Institute of Biomaterials
& Biomedical Engineering
UNIVERSITY OF TORONTO

January 3, 2010

Dear Members of the RESNA SSPC Committee:

I am writing to confirm that Ms. Sarah Power has completed the majority of the work presented in her paper, "Classifying Prefrontal Activity due to cognitive tasks using NIRS".

Please do not hesitate to contact me if you have any further questions.

Yours truly,

Tom Chau, PhD, PEng
Associate Professor & Graduate Coordinator, Clinical Engineering
Leader, NSERC CREATE: Academic Rehabilitation Engineering (CARE) Program
Institute of Biomaterials & Biomedical Engineering
University of Toronto

FACULTIES OF APPLIED SCIENCE AND ENGINEERING, MEDICINE, AND DENTISTRY
Rosebrugh Building, 164 College Street, Room 407, Toronto, Ontario, M5S 3G9 Canada
Tel: +1 416 946-8258 • Fax: +1 416 978-4317 • www.ibbme.utoronto.ca

CLASSIFYING PREFRONTAL ACTIVITY DUE TO COGNITIVE TASKS USING NIRS

Sarah D. Power and Tom Chau

Institute of Biomaterials and Biomedical Engineering, University of Toronto
Bloorview Research Institute, Holland Bloorview Kids Rehabilitation Hospital

INTRODUCTION

Many individuals with severe motor disabilities lack sufficient motor control to use speech, gestures, and movements to communicate and control external devices. Brain-computer interfaces (BCI) are controlled through brain activity alone, and may provide such individuals with an alternative means of communication and control. Near-infrared spectroscopy (NIRS) is an optical imaging technology that has recently been investigated as a potential measurement modality for non-invasive BCIs.

Generally, a user controls a BCI output by consciously eliciting distinct, repeatable patterns of activity in a particular region of the brain. This is usually done through performing different mental tasks. These characteristic patterns of activity are detected and interpreted by the system, which then produces the appropriate command signals to control a connected external device (e.g. computer mouse) in the way the user intended.

The ability to accurately detect and classify the patterns of activation associated with different mental tasks (and different BCI commands) is essential to BCI development. The aim of the present study was to investigate the feasibility of differentiating between two different cognitive tasks in the prefrontal cortex (PFC), namely mental arithmetic (MA) and music imagery (MI), using NIRS measurements. If these two activities can be successfully classified with acceptable accuracy, it could lead to the development of a 2-state synchronized NIRS-BCI.

METHODS

Participants

Ten able-bodied adults (four male, mean age = 26.2 +/- 6.9 years) were recruited from the staff and students at Holland Bloorview Kids Rehabilitation Hospital (Toronto, Canada). Individuals were excluded from participation if they had any condition that could affect either the measurements or their ability to follow the experimental protocol. Additionally, participants had to enjoy music and feel that performing imagery of self-selected musical pieces could elicit a positive emotional response. Ethical approval was obtained from Holland Bloorview Kids Rehabilitation Hospital and the University of Toronto. Participants provided signed consent.

Instrumentation

Signals were acquired using a multichannel frequency-domain NIRS instrument (Imagent Functional Brain Imaging System from ISS Inc., Champaign, IL). Ten NIR sources and three photomultiplier tube detectors were clipped into a flexible headband and secured against the participant's forehead. The ten sources were grouped into five pairs, each containing one 690 nm and one 830 nm source. Nine locations within a 27cm^2 trapezoidal area were probed as shown in Figure 1. The headband was placed on the forehead such that the bottom row of optodes sat just above the eyebrows, and the center row of optodes was in line with the participant's nose. Data was sampled at 31.25 Hz.

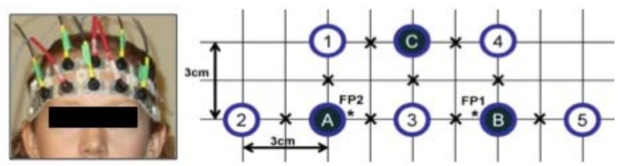

Figure 1: Source–detector configuration. Each solid circle represents a detector, while each open circle represents a source-pair consisting of one 690 nm and one 830 nm source fiber.

