the performance of a functional near-infrared spectroscopy (fNIRS)-based brain–computer interface based on relatively short task duration and multiclass classification. A custom-built eight-channel fNIRS system was used over the motor cortex areas in both hemispheres to measure the hemodynamic responses evoked by four different motor tasks (overt execution of arm lifting and knee extension for both sides) instead of finger tapping. The hemodynamic responses were classified using the naive Bayes classifier. Among the mean, max, slope, variance, and median of the signal amplitude and the time lag of the signal, several signal features are chosen to obtain highest classification accuracy. Ten runs of threefold cross-validation were conducted, which yielded classification accuracies of  $87.1\% \pm 2.4\%$  to  $95.5\% \pm 2.4\%$ ,  $77.5\% \pm 1.9\%$  to  $92.4\% \pm 3.2\%$ , and  $73.8\% \pm 3.5\%$  to  $91.5\% \pm 1.4\%$  for the binary, ternary, and quaternary classifications, respectively. Eight seconds of task duration for obtaining sufficient quaternary classification accuracy was suggested. The bit transfer rate per minute (BPM) based on the quaternary classification accuracy was investigated. A BPM can be achieved from 2.81 to 5.40 bits/min. © 2014 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.JBO.19.6.067009]

Keywords: functional near-infrared spectroscopy; brain–computer interface; multiclassification.

Paper 130903RR received Dec. 23, 2013; revised manuscript received May 29, 2014; accepted for publication Jun. 2, 2014; published online Jun. 26, 2014.

## 1 Introduction

Recently, brain–computer interface (BCI) has received great attention as a promising alternative communication method. People with motor disorders who cannot use typical augmentative technologies can have opportunities to communicate or control prostheses using BCI.<sup>1–3</sup> BCI has been investigated using electroencephalography (EEG) and electrocorticography.<sup>4–6</sup> Particularly, EEG has many advantages, such as noninvasiveness, high temporal resolution, and convenience for the user. However, susceptibility to electrical interference from the environment and difficulties with the fixation of electrodes have been reported as disadvantages.<sup>7,8</sup> Although magnetoencephalography and functional magnetic resonance imaging have been considered as alternative modalities, they are impractical, since the instruments are large, expensive, and lack flexibility.<sup>9,10</sup>

Functional near-infrared spectroscopy (fNIRS) is one of the promising technologies for BCI. fNIRS is a powerful tool for monitoring the real-time concentration changes of hemoglobin in the blood. fNIRS has distinct advantages such as affordability, portability, and flexibility of use.<sup>11</sup> Since the first investigation of fNIRS for measuring brain activity,<sup>12</sup> many studies have been devoted to overcoming the drawbacks of fNIRS for BCI. Since the signal quality is affected by motion artifacts, Sato et al.,<sup>13</sup> Cui et al.,<sup>14</sup> and Scholkman et al.<sup>15</sup> proposed novel methods for the reduction of motion artifacts. Robertson et al.<sup>16</sup> compared motion artifact reduction methods in previous studies, such as recursive linear square adaptive filtering, wavelet filtering,

independent component analysis, and linear regression. As the signal strength is attenuated by hair on the head, Holper et al.,<sup>17</sup> Coyle et al.,<sup>18</sup> and Kozel et al.<sup>19</sup> brushed out the hair from the measurement sites, and Young et al.<sup>20</sup> measured the signal from participants with shaved hair. Khan et al.<sup>21</sup> suggested novel brush optodes to overcome poor optical contact with the scalp. Because the physiological interferences (regarding cardiac and respiratory activities) overlap with the signals activated by stimulation, it is difficult to distinguish between them. Zhang et al.<sup>22,23</sup> used an adaptive filter to effectively remove the physiological interferences.

Many researchers investigated the classification of the fNIRS signal, Sitaram et al.,<sup>24</sup> Falk et al.,<sup>25</sup> Power et al.,<sup>26</sup> Gottemukkula and Derakhshani,<sup>27</sup> and Ayaz et al.<sup>28</sup> investigated the performance of binary classification methods using a support vector machine, hidden Markov models, Gaussian mixture models, the K-nearest neighbor classifier, and the naive Bayes (NB) classifier, respectively, which are all well-known traditional classifiers. The information transfer rate (ITR) can be used to evaluate the practicability of classification.<sup>29</sup> It is clear that the ITR depends on the accuracy of classification, which decreases with increases in the number of data classes. Because these investigations considered only the binary classification, the ITR using fNIRS for BCI was limited to 1 bit/trial.<sup>29</sup> Multiclass classification (i.e., ternary and quaternary classifications) can theoretically provide a better chance to increase the ITR up to 2 bits/trial, while the multiclass classification makes it

\*Address all correspondence to: Jichai Jeong, E-mail: [jcj@korea.ac.kr](mailto:jcj@korea.ac.kr)

difficult to obtain higher classification accuracy. It is important to apply the appropriate data processing methods to improve the performance of the system. Power et al.<sup>30</sup> used ternary classification for increasing the performance of fNIRS-based BCI. Moreover, since the bit transfer rate per minute (BPM) depends on the time, reducing the task and rest duration is another important issue. In this paper, we carried out the experiments with eight participants. During the experiments, the participants execute the four kinds of motor executions (overt left/right arm lifting and knee extension). To increase the ITR of fNIRS-based BCI, quaternary classification should be applied. Quaternary classification is usually more difficult to implement than simple binary classification. Because this is the first investigation on quaternary classification using hemodynamic responses measured from fNIRS, it may be necessary to use distinct hemodynamic responses evoked by overt motor execution instead of motor imagery.<sup>24</sup> The task duration is reduced while maintaining a sufficient level of classification accuracy. Two different kinds of light-emitting diodes (LEDs) and avalanche photodiodes (APDs) were utilized as the sources and detectors of the system. The detected signals were filtered with band-pass and Savitsky–Golay smoothing filters as a simple method to remove the physiological and system interference.<sup>19</sup> The hair was brushed out from the measurement sites to obtain enough intensity of the received light.

