Classification of prefrontal activity due to mental arithmetic and music imagery using hidden Markov models and frequency domain near-infrared spectroscopy

2010 J. Neural Eng. 7 026002

(http://iopscience.iop.org/1741-2552/7/2/026002)

J. Neural Eng. **7** (2010) 026002 (9pp) doi:10.1088/1741-2560/7/2/026002

Classification of prefrontal activity due to mental arithmetic and music imagery using hidden Markov models and frequency domain near-infrared spectroscopy

Sarah D Power[1,2], **Tiago H Falk**[1,2] **and Tom Chau**[1,2]

[1] Bloorview Research Institute, Bloorview Kids Rehab, Toronto, Ontario, Canada
[2] Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, Ontario, Canada

E-mail: tom.chau@utoronto.ca

Received 16 September 2009
Accepted for publication 19 January 2010
Published 18 February 2010
Online at stacks.iop.org/JNE/7/026002

Abstract
Near-infrared spectroscopy (NIRS) has recently been investigated as a non-invasive brain–computer interface (BCI). In particular, previous research has shown that NIRS signals recorded from the motor cortex during left- and right-hand imagery can be distinguished, providing a basis for a two-choice NIRS-BCI. In this study, we investigated the feasibility of an alternative two-choice NIRS-BCI paradigm based on the classification of prefrontal activity due to two cognitive tasks, specifically mental arithmetic and music imagery. Deploying a dual-wavelength frequency domain near-infrared spectrometer, we interrogated nine sites around the frontopolar locations (International 10–20 System) while ten able-bodied adults performed mental arithmetic and music imagery within a synchronous shape-matching paradigm. With the 18 filtered AC signals, we created task- and subject-specific maximum likelihood classifiers using hidden Markov models. Mental arithmetic and music imagery were classified with an average accuracy of 77.2% ± 7.0 across participants, with all participants significantly exceeding chance accuracies. The results suggest the potential of a two-choice NIRS-BCI based on cognitive rather than motor tasks.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Despite major advances in assistive technologies, a significant number of individuals with severe and multiple motor disabilities still have no means of expressing something as basic as functional intent. This has become a prevalent issue in rehabilitation engineering, and major efforts have been made in the development of novel 'access technologies' [1] which may provide such individuals with unconventional channels for communication and environmental control. Brain–computer interfaces (BCIs), for example, are controlled through brain activity alone and have emerged as promising access solutions for individuals who lack sufficient motor control to operate more conventional movement-based devices. Such individuals may include those with late-stage amyotrophic lateral sclerosis (ALS), brainstem stroke, or severe cerebral palsy [2].

In this paper, our focus is on the potential development of a safe, reliable, noninvasive BCI based on near-infrared spectroscopy (NIRS), an optical imaging technology that can be used to assess functional activity in the brain via the hemodynamic response. An NIRS-based BCI would allow an individual with severe motor disability to access an assistive device, such as a communication aid or environmental control unit, by consciously controlling functional activity in a

pre-defined region of the brain. When the user desires to activate the device, he or she would simply perform some simple mental task known to elicit a specific spatial and temporal activation pattern in the region under interrogation. This activity would be measured (by NIRS), classified, and the resultant signals used as input to control the assistive device.

Though NIRS has been used extensively in various other applications to assess functional brain activity [3–18], its potential as a basis for BCI technologies has only recently been explored. Initial studies involving able-bodied participants focused on the use of motor imagery as the paradigm of BCI control [19–21]. Because motor cortical activation due to motor tasks had been investigated with positive results in other NIRS studies [3–11] and because motor imagery had been shown to work well in previous BCIs based on electroencephalography (EEG) [22, 23], this form of brain activation was appealing for NIRS-BCI development.

In a five-subject study, Sitaram *et al* demonstrated the possibility of classifying NIRS signals arising from right- and left-hand motor imagery with accuracy around 89.1% [24]. The ability to classify these two activities opens the door for the development of an NIRS-BCI two-choice cursor control paradigm. The work described in [24] proposed a word speller interface in which the user employs left- or right-hand imagery to move the cursor to the left or right, respectively, to select the letter of choice.

However, for individuals with congenital motor impairments, or for whom a significant length of time has lapsed since the loss of motor function, eliciting functional activity via motor imagery in a manner adequate for BCI operation may be difficult, if not impossible. Indeed, imaging studies have identified significant brain motor function defects in individuals with complete and chronic spinal cord injury, including strongly reduced volume of activation, poor modulation of function and abnormal activation patterns due to motor imagery [25], while other studies have shown significant impairment in mental representation and manipulation of body parts in individuals with locked-in syndrome (LIS) [26]. Recent NIRS-BCI studies have addressed this issue by also considering higher cognitive tasks associated with the prefrontal cortex (PFC), an area of the brain less likely to be implicated in motor disability. Such studies have shown the feasibility of using NIRS to detect functional activity due to mental singing [27], various verbal tasks [28, 29], mental arithmetic [27, 29, 30] and other working memory tasks [31] for the purpose of BCI development. Measuring functional activity via the PFC rather than the motor cortex is also advantageous in that signal attenuation and motion artefacts due to hair, which have been reported to be the primary factor in system performance degradation [20], are mitigated.

The aim of the present study was to investigate the feasibility of differentiating between two different cognitive tasks in the PFC using NIRS measurements. If two such activities can be successfully classified with acceptable accuracy, it could provide an alternative to the motor imagery-based two-choice BCI paradigm for those populations mentioned above. The cognitive tasks between which we attempted to discriminate were mental arithmetic and music imagery.

