

Taking NIRS-BCIs Outside the Lab: Towards Achieving Robustness Against Environment Noise

Tiago H. Falk, *Member, IEEE*, Mirna Guirgis, Sarah Power, and Tom T. Chau, *Senior Member, IEEE*

Abstract—This paper reported initial findings on the effects of environmental noise and auditory distractions on the performance of mental state classification based on near-infrared spectroscopy (NIRS) signals recorded from the prefrontal cortex. Characterization of the performance losses due to environmental factors could provide useful information for the future development of NIRS-based brain-computer interfaces that can be taken beyond controlled laboratory settings and into everyday environments. Experiments with a hidden Markov model-based classifier showed that while significant performance could be attained in silent conditions, only chance levels of sensitivity and specificity were obtained in noisy environments. In order to achieve robustness against environment noise, two strategies were proposed and evaluated. First, physiological responses harnessed from the autonomic nervous system were used as complementary information to NIRS signals. More specifically, four physiological signals (electrodermal activity, skin temperature, blood volume pulse, and respiration effort) were collected in synchrony with the NIRS signals as the user sat at rest and/or performed music imagery tasks. Second, an acoustic monitoring technique was proposed and used to detect startle noise events, as both the prefrontal cortex and ANS are known to involuntarily respond to auditory startle stimuli. Experiments with eight participants showed that with a startle noise compensation strategy in place, performance comparable to that observed in silent conditions could be recovered with the hybrid ANS-NIRS system.

Index Terms—Ambient noise, autonomic nervous system, hidden Markov models, music imagery, near-infrared spectroscopy.

I. INTRODUCTION

IT is often extremely difficult for individuals with severe motor disabilities to communicate or interact with their environment. These individuals commonly make use of access

technologies to translate their intentions into useful control signals. Present technologies such as eye gaze [1], head [2], and tongue control devices [3], however, require some degree of voluntary motor control and are unsuitable for locked-in individuals who lack any functional motor skills. Although such individuals often have full cognitive awareness, their nonfunctioning bodies fail to respond appropriately to their intentions. To this end, brain-computer interfaces (BCI) have been investigated as alternative access solutions and have shown promising results in controlled environments [4], [5].

Near-infrared spectroscopy (NIRS) has been recently investigated as a noninvasive means of assessing functional activity in the brain via measured hemodynamic responses [6] with spatial resolutions (in the order of a few centimeters) similar to those obtained with electroencephalography (EEG)-based systems [7]. Unlike EEG, however, NIRS measurements do not require cumbersome skin preparation and electrode gels. Moreover, the thought processes required to intentionally generate the NIRS signals are relatively simple, and more directly reflect cognitive function [8]. NIRS determines the properties of the brain tissue by transmitting near-infrared electromagnetic radiation (650–950 nm wavelengths) through the skull and comparing the intensities of the returning and incident light. Certain functional mental activities are known to elicit specific spatial and temporal activation patterns in particular regions of the brain [9]. The metabolic demand of mental activity causes a change in regional concentrations of oxygenated and deoxygenated hemoglobin. Since the fraction of light absorbed versus the fraction transmitted is dependent on the concentrations of these chromophores, NIRS can be used to assess hemodynamic responses in regions such as the motor cortex (using motor imagery tasks [6]) and the prefrontal cortex (using music imagery [10], [11], mental arithmetic [12] or preference tasks [13]). Automated detection of hemodynamic changes can be used to control a binary BCI [9].

To improve the performance of NIRS-based BCI technologies, significant efforts have been placed on the detection and removal of *physiological* (e.g., respiration and cardiac) noise and motion artifacts from the recorded spectroscopy signals [9], [14], [15]. Little research, however, has been conducted to investigate the effects of *environmental* noise and auditory distractions on system performance. While the work described in [16] suggested that environmental noise had insignificant effects on the performance of a P300-based BCI, recent research has suggested that significant performance degradation is observed for NIRS-BCIs based on prefrontal cortex activation (e.g., via music imagery) [17]. This reduction in system accuracy may be due to factors such as 1) the performance of mental tasks being severely affected by environmental noise [18]–[22] or 2) the pre-

Manuscript received May 11, 2010; revised July 12, 2010 and September 02, 2010; accepted September 08, 2010. Date of publication September 27, 2010; date of current version April 08, 2011. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada, in part by the Ontario Centres of Excellence, in part by the Ward Family Foundation, and in part by the Canada Research Chairs program.