The hemodynamic responses by the concentration changes of the oxy-hemoglobin were estimated using the modified Beer–Lambert law (MBLL). To increase BPM, we tried to reduce the task duration while maintaining the classification accuracy of larger than 70%. Various tasks and rest durations were set to measure the hemodynamic responses. The measured hemodynamic responses were classified using binary, ternary, and quaternary classifications. Based on the classification results, the available ITR and BPM were calculated. So far, fNIRS studies have addressed the binary classification, whereas in the present study, binary, ternary, and quaternary classifications are addressed. Moreover, the effectiveness of quaternary classification has been determined using only eight channels for fNIRS-based BCI. The available BPM estimated

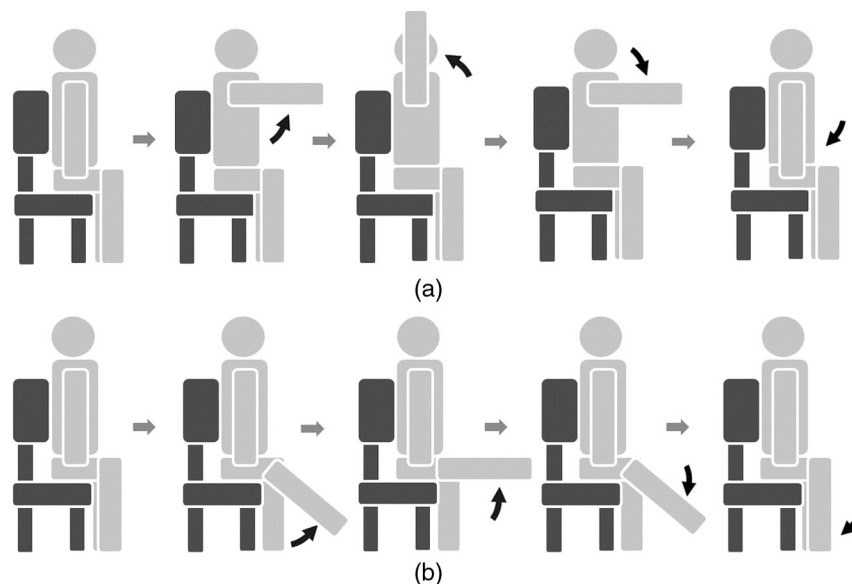
demonstrates the possibility of increasing the BPM of the fNIRS-based BCI.

This paper is organized as follows. The experimental protocol, system setup, and the data processing method used in this study are first introduced in Sec. 2. In Sec. 3, the proper task duration to increase BPM and the classification accuracy according to the number of classes of data are explained. In Sec. 4, the present results are compared with the previous studies, and the ITR and BPM are calculated based on the estimated classification accuracy. Finally, in Sec. 5, conclusions are presented.

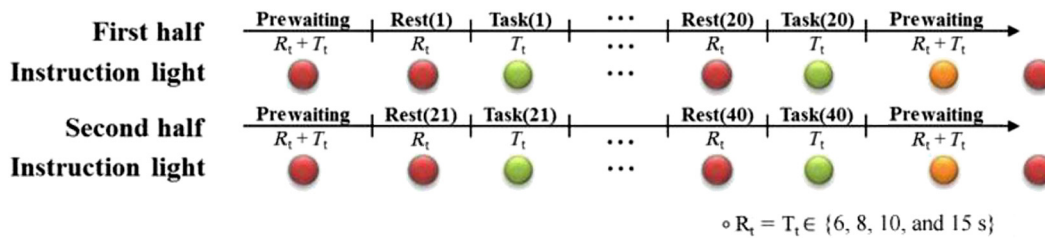
## 2 Methods

### 2.1 Participants and Experimental Protocols

Eight healthy participants participated in the study (seven males and one female, age:  $26.8 \pm 1.6$  years). None of the participants had histories of neurological or psychiatric issues. All the eight participants performed the experiments related to the four-class classification while only three participants among them were able to finish the experiments to determine the task duration. The experiments were approved by the Institutional Review Board of Korea University (KU-IRB-10-34-A-2). Written consent was obtained from all participants. Considering quaternary classification, it is necessary to obtain the hemodynamic responses evoked by four kinds of distinct behaviors. Since lifting up and down an arm or leg are more active movements than finger tapping, we adopted the active movements in our experiment. Figure 1 indicates the graphical motions of the active movements. Hemodynamic responses evoked by the motor tasks were collected from each participant while performing overt motor execution. Figure 2 shows the experimental design for motor execution. The task consisted of 40 cycles of 6, 8, 10, or 15-s motor task and rest. The durations of the task period ( $T_t$ ) and rest period ( $R_t$ ) were kept the same. The participant sat on a chair in front of a screen displaying instructions and the progress of the experiment. In the prewaiting period, a red light was



**Fig. 1** Graphical motions of (a) arm lifting and (b) knee extension. Participant sits on a chair and lifts up and down his/her arm or leg.



**Fig. 2** Experimental design of motor tasks for the measurement of hemodynamic responses using functional near-infrared spectroscopy (fNIRS). Each experiment was composed of a prewaiting time ( $T_t + R_t$  seconds), repetitions of rest periods for  $R_t$  seconds and task periods for  $T_t$  seconds, and postwaiting time ( $T_t + R_t$  seconds). Hemodynamic responses evoked by four kinds of motor tasks (left/right arm lifting and knee extension) were measured separately.

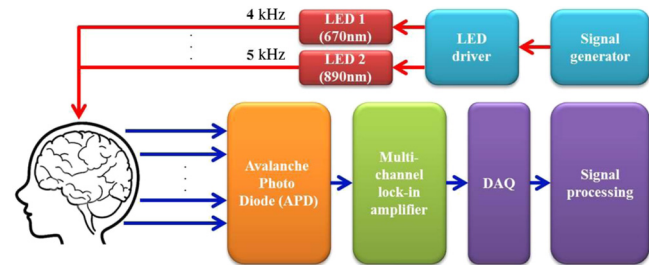
displayed. At the beginning of the task period, a green light was turned on. During this period, self-paced motor execution was performed. In the rest period, the color of the light was changed from green to red. The participant was asked not to think or move while staying in the rest state. The green and red lights were turned on and off repeatedly in the task and rest periods. In the postwaiting period, a yellow light was displayed and a red light was turned on when all the periods were finished.

A single experimental run was composed of a prewaiting period, 40 repetitions of rest and task periods, followed by a postwaiting period. Each participant first completed 20 repetitions of the rest and task periods in the first half. After a short rest, the participant completed the same number of repetitions again in the last half. Each half experimental run (20 repetitions) lasted 264, 352, 440, and 660 s for 6, 8, 10, and 15 s per one task period ( $T_t$ ), respectively. Each participant conducted a single task repeatedly in each experimental run. Hence, four sets of experimental runs (i.e., for left/right arm lifting and knee extension) were completed for each participant. The tasks were performed in the order of left arm lifting, right arm lifting, left knee extension, and right knee extension. Between each experimental run, the participants took a rest for a minute. Each participant was allowed to do some stretching and speak lightly, but was not allowed to leave the seat during the rest period.