Though the primary processing of arithmetic/numeric operations seems to occur in the parietal cortex [32–34], the lateral PFC has also been implicated [33–36]. The precise conditions that induce prefrontal activation during mental arithmetic are not well understood [37], but could be associated with working memory [33, 36], mental stress [18, 38], or other general cognitive operations that are instrumental, but not specific, to mental arithmetic [32, 33]. With respect to the music imagery task, we exploit the emotional component of music. Music is known to elicit [39, 40] and enhance [41] intense emotional responses that activate brain regions believed to be associated with emotional behaviors, including the PFC [42] and specifically, the orbitofrontal and frontopolar areas [43, 44]. The prefrontal hemodynamic response to subject-selected imagined singing has been previously observed in healthy participants via functional magnetic resonance imaging (fMRI) [45]. The prefrontal response to covert singing has also been observed in an individual with ALS, in the totally locked-in state (TLS), using optical topography [46]. Because hidden Markov models (HMMs) proved to be effective for the left- and right-hand classification problem [24], we explore their application to cognitive task classification.

2. Materials and methods

2.1. Participants

Ten able-bodied adults (four male, mean age = 26.2 ± 6.9) were recruited from the staff/students at Bloorview Kids Rehab (Toronto) to participate in this study. Potential subjects were self-selected according to the following inclusion/exclusion criteria.

- Subjects must not have any metabolic, cardiovascular, respiratory, psychiatric, or drug- or alcohol-related condition that could affect either the measurements or their ability to follow the experimental protocol.
- Subjects must have normal, or corrected-to-normal, vision.
- Subjects must enjoy music and feel that performing imagery of self-selected musical pieces could elicit a positive emotional response.

Subjects were asked to refrain from consuming coffee or alcohol, or smoking cigarettes, for several hours prior to the experimental sessions. Ethical approval was obtained from Bloorview Kids Rehab, and participants provided signed consent.

2.2. Instrumentation

A multichannel frequency-domain NIRS instrument (Imagent Functional Brain Imaging System from ISS Inc., Champaign, IL) was used for signal acquisition. Ten NIR sources and three photomultiplier tube detectors were secured against the forehead via a specially designed polyurethane headband. The ten sources were grouped into five pairs, each containing one 690 nm and one 830 nm source, so that a given location could be probed by both wavelengths simultaneously. With this

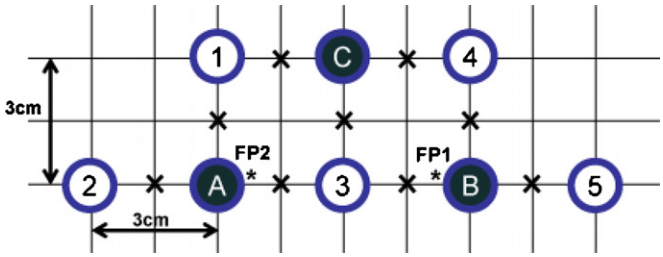

Figure 1. Source–detector configuration. Each solid circle represents a detector, while each open circle represents a source-pair comprising one 690 nm and one 830 nm source fibre. The detectors and source-pairs are spaced 3 cm apart. A '×' denotes a point of interrogation, while '∗' denotes the approximate FP1 and FP2 positions of the International 10–20 System.

source–detector configuration, we interrogated nine discrete locations within a trapezoidal area of 27 cm^2, as shown in figure 1. The source–detector pattern was centred on the forehead such that the centre line of sources/detectors was in line with the participant's nose, and the bottom row of sources/detectors was sitting just above the eyebrows. Detectors A and B lie approximately over the FP1 and FP2 positions of the International 10–20 System. In terms of optode placement, there is currently no standardized placement scheme for NIRS measurements [20] although it is believed that 3 cm is the ideal source–detector separation for measuring cortical hemodynamics [47]. As this distance increases, the optical signal weakens so that for separations exceeding 5 cm, the signal may be unusable [48]. In the given configuration, we considered only those signals arising from source–detector pairs (henceforth referred to as 'channels') with a separation of 3 cm, for a total of 18 channels (i.e. 3 detectors × 3 source-pairs per detector × 2 wavelengths per source-pair).

Light from the sources, modulated at 110 MHz, was delivered to the forehead via 400 μm diameter optical fibres and returned to the photomultiplier tube detectors by 3 mm diameter optical fibres. The detector amplifiers were modulated at 110.05 Hz, resulting in a cross-correlation (or heterodyning) frequency of 5 kHz at which data were recorded. The light sources were cyclically switched such that no two sources were on simultaneously. For each data collection cycle (i.e. one complete sequence through all ten sources), each source remained on for eight periods of the cross-correlation frequency (i.e. 1.6 ms), within which time eight acquisitions were made. The fast Fourier transform (FFT) was applied to the average of the eight waveforms to obtain estimates of ac intensity, dc intensity and phase delay for each channel. The resultant sampling frequency was 31.25 Hz (i.e. each source was sampled 31.25 times per second).

2.3. Experimental protocol

NIRS signals were collected from each participant as they performed two different cognitive tasks—music imagery and mental arithmetic. During the music imagery task, subjects

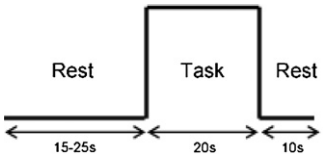

Figure 2. Single trial timing diagram. The initial rest interval varied between 15 and 25 s to reduce participant ability to anticipate the start of the task interval.

were asked to select three songs which they felt would self-elicit a strong emotional response. They were asked to rate each song on five-point scales of valance (1 = very unpleasant, 3 = neutral, 5 = very pleasant) and arousal (1 = very calming, 3 = neutral, 5 = very exciting), the objective being that the chosen songs would rate highly on each scale. When performing music imagery, participants were instructed to not just passively imagine the tune and/or lyrics, but to make an effort to feel the emotion elicited by the piece. It has been suggested that incorporating this self-monitoring element in an emotional induction task can result in an increase in the prefrontal hemodynamic response as compared to more passive emotional tasks [49]. During mental arithmetic, participants performed a sequence of simple mathematical calculations. These calculations always began with the subtraction of a small number (between 4 and 13) from a three digit number (both given, and different for each trial), and continued throughout the task interval with successive subtractions of the small number from the result of the previous subtraction (e.g. $967-13 = 954, 954-13 = 941, 941 - 13 = 928$, etc).