Experimental Protocol

NIRS signals were collected from each participant as they performed trials of MA and MI. For the MI task, participants were asked to pre-select several songs that they felt would elicit a strong positive (i.e., happy) emotional response in them. When performing MI, they were instructed to make an effort to feel the emotion elicited by the piece, rather than just passively reciting the lyrics or humming the tune. For the MA task, 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, 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). The initial calculation to begin the sequence was always given.

A single trial consisted of a 15–25 s rest interval, followed by a 20 s task interval, followed by a final 10 s rest interval. The duration of the initial rest interval was varied to reduce the participant's ability to anticipate the start of the task interval. During the resting state, subjects were instructed to relax, and to mentally recite the alphabet slowly. This slight load was meant to stabilize the prefrontal activity during the resting intervals. 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. Participants were told to remain in a resting state when the pictures did not match and to perform the indicated task when the pictures matched (which occurred during the rest and task intervals, respectively). During the task interval, the task to be performed was indicated in a small box below the pictures. Figure 2 shows representative display screens for each interval of a MA trial.

Fifty-one trials of each task were recorded per participant over three experimental sessions (one third of trials for each task were recorded per session). Within a given session, participants performed either all the math trials followed by the music trials (odd numbered participants), or vice versa (even numbered participants).

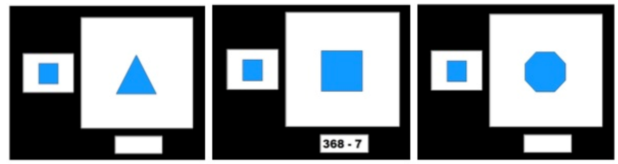

Figure 2: Example display screens and timing for mental arithmetic trial.

DATA ANALYSIS

Pre-processing

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) [1], cardiac activity (0.8–1.2 Hz) and the Mayer wave (approximately 0.1 Hz) [2]. We employed a wavelet filter that performed a ten-level decomposition using a Daubechies-12 wavelet. Given the knowledge that hemodynamic activity is predominantly low-frequency (peak response has been observed approximately 5–8 s post-stimulus [3]), the filtered signals were reconstructions retaining just the approximation signal and the last four detail signals.

For each of the 18 channels under consideration, the filtered ac light intensity in the period 2–20 s after the start of the task interval was extracted. Each 18 s segment was then normalized against its own mean intensity and scaled.

Classification

A hidden Markov model is a statistical model examining a Markov process in which the states are not directly observable, but rather are dependent on the observable outputs. Information about the state sequence can be gleaned from the output through the observation probability distribution associated with each state. Two different problems can be solved given a particular model: (1) determining the most likely state sequence from a given observation sequence, and (2) evaluating the probability of a given observation sequence [4]. For this MA vs. MI classification application, we were interested in the latter problem. In brief, we created a model for each of the two tasks, and then classified trials as MA or MI based on which model was more likely to have generated the trial data.

An HMM, representing an observation vector $\mathbf{u}_t$, is completely characterized by the following parameters:

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

In this case, the observation vectors were 18-dimensional, comprising the values of each of the 18 pre-processed ac light intensity signals at a single sampling point during the task interval. With a sampling rate of 31.25 Hz and interval duration of 18s, there was a total of 562 observation vectors per trial.

Classification accuracy was evaluated for each participant using five-fold cross-validation. For each fold, an HMM was modeled for each of the two tasks using the complete set of observation vectors from the training trials for that task, denoted as U_{train}^{task} (where task denotes either MA or MI). During training, the parameters A, b and Π were optimized for each HMM using the expectation-maximization algorithm [4, 5]. The parameters Π 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 [5] 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 are not optimized through model training, but rather need to be determined a priori. In this study, different combinations of these parameters were explored so that subject-specific classifiers could be developed. All combinations of Q and M that allowed for a training ratio (ratio of number of training points to number of estimated parameters) greater than 10 were considered. These combinations were as follows: Q=2, M={1-6}; Q=3, M={1-4}; Q=4, M={1-3}, and Q=5, M={1}.