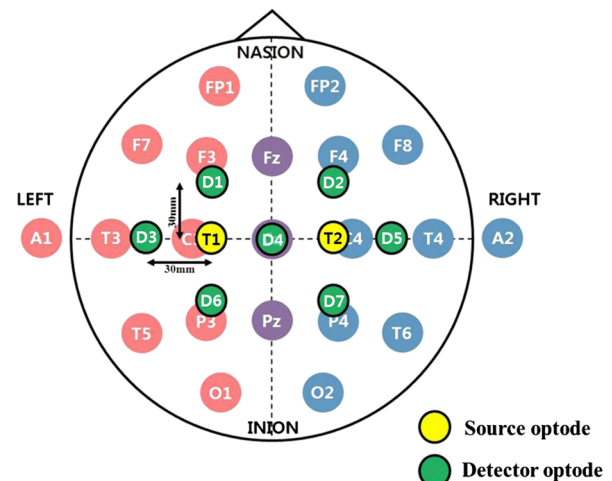
**2.2 System Setup and Data Processing**

All windows of the experiment room were covered by blackout curtains, and lights were turned off to keep out ambient light. An eight-channel custom-built continuous-wave fNIRS system was used to measure the hemodynamic responses. The fNIRS system uses LEDs operated at two different wavelengths of 670 nm (L6112-01, Hamamatsu photonics, Hamamatsu, Shizuoka, Japan) and 890 nm (L2656-03, Hamamatsu) which are similar to the wavelength of LEDs used by Coyle et al., because the fNIRS system used in this investigation was implemented based on their system.<sup>11</sup> The LEDs were placed in direct contact with the scalp. The lights were modulated with 4- and 5-kHz sinusoidal waves. The detectors used in the system were APDs (C5460-01, Hamamatsu), which collected the lights via 3-mm diameter multimode optical fibers. The detected light signals were first demodulated by a lock-in amplifier implemented by LabVIEW software (National Instruments, Austin, Texas). The data were recorded by a 16-bit data acquisition card (PXI-8106, National Instruments) and downsampled with a sampling frequency of 12.5 Hz. The whole system architecture is schematically illustrated in Fig. 3.

Figure 4 shows the arrangement of the fNIRS optodes. An fNIRS cap was made of elastic Lycra and shaped very similarly to a conventional EEG cap. The receiver optodes were fixed at a distance of 30 mm from the source optodes. The optodes were arranged on both sides of a hemisphere around the motor cortex areas around the  $C_3$ ,  $C_4$ , and  $C_2$  positions of the international 10 to 20 systems. The source and detector optode pairs consist of



**Fig. 3** Block diagram of an fNIRS-based brain-computer interface (BCI) system. 670- and 890-nm LEDs are modulated at 4 and 5 kHz, respectively. Avalanche photodiodes collect the lights, which are demodulated by lock-in amplifiers. The data are recorded by a 16-bit data acquisition card. Normalization, band-pass filtering, and Savitzky-Golay smoothing are applied for signal processing.

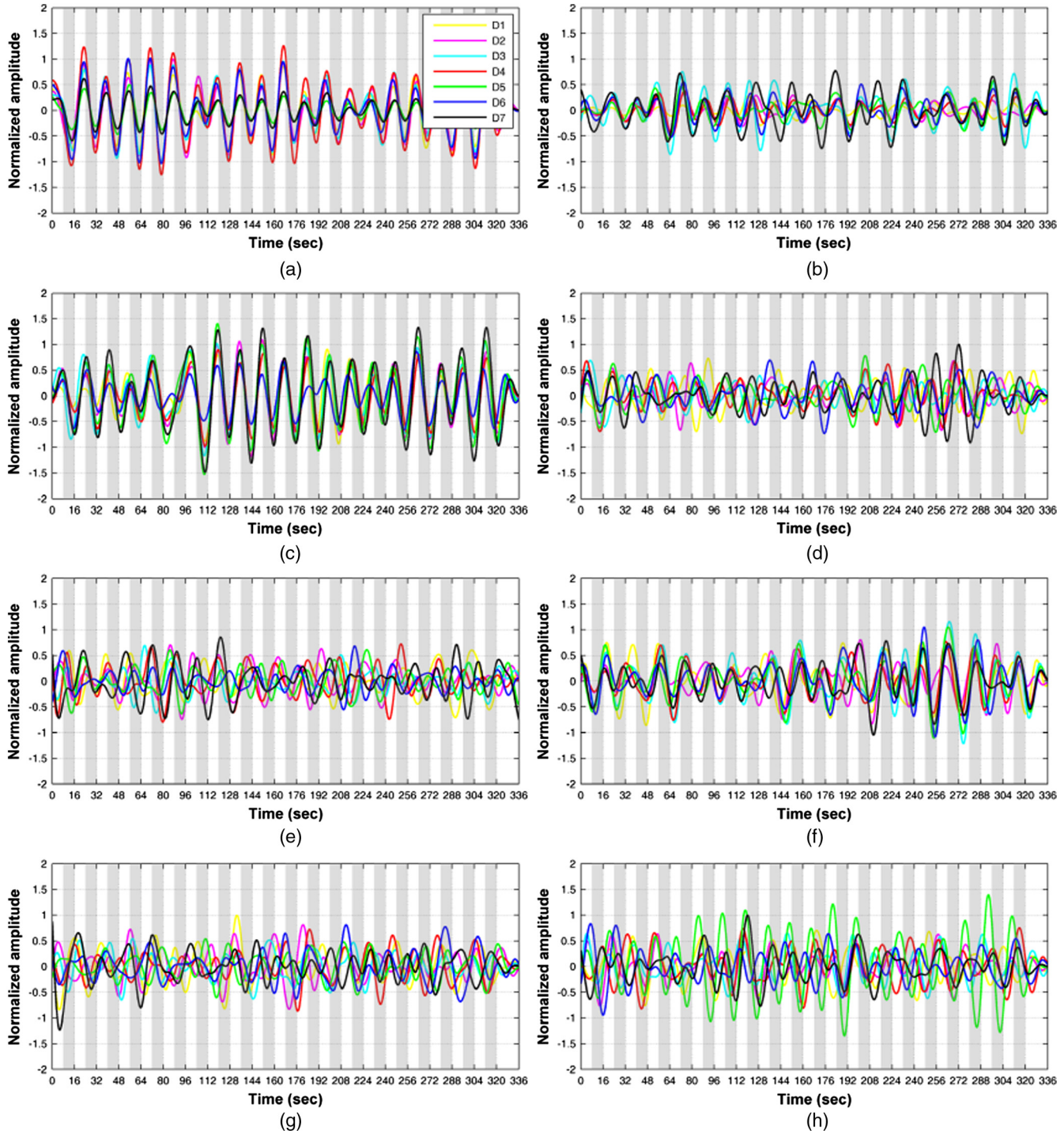


**Fig. 4** Optode arrangements for fNIRS. Yellow and green circles represent the source (S1 and S2) and detector optodes (D1 to D7), respectively. The source and detector are spaced 30 mm apart. The optodes were placed around the motor cortex area according to the international 10 to 20 systems.