Fifty-one trials of each task, namely mental arithmetic and music imagery, were recorded per participant, over three experimental sessions (17 trials of each task per session). In addition, 51 baseline trials (17 per session) were recorded in which the participant performed no cognitive task but remained in a resting state throughout the entire trial. These trials were collected for use in a future study and were not considered in the present discussion. For each session, all trials of a given task were performed first (task 1), followed by all the trials of the remaining task (task 2), with the baseline trials dispersed randomly throughout the entire session (the random order of baseline trials was different for the three sessions). For even-numbered participants, task 1 corresponded to mental arithmetic and task 2 to music imagery, while the reverse was true for odd-numbered participants. The 51 trials of each session were performed in five blocks (four blocks of 10 trials and one block of 11 trials), between which participants were given a short break.

During each experimental session, the participant sat on a chair in a darkened, quiet room facing a computer screen on which task cues were visually presented. Each trial consisted of a 15–25 s rest interval, a 20 s task interval and a final 10 s rest interval. Figure 2 depicts the timing diagram of a single trial. During the resting state, subjects were instructed to relax, and to mentally recite the alphabet very slowly and calmly. This slight load was used to stabilize the prefrontal activity during the resting intervals, as suggested in [27]. The first rest interval

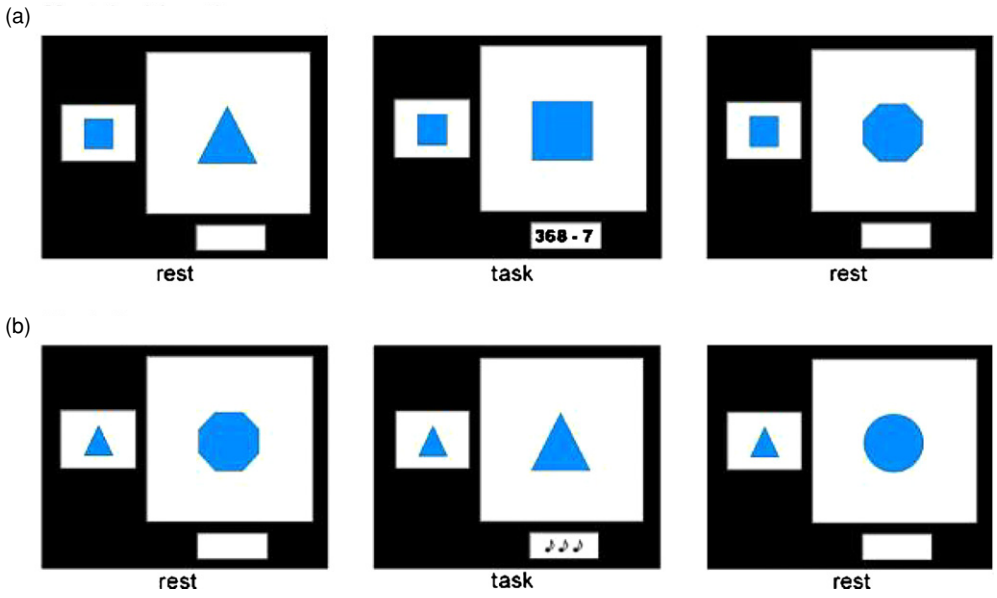

Figure 3. Sample trial display screens for (a) mental arithmetic and (b) music imagery. The participant was instructed to remain in a resting state when the pictures were not matching (which occurred during the rest intervals), and to perform the indicated task when the pictures matched (which occurred during the task interval).

was varied from 15 to 25 s to reduce the participant's ability to anticipate the start of the task interval.

To provide cues for the participant to transition between the rest and task states, the experiment was designed as a picture matching task. Two pictures of common geometric shapes were presented on the screen, the smaller of which remained constant throughout the trial while the other changed at the beginning of each interval. Participants were told to remain in a resting state whenever the pictures were not matching (which occurred during the rest intervals), and to perform the indicated task when the pictures matched (which occurred during the task interval). During the task interval, the task to be performed was indicated via a small box on the screen—for music imagery trials this box displayed music notes, while for mental arithmetic trials the equation the subject was to perform was displayed. Figures 3(a) and (b) depict representative display screens for each interval of a mental arithmetic trial and a music imagery trial, respectively.

2.4. NIRS data pre-processing

When classifying activity, the light intensity signals can be used directly, or can first be mathematically converted to haemoglobin concentrations. Both methods have appeared in the literature—both [27] and [50] classified light intensity directly, while [24] and [20] performed the concentration conversions prior to classification. Neither method has been shown to be more discriminating than the other [50]. Because it is less computationally intensive and therefore may be more conducive to future adaptation to an online system, we chose to classify the light intensity signals directly.

Prior to classification, the raw ac light intensity signals were low-pass filtered to mitigate physiological noise due, primarily, to respiration (0.2–0.3 Hz) [3], cardiac signals (0.8–1.2 Hz) and the Mayer wave (approximately 0.1 Hz) [51]. We employed a wavelet filter that performed a ten-level decomposition using a Daubechies-12 wavelet. The filtered signals were reconstructions retaining just the approximation signal and the last four detail signals, given the knowledge that hemodynamic activity is predominantly low-frequency (peak response has been observed approximately 5–8 s post-stimulus [11, 19, 21]). Wavelet filtering has been previously suggested for functional NIRS signals [51].