During the testing phase, each trial from the test set was tested against each of the two HMMs (i.e., HMM_{MA} and HMM_{MI}). For each test trial, *i*, the forward-backward algorithm described in [4] was used to determine the log-likelihood (LL) of each model having generated the test data, denoted by the observation matrix $U_{test,i}^{task}$. The HMM yielding the highest log-likelihood value represented the model from which the observed data most likely arose, and thus the trial was classified as either MA or MI. Figures 3(a) and (b) depict the model training and classification schemes, respectively.

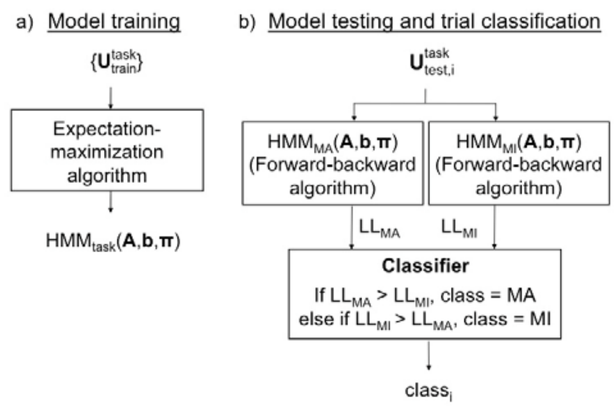

Figure 3: a) Model training and b) classification procedure.

RESULTS

The results for the best-performing model parameters (Q and M) for each participant are listed in Table 1. These parameters are reported, along with the average accuracy and standard deviation over the five-fold cross-validation for each participant. Across participants, the average accuracy achieved was 77.2%. The classification rates for all individuals were significantly greater than chance, as determined by t-tests ($p < 0.0323$).

DISCUSSION

We were able to distinguish between mental arithmetic and music imagery with an accuracy

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

Subject #	Model Parameters		Classification Accuracy (mean +/- std)	p-value
	Q	M		
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	

exceeding chance, thus clearly the tasks must elicit different spatial and/or temporal activation patterns within the PFC (at least for these individuals) that were captured by the HMMs. This hypothesis is substantiated by the plot of the hemodynamic response over the task interval at each of the nine interrogation locations, which can be found in [6].

We anticipate that the described algorithm will be suitable for use in a two-choice NIRS-BCI operating under a synchronized control paradigm [7]. There would be system-defined control intervals during which the user would be asked to perform either MI or MA to indicate one of two choices (e.g., yes or no). 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.

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. However, due to inter-session variability in user response patterns and in sensor placement, it may prove ineffective to have a single, a priori, classifier training session. The classifiers may have to be re-trained prior to each use using newly-collected data (either exclusively, or as a supplement to the previously-collected data). Of course, for reasons of practicality the amount of data collected at each session will have to be limited. Further work will involve adapting the developed algorithm for online use, and designing an effective classifier/user training protocol.

For a more thorough discussion of the results, the reader is directed to [6].

CONCLUSIONS

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 subject-specific maximum-likelihood HMM classifiers warrant further investigation of a two-choice NIRS-BCI paradigm based on the classification of different prefrontal cognitive tasks.

ACKNOWLEDGEMENTS

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.

REFERENCES

[1] M. A. Franceschini, S. Fantini, V. Toronov, M. E. Filiaci, and E. Gratton, "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, 2000.

[2] F. Matthews, B. A. Pearlmutter, T. E. Ward, C. Soraghan, and C. Markham, "Hemodynamics for brain-computer interfaces," *IEEE Signal Process. Mag.*, vol. 25, pp. 87-94, 2008.

[3] D. A. Benaron et al., "Noninvasive functional imaging of human brain using light," *J. Cereb Blood F. Met.*, vol. 20, pp. 469-77, 2000.

[4] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. IEEE*, vol. 77, pp. 257-86, 1989.

[5] C. M. Bishop, *Pattern recognition and machine learning*, Springer Science+Business Media, New York, 2006.

[6] 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.*, vol. 7(2), 9pp, 2010.

[7] S. G. Mason, and G. E. Birch, "A brain-controlled switch for asynchronous control applications," *IEEE Trans. Biomed. Eng.*, vol. 47, pp. 1297–307, 2000.