eight channel arrangements. The concentration changes of oxy-hemoglobin were calculated from the changes in raw optical intensity data received using the MBLL given by<sup>31</sup>

$$\Delta C_{\text{HbO}} = \frac{\alpha_{\text{HbR}}^{\lambda_1} \frac{\Delta A^{\lambda_2}}{L^{\lambda_2}} - \alpha_{\text{HbR}}^{\lambda_2} \frac{\Delta A^{\lambda_1}}{L^{\lambda_1}}}{\alpha_{\text{HbR}}^{\lambda_1} \alpha_{\text{HbO}}^{\lambda_2} - \alpha_{\text{HbR}}^{\lambda_2} \alpha_{\text{HbO}}^{\lambda_1}}, \quad (1)$$

where  $\alpha_{\text{HbR}}$  and  $\alpha_{\text{HbO}}$  denote the absorption coefficients of oxy- and deoxy-hemoglobin at wavelengths  $\lambda_1$  and  $\lambda_2$ , respectively.  $\Delta A$  denotes absorbance and  $L$  [differential path length factor (DPF)  $\times$  distance between source and detector] is the total mean path length at wavelengths  $\lambda_1$  and  $\lambda_2$ .<sup>31</sup> Because we did not estimate a precise DPF value, a DPF of approximately 6.0 was used for both wavelengths.<sup>24</sup>

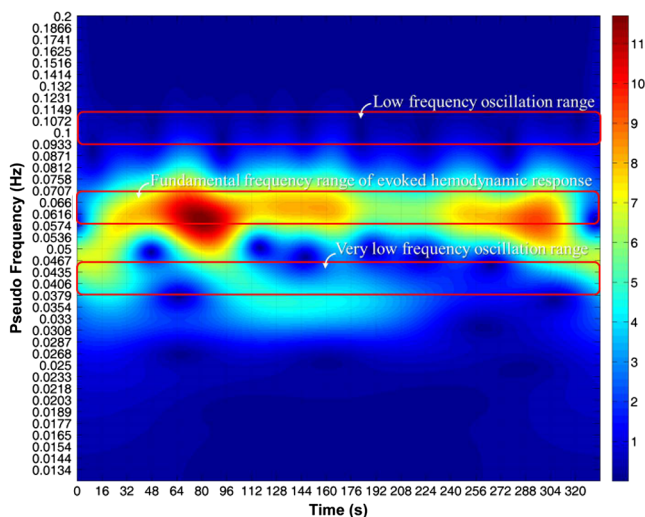


**Fig. 5** Hemodynamic responses ( $\Delta C_{\text{HbO}}$ ) measured from eight participants evoked by left arm lifting ( $T_t = R_t = 8$  s). (a)–(h)  $\Delta C_{\text{HbO}}$  from participants 1 to 8, respectively. Shaded regions represent task durations. Twenty motor task executions are repeated. Yellow, magenta, cyan, red, green, blue, and black lines indicate the hemodynamic responses measured at D1, D2, D3, D4, D5, D6, and D7 positions, respectively.

The calculated concentration changes of oxy-hemoglobin of the task and rest periods,  $\mathbf{X}$ , were normalized by subtracting the mean  $E(\mathbf{X})$  of the amplitude of the signal in the whole periods of each channel, and dividing by the standard deviation  $\sigma$  of the amplitude of the signal in the whole periods of each channel [i.e.,  $\{\mathbf{X} - E(\mathbf{X})\}/\sigma$ ]. We used a zero-phase distortion filter (third-order Chebyshev I with a ripple factor of 0.5 dB) with a passband of 0.03 to 0.07 Hz to remove the baseline drift and physiological noise because the fundamental frequency of the evoked hemodynamic response was expected to be in the passband. If we increase the bandwidth of the BPF, the evoked hemodynamic responses can be distorted by low frequency oscillation signals.<sup>23,32</sup> The correlation-based signal improvement method was used for minimizing the effect of head movements,<sup>14</sup> and a third-order Savitzky–Golay smoothing filter was used to determine the smoothed values for each point.<sup>15</sup> Figure 5 shows the measured hemodynamic responses ( $\Delta C_{\text{HbO}}$ ) from all the participants evoked by left arm lifting. Participant 1 result shows the clear change of the hemodynamic responses because the participant has shaved his/her hair so as not to weaken the intensity of the light. Since participants 2 and 3 have thin hair, we could obtain the relatively distinct evoked hemodynamic responses. Figure 6 shows the time-frequency analysis results of  $\Delta C_{\text{HbO}}$  measured from participant 1 using a continuous-time wavelet transform (Morlet wavelet). The fundamental frequency components of  $\Delta C_{\text{HbO}}$  are explicitly presented.

### 2.3 Feature Space

Among (1) mean, (2) max, (3) slope, (4) delay, (5) var, and (6) median, several features were chosen individually to obtain the highest classification accuracy. Individually chosen features are presented in Table 1. We set the calculation window with a length of  $T_t$  seconds to calculate the features after 8 s from the onset of a motor task. For example, when  $T_t = 15$  s, the features were calculated for 15 s after 8 s from the onset of a motor task. Forty-two features were calculated for each participant



**Fig. 6** Time-frequency analysis of the hemodynamics response ( $\Delta C_{\text{HbO}}$  at D2 of participant 1) using a continuous wavelet transform (Morlet). Low (around 0.1 Hz) and very low (around 0.04 Hz) frequency oscillations are much weaker than the evoked hemodynamic response.

**Table 1** Chosen features of fNIRS signals for classification.

Item	Feature	Remark
1. Mean	Mean value of signal amplitude	The most frequently chosen feature in the previous studies
2. Max	Maximum value of signal amplitude	Maximum value of signal amplitude within the calculation window
3. Slope	Average slope of signal amplitude	Average slope of signal amplitude within the calculation window
4. Delay	Channel activation delay time	Time lag of the channel activation until the signal amplitude exceeds the threshold value of zero
5. Var	Variance of signal amplitude	Expected value of the square of the deviations
6. Median	Median value of signal amplitude	Middle value in numerical order in the list of numbers appeared in the data