For each of the 18 channels under consideration, the pre-processed ac light intensity in the period 2–20 s after the start of the task interval was extracted. Each 18 s segment (which at a sampling frequency of 31.25 Hz was 562 data points in length) was then normalized against its own mean intensity and scaled. These time series segments ($x_k(t)$, where k denotes the channel number) were concatenated to form a per-trial 18 $\times$ 562 observation matrix, $\mathbf{U}$, to be used for HMM training/testing. Each column of $\mathbf{U}$ represents the 18-dimensional feature vector, $\boldsymbol{\mu}_t$, for a given sampling point, t. The vector $\boldsymbol{\mu}_t$ comprises the ac light intensity measured at each of the 18 channels. Note that due to a transient problem with the instrumentation that corrupted the data from one of the channels (but left the other channels unaffected), the feature vectors for four of the subjects were 17-dimensional:

$$\mathbf{U} = \begin{Bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_{18}(t) \end{Bmatrix} = \begin{Bmatrix} \boldsymbol{\mu}_1 & \boldsymbol{\mu}_2 & \cdots & \boldsymbol{\mu}_{562} \end{Bmatrix}.$$

2.5. Pattern classification

In a Markov model, the states are directly observable, and therefore the state transition probabilities are the only model parameters. A hidden Markov model is a statistical model that examines a Markov process in which the states are not directly

observable (thus the adjective 'hidden'), but depend on the observable outputs. An observation probability distribution is associated with each state, thus allowing information about the state sequence to be inferred from the output sequence (in the present discussion, the matrix **U**). Given a particular model, there are two different problems that can be solved: (1) determining the most likely state sequence from a given observation sequence and (2) evaluating the probability of a given observation sequence [52]. For this application, in which the objective was to distinguish activation due to mental arithmetic from that due to music imagery, we were interested in the latter problem. In brief, we created a model for each of the two tasks, and then classified trials as mental arithmetic or music imagery based on which model was more likely to have generated the trial data.

Hidden Markov models are useful for modelling sequential data, such as those arising from time series measurements. It is because of their ability to capture temporal information that we chose to employ HMMs in this classification problem. We anticipated that detailed information about the unique spatio-temporal activation patterns of each task would allow for the most reliable classification. This is substantiated by the results reported by Sitaram *et al* [24]. They found that HMMs performed significantly better than support vector machines in the classification of left- and right-hand motor imagery (as measured by NIRS), and suggested that this was because 'a probability network like the HMM might model the dynamic nature of the hemodynamic time series more effectively' [24].

Here, we will discuss only those basic HMM principles needed to understand the methods employed in this study. The reader is referred to [52] for a detailed discussion of the theory, implementation and practical application of HMMs.

An HMM, representing an observation vector $\boldsymbol{\mu}_t$, is completely characterized by the following parameters.

(1) The number of discrete states, Q.
(2) A state transition probability matrix, $\mathbf{A} = \{a_{ij}\}$, of transition probabilities between states i and j.
(3) The initial state distribution vector, $\boldsymbol{\pi}$.
(4) The vector of observation probability distributions in each state j, denoted as $\mathbf{b} = \{b_j(\boldsymbol{\mu})\}$ where $j = 1, \ldots, Q$. Commonly, Gaussian mixture models (GMM), given by a weighted sum of M component Gaussian densities, are employed as observation probability distributions.

In this study, unique HMMs were modelled and tested for each participant. Due to the variability in hemodynamic response patterns among individuals, the need for subject-specific classifier training has been recognized in previous NIRS-BCI studies [24, 50]. The complete data set for an individual (51 trials each of mental arithmetic and music imagery) was divided into training and test sets (using an 80/20 split per class). An HMM was modelled for each of the two tasks using the complete set of observation vectors from the training trials for that task, denoted as $\left\{\mathbf{U}_{\text{train}}^{\text{task}}\right\}$ (where task denotes either math or music). During training, the parameters **A** (state transition probabilities), **b** (output distribution parameters) and $\boldsymbol{\pi}$ (initial state probabilities) were optimized for each HMM using the expectation-maximization algorithm [52, 53]. The parameters $\boldsymbol{\pi}$ and **A** were initialized such that all states and all transitions between states, respectively, were equally probable. For the parameter **b**, the k-means clustering algorithm [53] was used to obtain the initialization values for the weights, means and covariances (full) of the M Gaussian components in the GMM. The parameters Q and M need to be determined *a priori*. In this study, different combinations of these parameters were explored such that subject-specific classifiers could be developed. All combinations of Q and M that yielded a training ratio (ratio of number of training points to number of estimated parameters) [54] greater than 10 were tested. These parameter combinations were as follows: $Q = 2$, $M = \{1, 2, 3, 4, 5, 6\}$; $Q = 3$, $M = \{1, 2, 3, 4\}$; $Q = 4$, $M = \{1, 2, 3\}$ and $Q = 5$, $M = \{1\}$.

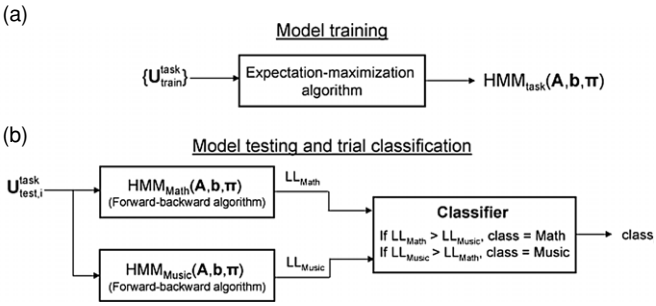

Figure 4. Training and classification procedure. (a) Model training: an HMM was modelled for each of the two tasks using the complete set of observation vectors from the training trials for that task, denoted as $\{\mathbf{U}_{\text{train}}^{\text{task}}\}$ (where task denotes either math or music). (b) Model testing and trial classification: for each test trial, i, the log-likelihood (LL) of each model generating the trial observation matrix, denoted by $\mathbf{U}_{\text{test},i}^{\text{task}}$, was determined using the forward–backward algorithm described in [52]. The predicted class was the one corresponding to the model that yielded the greater LL value.

During the classification phase, each trial from the test set (including both mental arithmetic and music imagery trials) was tested against each of the two HMMs (i.e. HMM_{math} for mental arithmetic and $\text{HMM}_{\text{music}}$ for music imagery). For each test trial, i, the log-likelihood of each model generating the test data, denoted by the observation matrix $\mathbf{U}_{\text{test},i}^{\text{task}}$, was determined using the forward–backward algorithm described in [52, 53]. The HMM yielding the highest log-likelihood value represented the model from which the observed data most likely arose, and thus the trial could be classified as belonging to either the mental arithmetic or music imagery task. On the rare occasion (less than 2% of cases) that log-likelihoods for each of the two models were equal, the predicted class was randomly assigned. Classification accuracy was evaluated for each participant using fivefold cross-validation. Figures 4(a) and (b) depict the model training and classification schemes, respectively. The publicly accessible HMM toolbox for Matlab [55] was used for the simulations.

3. Results

The results for the best-performing model parameters (Q and M) for each participant are given in table 1. The table

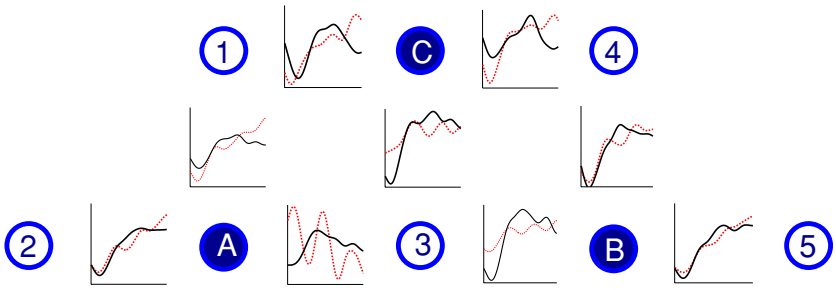

Figure 5. Light intensity versus time plots showing the hemodynamic response over the task interval (2–20 s post-stimulus) at each of the nine interrogation locations (corresponding to the points indicated by a '×' in figure 1), for music imagery (solid black) and mental arithmetic (dashed red). For simplicity, only the signals from the 830 nm sources are shown, but note that the 690 nm signals exhibit similar patterns. For each task, the signals shown are the average over all 51 trials for a single participant (participant 2).

Table 1. Per-participant classification accuracies and optimized HMM parameters.

	Model parameters		Classification accuracy (%)		
Participant no	Q	M	Mean	SD	p
1	2	4	71.7	6.6	0.0018
2	3	3	86.4	7.9	0.0005
3	4	1	80.2	6.1	0.0004
4	2	4	76.5	4.2	0.0001
5	4	1	82.1	9.2	0.0015
6	2	1	80.3	4.0	0.0001
7	3	2	80.4	7.9	0.001
8	4	3	77.0	5.7	0.0004
9	2	2	77.3	6.2	0.0006
10	2	5	60.5	7.3	0.0323
	Average:		77.2 ± 7.0		

reports these parameters as well as the average accuracy and standard deviation over the fivefold cross-validation for each participant. The average accuracy of the classifier over all participants was 77.2%. *T*-tests showed that the classification rates for all individuals were significantly higher than chance ($p < 0.0323$). The individual *p*-values are reported in table 1.

4. Discussion

4.1. Differentiating between mental arithmetic and music imagery

In this study, we attempted to classify activity in the PFC resulting from two different cognitive tasks, specifically mental arithmetic and music imagery. The encouraging classification results obtained using pre-processed ac light intensity signals and an HMM for each class warrant further investigation of a two-choice NIRS-BCI paradigm based on the classification of different prefrontal cognitive tasks.

Perhaps the reason we were able to differentiate between mental arithmetic and music imagery is because the tasks elicit slightly different spatial and temporal activation patterns within the PFC, which were captured by the HMMs. This hypothesis is substantiated by figure 5 which shows, for each task, the hemodynamic response over the task interval at each of the nine interrogation locations (the signals shown are the average of the 51 trials for each task, for a single subject). Note the differences in the response between the two tasks.

Two individuals—participants 1 and 10—had accuracies that were substantially lower than the mean. It was noted during fitting of the instrumentation at the experimental sessions that the signal levels for participant 1 were atypically low, even after attempts were made to improve coupling of the sources and detectors to the forehead. Also, during the first experimental session, this participant complained of severe discomfort due to the headband, and the experimenter had to remove and replace the apparatus three different times during the course of the session. These complications may explain the lower classification accuracy achieved for this participant. No such circumstances were noted for participant 10, so the lower accuracy in this case could potentially be related to the HMM parameters. Though a number of different combinations of the parameters Q and M were tested, it is possible that this set did not include the optimal values for this particular participant. To investigate additional values for Q and M for this participant, while still maintaining an acceptable training ratio, more data would be required.

4.2. Participant-specific HMMs

In this study, we chose to optimize not only the **A**, **b** and $\boldsymbol{\pi}$ parameters on a per-participant basis, but to optimize Q and M as well. The parameter Q represents the number of states defined by the observations. It is likely that throughout the course of the task interval, the participant did not jump immediately from a resting state to an active state, but rather that there was an initial period of increasing activity. Furthermore, during this 'ramping up' stage, and once a peak level of activity was reached, it is unlikely that the participant was able to sustain a uniform pattern of activity, but rather that there were irregularities in the level and spatial distribution of the activity within the task interval. These irregularities, along with possible extraneous activity due to mind-wandering, could result in a wider variety of different states. The intermediate activity between rest and peak response will vary significantly between individuals, and thus it is not surprising that the individually trained HMMs for the different participants would be optimally defined by

different numbers of states, Q. For similar reasons, we would also expect inter-subject variability in the number of Gaussian components, M, defining the observation distribution.

4.3. Comparing to past work

The accuracies achieved here are lower than those reported in [24] for the classification of left- and right-hand motor imagery by HMMs (although it is higher than the accuracy of 73.1% they achieved using support vector machines to perform the classification), but this was not unexpected. It is known that movement, and the imagery of movement, elicits activity in the portion of the motor cortex contralateral to the side of movement. In the motor imagery classification problem, the NIRS signals were collected bilaterally over the motor cortex, and thus the problem was one of differentiating activation in distinctly different brain regions (i.e. the left and right motor cortices). Though this is certainly non-trivial, there is an added complexity in our problem in which the goal was to differentiate activation occurring due to different tasks in the same brain region. Also important to note is the fact that the present study had twice the number of subjects as the motor imagery study. Because the tasks we have investigated require only cognitive capability, they may be a more appropriate alternative for use in a BCI for individuals with severe motor disabilities, for whom motor imagery tasks may be difficult, or even impossible, to perform. Further work must be done to investigate the feasibility of classifying activity due to mental arithmetic and music imagery in individuals from the target population. The prospects are promising, however, as these tasks have already been shown to be individually detectable by NIRS in individuals with ALS [27, 46].