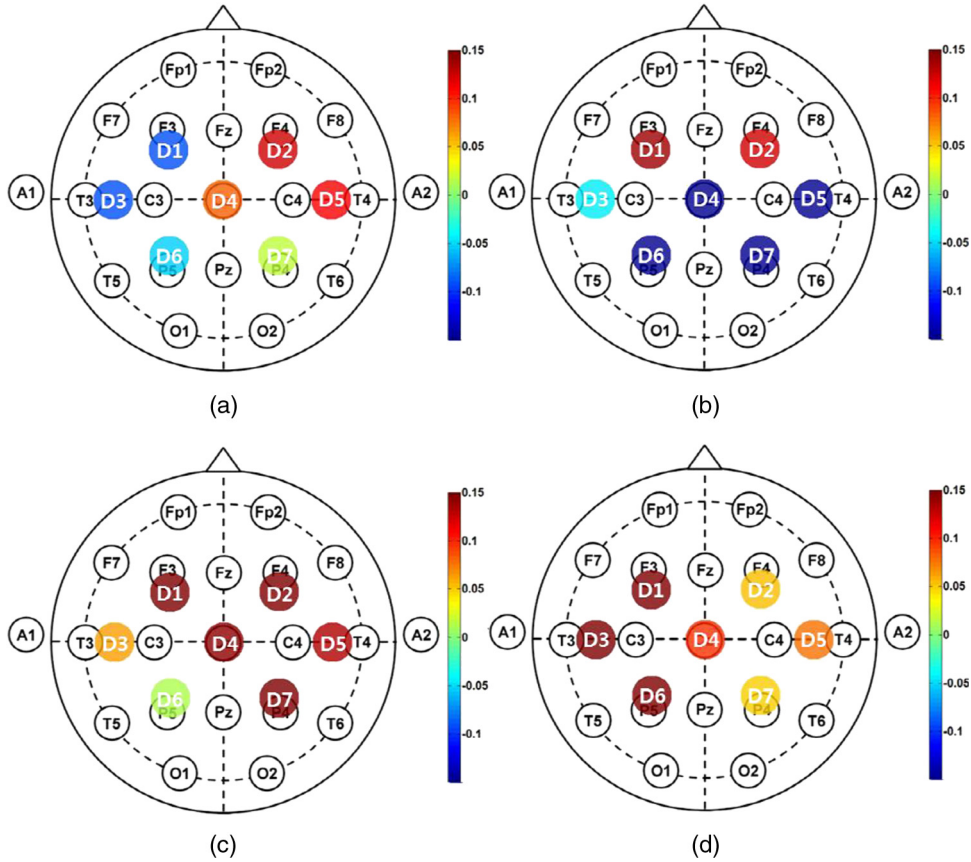
[6 features  $\times$  7 detectors  $\times$  1 signal ( $\Delta C_{\text{HbO}}$ )]. To the best of our knowledge, this is the first investigation using the max, slope, and median in the four-class classification of fNIRS signals. Figure 7 shows the grand average of the spatial hemodynamic responses of all the participants evoked by the four kinds of motor tasks. The contralateral hemispheres were more activated by the motor tasks. Since the knee and leg areas of Homunculus are near  $C_z$ , the knee extension activated more around  $C_z$  (D4 position) than the arm lifting.<sup>33</sup> From this physiological reason, we can classify the hemodynamic responses evoked by arm lifting and knee extension using the  $\Delta C_{\text{HbO}}$  at D4.

### 2.4 Dimension Reduction Method

A high-dimensional feature set is difficult to handle, since high dimensionality may increase the noise in the data set, which may degrade the performance of the classifier. This is known as “the curse of dimensionality.”<sup>34</sup> To prevent this phenomenon, principal component analysis (PCA) and linear discriminant analysis (LDA) were used to reduce the dimension of each channel data.<sup>35</sup> PCA converts a set of the correlated data into a set of the uncorrelated components. LDA maximizes the ratio of between-class variance to within-class variance in the data set.<sup>36</sup> After that, we considered every possible dimensionality of the feature set, and the dimensionality of the feature set which shows the highest classification accuracy was found and then used. Dimension reduction was performed using the entire measured data. Figure 8 shows the signal processing flow of fNIRS signals.

### 2.5 Classification and Outlier Elimination

The NB classification based on the Bayesian theorem was used. NB assumes that all the features are independent. Based on the Bayes’ theorem, the sample data were classified under the maximum likelihood rule (ML rule). Outliers were eliminated to remove the unreliable data. Ten out of 40 sample data were selected as outliers. Outliers were the farthest data from the



**Fig. 7** Grand average of the spatial hemodynamic responses ( $\Delta C_{HbO}$ ) of all the participants evoked by motor task of (a) left arm lifting, (b) right arm lifting, (c) left knee extension, and (d) right knee extension. (a)–(d) The relatively large  $\Delta C_{HbO}$  on the contralateral hemisphere channels. (c) and (d) The large  $\Delta C_{HbO}$  compared with (a) and (b). Knee extension evokes larger  $\Delta C_{HbO}$  than arm lifting at D4.

average  $\Delta C_{HbO}$  of the measured data. After excluding the outliers, 10 runs of the threefold cross-validation method were used to validate the results using the training and validation (test) subsets. Thirty samples were randomly divided into three subsets. Each subset consisted of 10 samples. Two subsets were designated as the training subsets and the remaining subset was used for validation of the trained classifier. In the training step, NB estimates the parameters of the probability distribution of the samples by using the training samples in the training subsets. In the classification step, using the test samples in the validation subsets, NB calculates the posterior probability of the sample and determines the class of the test samples in accordance with the maximum posterior probability (i.e., the ML rule).<sup>37</sup> Because NB only requires a small amount of training data to estimate the parameters necessary for classification, NB is extremely fast compared to other more sophisticated methods.<sup>38</sup>

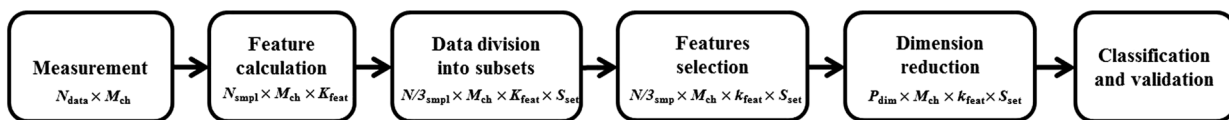
## 2.6 Data Transfer Rate

The ITR was calculated in bits/trial as follows:<sup>29</sup>

$$ITR = \log_2 N + P \log_2 P + (1 - P) \log_2 \frac{1 - P}{N - 1}, \quad (2)$$

where  $N$  is the number of classes of the data and  $P$  is the classification accuracy to evaluate the performance of BCI. Based on ITR, the BPM can be given by

$$BPM = ITR \times (\text{duration of task and rest period}) \text{ (bits/min)}. \quad (3)$$



**Fig. 8** Signal processing flow of fNIRS signals. Dimensionalities of the data used in each step are represented.  $N_{data}$ ,  $M_{ch}$ ,  $N_{smp1}$ ,  $K_{feat}$ ,  $S_{set}$ , and  $P_{dim}$  are the number of data points, channels, samples, features, and dimensions, respectively.  $k_{feat}$  is the number of features selected for classification ( $1 \leq k_{feat} \leq K_{feat}$ ).

### 3 Classification Results

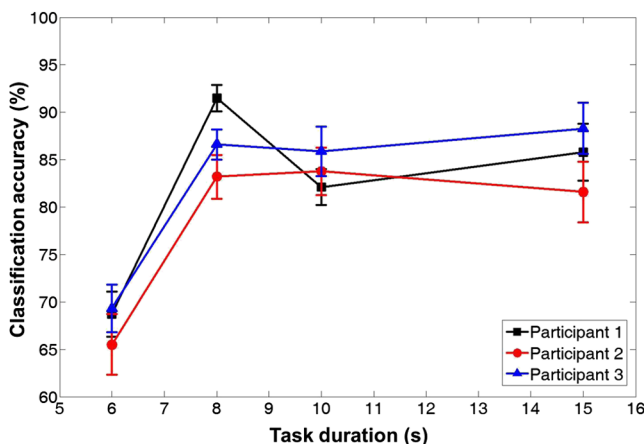
#### 3.1 Task Duration

Naseer and Hong<sup>39</sup> considered various task durations with constant rest durations based on a binary classification. We estimate the proper task duration to increase BPM while maintaining a quaternary classification accuracy of larger than 70%. Figure 9 shows the average quaternary classification accuracy of four different tasks for participants 1–3 as a function of task duration. The highest classification accuracies were observed at different task durations. For participant 1, a maximum classification accuracy of 91.5% was observed at a task duration of 8 s. At a task duration of 6 s, the classification accuracy was estimated to be below 70% for all the participants according to the threshold of the sufficient classification accuracy for BCI.<sup>40–42</sup> Among the task durations that achieve classification accuracy above 70%,  $T_t$  values of 8 s each can be chosen as the task duration.

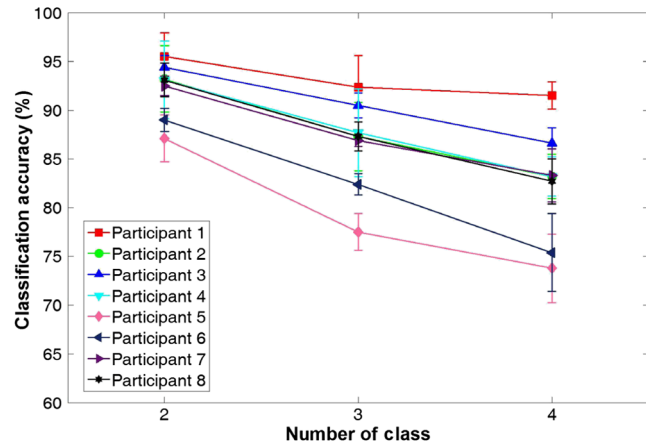
#### 3.2 Multiclass Classification

The classification performance of the binary, ternary, and quaternary classifications was investigated at the  $T_t = 8$  s. For binary classification, the six combinations ( ${}_4C_2 = 6$ ) of two data classes among four data classes (LA: left arm lifting, RA: right arm lifting, LK: left knee extension, and RK: right knee extension) were compared [i.e., LA versus RA), (LA versus LK), (LA versus RK), (RA versus LK), (RA versus RK), and (LK versus RK)]. Four combinations ( ${}_4C_3 = 4$ ) of three data classes [i.e., LA versus RA versus LK), (LA versus RA versus RK), (LA versus LK versus RK), and (RA versus LK versus RK)], and one combination ( ${}_4C_4 = 6$ ) of four data classes [i.e., LA versus RA versus LK versus RK)] were considered for ternary and quaternary classifications, respectively.

Figure 10 shows the average classification accuracy of four different classes of the data for eight participants according to the number of classes of the data. The classification accuracy decreases with increases in the number of data classes. The binary classification shows the highest classification accuracy for each participant, ranging from 87.1% to 95.5%, and the



**Fig. 9** Average quaternary classification accuracies as a function of the task duration. Square, circle, and triangle symbols indicate the classification accuracies of participants 1, 2, and 3, respectively. Error bar indicates the standard deviation. At a window length of 6 s, the classification accuracies of all the participants were below 70%.



**Fig. 10** Average classification accuracies of eight participants according to the number of data classes ( $N = 2, 3,$  and  $4$ ) for  $T_t = R_t = 8$  s. Error bar indicates the standard deviation. The highest and lowest classification accuracies are observed when the number of classes of the data is 2 ( $N = 2$ ) and 4 ( $N = 4$ ), respectively.

quaternary classification shows the lowest classification accuracy for each participant, from 73.8% to 91.5%. Table 2 shows the average classification accuracies according to the number of data classes. The features set of each participant which showed the highest classification accuracy was selected. Because 10 runs of cross-validation were conducted, the results are presented as the means  $\pm$  standard deviation. The highest classification accuracy was obtained using PCA for participants 1, 3, 6, 7, and 8, because the LDA could not be applied to the high-dimensional feature vector because of the singular scatter matrix problem.<sup>43</sup> However, the classification accuracy obtained by LDA and applied to a low-dimensional feature vector was comparable to that obtained by PCA.

#### 3.3 Data Transfer Rate

The ITR for all the participants with the number of data classes ranging from two to four are presented as a function of classification accuracy in Fig. 11. The maximum and minimum classification accuracies were achieved when the number of data classes was two ( $N = 2$ ) and four ( $N = 4$ ), respectively. For all the participants, the ITR increases as the number of classes increases. The maximum ITR values were 1.44, 1.08, 1.22, 1.08, 0.75, 0.80, 1.08, and 1.06 bits/trial for participants 1 through 8, respectively. At  $T_t = 8$  s, the available BPMs were 2.81 to 5.40 bits/min [0.75 to 1.44 bits/trial  $\times$  3.75 trials/min (0.27 min/trial)].

## 4 Discussion

#### 4.1 Comparison to Previous Studies

In many studies, classification has been performed using finger tapping as the motor imagery task. For an EEG-based BCI, quaternary classification using motor imagery tasks has been investigated, while binary or ternary classifications have been considered for fNIRS-based BCI.<sup>29,30</sup> Although motor imagery is more suitable for BCI, in this study, overt motor execution was performed for motor tasks to obtain distinct hemodynamic responses, because overt execution leads to more distinct activated signals. In spite of the quaternary classification, the



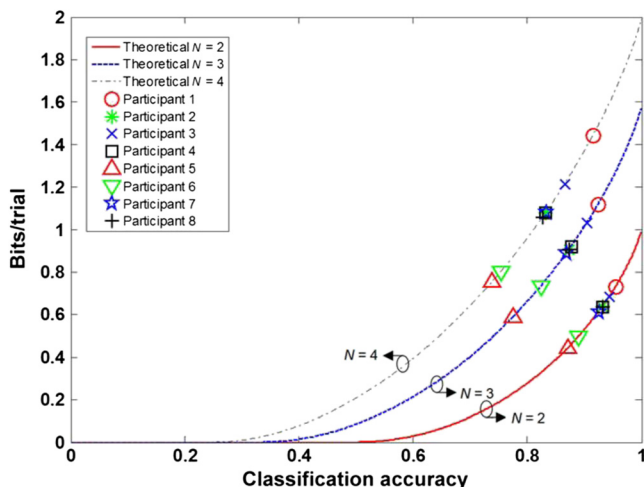
**Table 2** Average classification accuracies according to the number of data classes for binary, ternary, and quaternary classifications for  $T_t = R_t = 8$  s. Principal component analysis (PCA) and linear discriminant analysis (LDA) are applied, and the reduced dimensions by PCA or LDA are represented in the brackets. Among mean (1), max (2), slope (3), delay (4), var (5), and median (6), the feature sets are determined for obtaining the highest classification accuracy. Because 10 runs of cross-validation were conducted, the classification accuracies are presented in the form of mean  $\pm$  standard deviation (SD).