4.4. Two-choice paradigm

We have designed the algorithm described in section 2.5 to be suitable for use in a two-choice BCI operating under a synchronized control paradigm [56]. The goal would be to collect, for each task, a sample of data with which to initially tune the HMM parameters offline. The user would then begin learning to operate the BCI online. There would be system-defined control intervals during which the user would be asked to perform either music imagery or mental arithmetic to indicate one of two choices—yes or no, for example. The system would evaluate the user's brain activity only during these defined control periods, with the HMM classifier being used to determine which task was performed, and thus how the user responded (i.e. yes or no) during a given control period. At the end of each session, as more data became available, the model parameters would be updated.

4.5. NIRS-BCI: challenges

Though a majority of BCI research has focused on electroencephalography (EEG) as the measurement modality, NIRS has been gaining attention due to the many advantages it affords. It has good spatial resolution (within the order of a centimetre depending on probe geometry), and is not affected by interference from electrophysiological artefacts such as EMG, ECG and EOG [21]. Furthermore, the thought processes required to intentionally generate the signals are relatively simple, and the signals more directly reflect a cognitive function than their EEG counterparts [19]. For BCI applications, NIRS is also advantageous over fMRI due to its portability and lower cost.

However, there are also some challenges associated with NIRS in the context of BCI applications. One example is speed—the information transfer rate of an NIRS-BCI will be significantly limited due to the inherent latency of the hemodynamic response. The peak response has been shown to occur approximately 5–8 s post-stimulus [11, 19, 21]. With the technique presented here, a 20 s activation period would be needed to classify the activity as resulting from mental arithmetic or music imagery, which would translate to a maximum information transfer rate of just 3 bits min^{-1}. While this transfer rate is comparable to the rates achieved in previous NIRS-BCI studies [20], one could potentially accelerate communication by reducing the window size for classification [57]. Further work must be done to determine the temporal window which will provide the optimal trade-off between speed and accuracy. BCIs based on the electroencephalographic response, which is inherently much faster than the hemodynamic response, can achieve transfer rates between 10 and 25 bits min^{-1} [58].

Also challenging is the fact that, as noted by [51], the power of NIRS signals is dominated by both physical and physiological noise—introduced by subject movement, heart rate, respiration and other physiological trends like the Mayer wave—making functional activity very difficult to identify without significant pre-processing. We mitigated the effects of systemic blood flow through the use of a low-pass wavelet filter, but we did not incorporate any techniques for motion artifact removal. In this study, subject movement was intentionally limited, but in a practical situation where the user could have involuntary movements, motion artefact would have to be explicitly addressed. Techniques such as adaptive filtering, Weiner filtering, or principal component analysis could offer potential solutions for efficient motion artifact removal [51].

4.6. Study limitations

It should also be noted that the conditions of this study were controlled beyond just restricting subject movement. The experimental sessions were conducted in a quiet room with minimal extraneous auditory and visual distractions. It is unclear what the effects of such distractions would have been on both the cortical activation patterns and the user's ability to focus on the required tasks. Further work must be done to ensure that acceptable performance levels can be maintained in more natural, uncontrolled environments. Furthermore, in this study, the subject's only role was to perform the given task when cued to do so. In this case, all of his or her attention could be focused on performance of the task. However, in a real functional application this would not be the case. Using the BCI in a practical way would introduce additional cognitive load, associated with, for example, spelling a word, making

a decision, concentrating on the changing display, etc. Task performance would likely be affected when the task-related activity must be generated amid this unrelated mental activity.

Another limitation of this study was that data were collected over just three experimental sessions. We do not know how the differentiability of the two tasks would change over time, as the participant gained proficiency with the tasks. Would this proficiency allow the user to elicit greater, more characteristic patterns of activity for each task, or would it cause the task to be less cognitively demanding, resulting in a reduction in activity levels and a more general activation pattern? Longer-term studies must be conducted to investigate these issues.

Finally, the mental arithmetic task may be difficult for children. Further work must be done to investigate the differentiability of other cognitive tasks (e.g. verbal tasks) that may be more suitable for those who are unable to perform the arithmetic task.

5. Conclusion

In summary, prefrontal NIRS signals due to two different cognitive tasks, namely, mental arithmetic and music imagery, were measured in a controlled laboratory environment and classified using maximum likelihood classifiers based on task-specific hidden Markov models. Classification accuracies well in excess of chance were attained for all participants. While these results are encouraging, issues such as the slow hemodynamic response time, motion artefacts and user distraction need to be considered in further investigations of NIRS brain–computer interfaces driven by cognitive tasks.

Acknowledgments

This study was made possible by the Canada Research Chairs program, the Ontario Centres of Excellence and the Natural Sciences and Engineering Research Council of Canada. The authors would like to acknowledge Mr Duluxan Sritharan for his assistance with the data collection for this study.