Participant number	Dimension reduction (dim.)	Chosen feature set	Binary (%) (mean $\pm$ SD)	Ternary (%) (mean $\pm$ SD)	Quaternary (%) (mean $\pm$ SD)
1	PCA (4)	1, 3	95.5 $\pm$ 2.4	92.4 $\pm$ 3.2	91.5 $\pm$ 1.4
2	LDA (3)	2, 3, 6	93.2 $\pm$ 3.4	87.3 $\pm$ 3.5	83.2 $\pm$ 2.3
3	PCA (7)	1, 3	94.4 $\pm$ 1.2	90.5 $\pm$ 1.3	86.6 $\pm$ 1.6
4	LDA (3)	2, 3, 6	93.1 $\pm$ 4.0	87.7 $\pm$ 4.5	83.2 $\pm$ 2.0
5	LDA (7)	3, 6	87.1 $\pm$ 2.4	77.5 $\pm$ 1.9	73.8 $\pm$ 3.5
6	PCA (6)	1, 3, 5	89.0 $\pm$ 1.2	82.4 $\pm$ 1.1	75.4 $\pm$ 4.0
7	PCA (6)	3, 4, 6	92.5 $\pm$ 1.0	86.9 $\pm$ 0.6	83.3 $\pm$ 2.7
8	PCA (7)	3, 6	93.1 $\pm$ 1.7	87.3 $\pm$ 1.5	82.7 $\pm$ 2.3

classification accuracies achieved in this study were comparable to those of the binary classification in other studies.<sup>17,24,26</sup>

## 4.2 Experimental Consistency

Since the classification results can potentially be affected by the global changes in the signals by the contact condition of the optodes, ambient light, participant fatigue, attention, and physiological change, a custom-built head cap was used to maintain the optode contact conditions and to totally prevent the effects of ambient light on the experimental results. Physiological changes of the participants were filtered out at the preprocessing step. Because the experiments lasted for a long time, an exhausted participant could not participate in some experiments, as shown in Fig. 9.



**Fig. 11** Achievable information transfer rates (ITRs) as a function of the different numbers of data classes for all the participants ( $N = 2, 3,$  and  $4$ ). Dash-dotted, dotted, and solid lines are the theoretical values of the ITR using BCI when the number of data classes is 2, 3, and 4, respectively.

## 4.3 Data Transfer Rate

An experiment was designed to investigate four-choice fNIRS-based BCI, since the achievable BPM using a binary system is just 1 to 2 bits/min. The suggested quaternary system could achieve a BPM of up to 5.33 bits/min with lower accuracy as a trade-off. Although Blankertz et al. achieved 35 bits/min using a 128-channel EEG system,<sup>4</sup> it is impossible for an fNIRS-based BCI system to achieve a comparable ITR because of the inherent hemodynamic delay. A task duration and rest duration time longer than 16 s ( $T_t + R_t$ ) is enough to obtain sufficient accuracy based on the results. Therefore, 7.5 bits/min [2 bits/trial (theoretically maximum B)  $\times$  3.75 trials/min] may be the maximum BPM using a four-choice fNIRS-based BCI system. This result shows the limitation of the fNIRS-based BCI using the quaternary system.

## 5 Conclusions

The performance of BCI has been improved in terms of the ITR using fNIRS by reducing the task duration as much as possible. The task duration should be 8 s while maintaining a sufficient level of classification accuracy for fNIRS-based BCI. The calculated average classification accuracy was as high as 95.5% for binary classification, 92.4% for ternary classification, and 91.5% for quaternary classification. Using the four-choice system, sufficient classification accuracy to increase the ITR was obtained. As a result, an available BPM of 2.78 to 5.33 bits/min was achieved using the BCI. From the experimental results, we have successfully classified the hemodynamic responses evoked by short motor tasks using multiclass classification to increase the BPM.

## Acknowledgments

This research was supported in part by Korea University Grant and Brain Korea 21 plus project in 2014.