References

[1] Tai K, Blain S and Chau T 2008 A review of emerging access technologies for individuals with severe motor impairments *Assist. Technol.* **20** 204–19

[2] Wolpaw J R, Birbaumer N, Heetderrks W J, McFarland D J, Peckham P H, Schalk G, Donchin E, Quatrano L A, Robinson C J and Vaughan T M 2000 Brain–computer interface technology: a review of the first international meeting *IEEE Trans. Rehabil. Eng.* **8** 164–73

[3] Franceschini M A, Fantini S, Toronov V, Filiaci M E and Gratton E 2000 Cerebral hemodynamics measured by near-infrared spectroscopy at rest and during motor activation *Proc. Optical Society of America In Vivo Optical Imaging Workshop* pp 73–80

[4] Huppert T J, Hoge R D, Diamond S G, Franceschini M A and Boss D A 2006 A comparison of BOLD, ASL, and NIRS hemodynamic responses to motor stimuli in adult humans *NeuroImage* **29** 368–82

[5] Huppert T J, Hoge R D, Franceschini M A and Boas D A 2005 A spatial-temporal comparison of fMRI and NIRS hemodynamic responses to motor stimuli in adult humans *Proc. SPIE* **5693** 191–202

[6] Jasdzewsky G, Strangman G, Wagner J, Kwong K K, Poldrack R A and Boas D A 2003 Differences in the hemodynamic response to event-related motor and visual paradigms as measured by near-infrared spectroscopy *NeuroImage* **20** 479–88

[7] Kuboyama N, Nabetani T, Shibuya K, Machida K and Ogaki T 2004 The effect of maximal finger tapping on cerebral activation *J. Physiol. Anthropol. Appl. Hum. Sci.* **23** 105–10

[8] Boas D A, Strangman G, Culver J P, Hoge R D, Jasdzewsky R A, Poldrack R A, Rosen B R and Mandeville J B 2003 Can the cerebral metabolic rate of oxygen be estimated with near-infrared spectroscopy? *Phys. Med. Biol.* **48** 2405–18

[9] Strangman G, Culver J P, Thompson J H and Boas D A 2002 A quantitative comparison of simultaneous BOLD fMRI and NIRS recordings during functional brain activation *NeuroImage* **17** 719–31

[10] Mehagnoul-Schipper D J, van der Kallen B F W, Colier W N J M, van der Sluijs M C, van Erning L J T O, Thijssen H O M, Oeseburg B, Hoefnagels W H L and Jansen R W M M 2002 Simultaneous measurements of cerebral oxygenation changes during brain activation by near-infrared spectroscopy and functional magnetic resonance imaging in healthy young and elderly subjects *Hum. Brain Mapp.* **16** 14–23

[11] Benaron D A *et al* 2000 Noninvasive functional imaging of human brain using light *J. Cereb. Blood. F. Met.* **20** 469–77

[12] Herrmann M J, Plichta M M, Ehlis A C and Fallgatter A J 2005 Optical topography during a go–nogo task assessed with multi-channel near-infrared spectroscopy *Behav. Brain Res.* **160** 135–40

[13] Herrmann M J, Walter A, Ehlis A C and Fallgatter A J 2006 Cerebral oxygenation changes in the prefrontal cortex: effects of age and gender *Neurobiol. Aging* **27** 888–94

[14] Quaresima V, Ferrari M, Torricelli A, Spinelli L, Pifferi A and Cubeddu R 2005 Bilateral prefrontal cortex oxygenation responses to a verbal fluency task: a multichannel time-resolved near-infrared topography study *J. Biomed. Opt.* **10** (011012)1–12

[15] Boecker M, Buecheler M M, Schroeter M L and Gauggel S 2007 Prefrontal brain activation during stop-signal response inhibition: an event-related functional near-infrared spectroscopy study *Behav. Brain Res.* **176** 259–66

[16] Yang H, Zhou Z, Liu Y, Ruan Z, Gong H, Luo Q and Lu Z 2007 Gender difference in hemodynamic responses of prefrontal area to emotional stress by near-infrared spectroscopy *Behav. Brain Res.* **178** 172–76

[17] Tanida M, Sakatani K, Takano R and Tagai K 2004 Relation between asymmetry of prefrontal cortex activities and the autonomic nervous system during a mental arithmetic task: near infrared spectroscopy study *Neurosci. Lett.* **369** 69–74

[18] Tanida M, Katsuyama M and Sakatani K 2007 Relation between mental stress-induced prefrontal cortex activity and skin conditions: a near-infrared spectroscopy study *Brain Res.* **1184** 210–6

[19] Coyle S, Ward T, Markham C and McDarby G 2004 On the suitability of near-infrared (NIR) systems for next-generation brain–computer interfaces *Physiol. Meas.* **25** 815–22

[20] Coyle S M, Ward T E and Markham C M 2007 Brain–computer interface using a simplified functional near-infrared spectroscopy system *J. Neural Eng.* **4** 219–26

[21] Sitaram R, Hoshi Y and Guan C 2005 Near infrared spectroscopy based brain–computer interface *Proc. SPIE* **5852** 434–42

[22] Pfurtscheller G, Neuper C, Schlögl A and Lugger K 1998 Separability of EEG signals recorded during right and left motor imagery using adaptive autoregressive parameters *IEEE Trans. Rehabil. Eng.* **6** 316–25
[23] Pfurtscheller G, Neuper C, Guger C, Harkam W, Ramoser H, Schlögl A, Obermaier B and Pregenzer M 2000 Current trends in Graz brain–computer interface (BCI) research *IEEE Trans. Rehabil. Eng.* **8** 216–19
[24] Sitaram R, Zhang H, Guan C, Thulasidas M, Hoshi Y, Ishikawa A, Shimizu K and Birbaumer N 2007 Temporal classification of multichannel near-infrared spectroscopy signals of motor imagery for developing a brain–computer interface *NeuroImage* **34** 1416–27
[25] Cramer S C, Orr E L R, Cohen M J and Lacourse M G 2007 Effects of motor imagery training after chronic, complete spinal cord injury *Exp. Brain Res.* **177** 233–42
[26] Conson M, Sacco S, Sara M, Pistoia F, Grossi D and Trojano L 2008 Selective motor imagery defect in patients with locked-in syndrome *Neuropsychologia* **46** 2622–28
[27] Naito M, Michioka Y, Ozawa K, Ito Y, Kiguchi M and Kanazawa T 2007 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.* **90** 1028–37
[28] Ogata H, Mukai T and Yagi T 2007 A study on the frontal cortex in cognitive tasks using near-infrared spectroscopy *Proc. IEEE EMBS (Lyon, France)* pp 4731–4
[29] Utsugi K, Obata A, Sato H, Katsura T, Sagara K, Maki A and Koizumi H 2007 Development of an optical brain–machine interface *Proc. IEEE EMBS (Lyon, France)* pp 5338–41
[30] Bauernfeind G, Leeb R, Wriessnegger S C and Pfurtscheller G 2008 Development, set-up and first results for a one-channel near-infrared spectroscopy system *Biomed. Tech./Biomed. Eng.* **53** 36–43
[31] Ayaz H, Izzetoglu M, Bunce S, Heiman-Patterson T and Onaral B 2007 Detecting cognitive activity related hemodynamic signal for brain computer interface using functional near infrared spectroscopy *Proc. IEEE EMBS (Lyon, France)* pp 342–5
[32] Richter M M, Zierhut K C, Dresler T, Plichta M M, Ehlis A C, Reiss K, Pekrun R and Fallgatter A J 2009 Changes in cortical blood oxygenation during arithmetical tasks measured by near-infrared spectroscopy *J. Neural Transm.* **116** 267–73
[33] Gruber O, Indefrey P, Steinmetz H and Kleinschmidt A 2001 Dissociating neural correlates of cognitive components in mental calculation *Cereb. Cortex* **11** 350–59
[34] Menon V, Rivera S M, White C D, Glover G H and Reiss A L 2000 Dissociating prefrontal and parietal cortex activation during arithmetic processing *NeuroImage* **12** 357–65
[35] Remy F, Mirrashed F, Campbell B and Richter W 2004 Mental calculation impairment in Alzheimer's disease—a functional magnetic resonance imaging study *Neurosci. Lett.* **358** 25–8
[36] Zago L, Pesenti M, Mellet E, Crivello F, Mazoyer B and Tzourio-Mazoyer N 2001 Neural correlates of simple and complex mental calculation *Neuroimage* **13** 314–27
[37] De Pisapia N, Slomski J A and Braver T S 2007 Functional specializations in lateral prefrontal cortex associated with the integration and segregation of information in working memory *Cereb. Cortex* **17** 993–1006
[38] Tanida M, Katsuyama M and Sakatani K 2008 Effects of fragrance administration on stress-induced prefrontal cortex activity and sebum secretion in the facial skin *Neurosci. Lett.* **432** 157–61
[39] Altenmüller E, Schürmann K, Lim V K and Parlitz D 2002 Hits to the left, flops to the right: different emotions during listening to music are reflected in cortical lateralisation patterns *Neuropsychologia* **40** 2242–56
[40] Krumhansl C L 1997 An exploratory study of musical emotions and psychophysiology *Can. J. Exp. Psychol.* **51** 336–53
[41] Baumgartner T, Esslen M and Jancke L 2006 From emotion perception to emotion experience: emotions evoked by pictures and classical music *Int. J. Psychophysiol.* **60** 34–43
[42] Boso M, Politi P, Barale F and Enzo E 2006 Neurophysiology and neurobiology of the musical experience *Funct. Neurol.* **21** 187–91
[43] Blood A J, Zatorre R J, Bermudez P and Evans A C 1999 Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions *Nat. Neurosci.* **2** 382–7
[44] Blood A J and Zatorre R J 2001 Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion *Proc. Natl Acad. Sci. USA* **98** 11818–23
[45] Kleber B, Birbaumer N, Veit R, Trevorrow T and Lotze M 2007 Overt and imagined singing of an Italian aria *NeuroImage* **36** 889–900
[46] Fuchino Y, Nagao M, Katura T, Bando M, Naito M, Maki A, Nakamura K, Hayashi H, Koizumi H and Yoro T 2008 High cognitive function of an ALS patient in the totally locked-in state *Neurosci. Lett.* 85–89
[47] Yamamoto T, Maki A, Kadoya T, Tanikawa Y, Yamada Y, Okada E and Koizumi H 2002 Arranging optical fibres for the spatial resolution improvement of topographical images *Phys. Med. Biol.* **47** 3429–40
[48] Gratton G, Brumback C R, Gordon B A, Pearson M A, Low K A and Fabiani M 2006 Effects of measurement method, wavelength, and source–detector distance on the fast optical signal *NeuroImage* **32** 1576–90
[49] Herrmann M J, Ehlis A C and Fallgatter A J 2003 Prefrontal activation through task requirements of emotional induction measured with NIRS *Biol. Psychol.* **64** 255–63
[50] Luu S and Chau T 2009 Decoding subjective preference from single-trial near-infrared spectroscopy signals *J. Neural Eng.* **6** 1–8
[51] Matthews F, Pearlmutter B A, Ward T E, Soraghan C and Markham C 2008 Hemodynamics for brain–computer interfaces *IEEE Signal Process. Mag.* **25** 87–94
[52] Rabiner L R 1989 A tutorial on hidden Markov models and selected applications in speech recognition *Proc. IEEE* **77** 257–86
[53] Bishop C M 2006 *Pattern Recognition and Machine Learning* (New York: Springer Science+Business Media) pp 610–35
[54] Falk T H 2005 An improved GMM-based voice quality predictor *Proc. Interspeech (Lisbon, Portugal)* pp 2733–6
[55] Murphy K 2005 Hidden Markov Model (HMM) Toolbox for Matlab http://www.cs.ubc.ca/~murphyk/Software/HMM/hmm.html
[56] Mason S G and Birch G E 2000 A brain-controlled switch for asynchronous control applications *IEEE Trans. Biomed. Eng.* **47** 1297–307
[57] Tai K and Chau T 2009 Single-trial classification of NIRS signals during emotional induction tasks: Towards a corporeal machine interface *J. Neuroeng. Rehabil.* **6** 39
[58] Wolpaw J R, Birbaumer N, McFarland D J, Pfurtscheller G and Vaughan T M 2002 Brain–computer interfaces for communication and control *Clin. Neurophysiol.* **113** 767–91