## References

1. M. A. Lebedev and M. A. L. Nicolelis, "Brain-machine interfaces: past, present and future," *Trends Neurosci.* **29**(9), 536–546 (2006).
2. X. Gao et al., "A BCI-based environmental controller for the motion-disabled," *IEEE Trans. Neural Syst. Rehabil. Eng.* **11**(2), 137–140 (2003).
3. G. Pfurtscheller et al., "Graz-BCI: state of the art and clinical applications," *IEEE Trans. Neural Syst. Rehabil. Eng.* **11**(2), 177–180 (2003).
4. B. Blankertz et al., "The Berlin brain-computer interface: EEG-based communication without subject training," *IEEE Trans. Neural Syst. Rehabil. Eng.* **14**(2), 147–152 (2006).
5. H. Serby, E. Yom-Tov, and G. F. Inbar, "An improved P300-based brain-computer interface," *IEEE Trans. Neural Syst. Rehabil. Eng.* **13**(1), 89–98 (2005).
6. T. Pistohl et al., "Prediction of arm movement trajectories from ECoG-recordings in humans," *J. Neurosci. Methods* **167**(1), 105–114 (2008).
7. J. Zeng and Y. Sun, "Motor disability assistance based on novel non-invasive electroencephalographic decoding," in *GCIS 2012*, pp. 92–94, IEEE, New Jersey (2012).
8. N. A. Chadwick, D. A. McMeekin, and T. Tan, "Classifying eye and head movement artifacts in EEG signals," in *DEST 2011*, pp. 285–291, IEEE, New Jersey (2011).
9. J. Mellinger et al., "An MEG-based brain-computer interface (BCI)," *NeuroImage* **36**(3), 581–593 (2007).
10. S. S. Yoo et al., "Brain-computer interface using fMRI: spatial navigation by thoughts," *Neuroreport* **15**(10), 1591–1595 (2004).
11. S. Coyle et al., "On the suitability of near-infrared (NIR) systems for next-generation brain-computer interfaces," *Physiol. Measure.* **25**(4), 815–822 (2004).
12. F. F. Jobsis, "Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters," *Science* **198**(4323), 1264–1267 (1977).
13. H. Sato et al., "Wavelet analysis for detecting body-movement artifacts in optical topography signals," *NeuroImage* **33**(2), 580–587 (2006).
14. X. Cui, S. Bray, and A. L. Reiss, "Functional near infrared spectroscopy (NIRS) signal improvement based on negative correlation between oxygenated and deoxygenated hemoglobin dynamics," *NeuroImage* **49**(4), 3039–3046 (2010).
15. F. Scholkman et al., "How to detect and reduce movement artifacts in near-infrared imaging using moving standard deviation and spline interpolation," *Physiol. Measure.* **31**(5), 649–662 (2010).
16. F. C. Robertson, T. S. Douglas, and E. M. Meintjes, "Motion artifact removal for functional near infrared spectroscopy: a comparison of methods," *IEEE Trans. Biomed. Eng.* **57**(6), 1377–1387 (2010).
17. 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**(4), 1105–1113 (2009).
18. S. Coyle et al., "A mechanical mounting system for functional near-infrared spectroscopy brain imaging studies," in *Opto-Ireland 2005: Optical Sensing and Spectroscopy*, pp. 618–627, SPIE Europe, Cardiff, UK (2005).
19. F. A. Kozel et al., "Using simultaneous repetitive transcranial magnetic stimulation/functional near infrared spectroscopy (rTMS/fNIRS) to measure brain activation and connectivity," *NeuroImage* **47**(4), 1177–1184 (2009).
20. A. E. R. Young et al., "Behaviour of near-infrared light in the adult human head: implications for clinical near-infrared spectroscopy," *Br. J. Anaesth.* **84**(1), 38–42 (2000).
21. B. Khan et al., "Improving optical contact for functional near-infrared brain spectroscopy and imaging with brush optodes," *Biomed. Opt. Express* **3**(5), 878–898 (2012).
22. Q. Zhang, E. N. Brown, and G. E. Strangman, "Adaptive filtering for global interference cancellation and real-time recovery of evoked brain activity: a Monte Carlo simulation study," *J. Biomed. Opt.* **12**(4), 044014 (2007).
23. Q. Zhang, G. E. Strangman, and G. Ganis, "Adaptive filtering to reduce global interference in non-invasive NIRS measures of brain activation: how well and when does it work?," *NeuroImage* **45**(3), 788–794 (2009).
24. 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).
25. T. H. Falk et al., "Improving the performance of NIRS-based brain-computer interfaces in the presence of background auditory distractions," in *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, IEEE, New Jersey (2010).
26. 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), 026002 (2010).
27. V. Gottemukkula and R. Derakhshani, "Classification-guided feature selection for NIRS-based BCI," in *Proc. 5th Int. IEEE/EMBS Conf. on Neural Engineering (NER)*, pp. 72–75, IEEE, New Jersey (2011).
28. H. Ayaz et al., "Assessment of cognitive neural correlates for a functional near infrared-based brain-computer interface system," in *Foundations of Augmented Cognition. Neuroergonomics and Operational Neuroscience*, D. Schmorow, I. Estabrooke, and M. Grootjen, Eds., pp. 699–708, Springer, Berlin, Heidelberg (2009).
29. B. Obermaier et al., "Information transfer rate in a five-classes brain-computer interface," *IEEE Trans. Neural Syst. Rehabil. Eng.* **9**(3), 283–288 (2001).
30. S. D. Power, A. Kushki, and T. Chau, "Automatic single-trial discrimination of mental arithmetic, mental singing and the no-control state from prefrontal activity: toward a three-state NIRS-BCI," *BMC Res. Notes* **5**, 141 (2012).
31. L. Kocsis, P. Herman, and A. Eke, "The modified Beer-Lambert law revisited," *Phys. Med. Biol.* **51**(5), N91–N98 (2006).
32. R. W. S. Alan, V. Oppenheim, and J. R. Buck, *Discrete-Time Signal Processing*, Prentice-Hall, Inc., Upper Saddle River, New Jersey (1999).
33. G. D. Schott, "Penfield's homunculus: a note on cerebral cartography," *J. Neurol. Neurosurg. Psychiatry* **56**(4), 329–333 (1993).
34. R. E. Bellman, *Dynamic Programming*, Courier Dover Publications, Mineola, NY (2003).
35. A. Bashashati et al., "A survey of signal processing algorithms in brain-computer interfaces based on electrical brain signals," *J. Neural Eng.* **4**(2), R32–R57 (2007).
36. K. Fukunaga, *Statistical Pattern Recognition*, Academic Press, San Diego, California (1990).
37. H. Zhang, "The optimality of naive Bayes," in *Proc. 17th Int. Florida Artificial Intelligence Research Symposium*, FLAIRS, Florida (2004).
38. D. Lewis, "Naive (Bayes) at forty: the independence assumption in information retrieval," in *Machine Learning: ECML-98*, C. Nédellec and C. Rouveirol, Eds., pp. 4–15, Springer, Berlin, Heidelberg (1998).
39. N. Naseer and K. S. Hong, "Classification of functional near-infrared spectroscopy signals corresponding to the right- and left-wrist motor imagery for development of a brain-computer interface," *Neurosci. Lett.* **553**, 84–89 (2013).
40. A. Kubler et al., "Patients with ALS can use sensorimotor rhythms to operate a brain-computer interface," *Neurology* **64**(10), 1775–1777 (2005).
41. A. Kubler et al., "BCI meeting 2005—workshop on clinical issues and applications," *IEEE Trans. Neural Syst. Rehabil. Eng.* **14**(2), 131–134 (2006).
42. F. Nijboer et al., "An auditory brain-computer interface (BCI)," *J. Neurosci. Methods* **167**(1), 43–50 (2008).
43. H. Gao and J. W. Davis, "Why direct LDA is not equivalent to LDA," *Pattern Recognit.* **39**(5), 1002–1006 (2006).

**Jaeyoung Shin** received a BS and PhD degree in radio communications engineering from Korea University, Seoul, Korea, in 2007 and 2013, respectively. His current research interests include power-line communications, functional near-infrared spectroscopy, and brain computer interface.

**Jichai Jeong** received a BS degree from Korea University, Seoul, Korea, in 1980, the M.S. degree from KAIST, Daejeon, Korea in 1982, and a PhD degree in electrical engineering from Carnegie Mellon University, Pittsburgh, USA, in 1988. He has been with the Faculty of Department of Brain and Cognitive Engineering, Korea University since 2008. His current research interests include brain-computer interface, optical coherence tomography.