T. H. Falk was with the Bloorview Research Institute and Holland Bloorview Kids Rehabilitation Hospital, Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, ON M4G 1R8, Canada. He is now with is now with the Institut National de la Recherche Scientifique (Energy, Materials, and Telecommunications), University of Quebec, Montreal, QC H5A 1K6, Canada (e-mail: tiago.falk@ieee.org).

M. Guirgis, S. Power, and T. T. Chau are with the Bloorview Research Institute and Holland Bloorview Kids Rehabilitation Hospital, Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, ON M4G 1R8, Canada.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TNSRE.2010.2078516

frontal cortex being involved in the processing of distracting auditory stimuli [23], [24].

In order to develop robust technologies that can be taken beyond controlled experimental settings and into everyday surroundings, the effects of environmental factors need to be ascertained and automated compensation strategies need to be implemented. In this paper, initial findings are reported that show NIRS-based systems achieving chance-level performance in noisy environments. In order to take the first step towards the development of noise-robust tools and to enhance the classification of mental activity under noisy conditions, two strategies are investigated. The first strategy proposes the incorporation of signals harnessed from the autonomic nervous system (ANS) as complementary information to brain hemodynamic responses for improved mental activity classification. This is motivated by previous research that has suggested that physiological responses such as changes to skin electrodermal activity [25], skin temperature [26], heart rate [27], respiration rates [28], and salivary pH [29] can be elicited by mental imagery tasks similar to those used to control NIRS-BCI systems. The reader is referred to [30] and the references therein for a more comprehensive review of volitional ANS control techniques. It is hypothesized that user intention can be better modeled once the two modalities are combined.

Notwithstanding, inclusion of ANS-harnessed signals may not significantly improve mental state classification performance in the presence of auditory startle stimuli. Much like the prefrontal cortex, the ANS is known to (involuntarily) respond to auditory startle stimuli [31], [32]. To overcome this shortcoming, an acoustic monitoring technique termed “environment sniffing” [33] is proposed and used to suppress classification errors caused by auditory startle reflexes.

The remainder of this paper is organized as follows. Section II describes the materials and methods used in the experiments. Section III introduces the proposed ANS-NIRS paradigm which incorporates complementary information from the ANS as well as an acoustic startle noise detection and compensation strategy. Sections IV–VI present experimental results, discussion, and conclusions, respectively.

II. MATERIALS AND METHODS

A. Participants

Ten able-bodied adults (three male) were initially recruited but only eight participants (mean age of 31.5 ± 10.8 years) concluded the study. Ethics approval was obtained from the affiliated institutes and participants provided written consent. Candidates with metabolic, cardiovascular, respiratory, hearing, or drug- or alcohol-related conditions that could affect either the measurements or their ability to follow the experimental protocol (see Section II-C) were excluded.

B. Instrumentation

1) *NIRS Signals*: A multichannel frequency-domain NIRS instrument (Imagent Functional Brain Imaging System from ISS Inc., Champaign, IL) was used for signal acquisition. Sixteen light sources, eight at 690 nm and eight at 830 nm, were used along with two photomultiplier tube detectors. Light

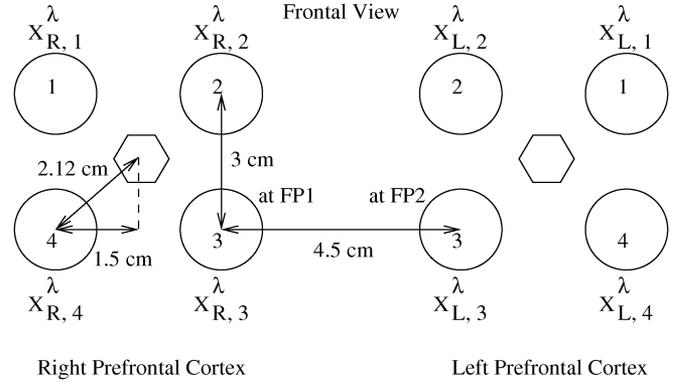


Fig. 1. Positioning of source pairs (circles) and detectors (hexagons). Each source pair represents one $\lambda = 690$ nm and one $\lambda = 830$ nm source. The signal from each source is represented as $x_{p,i}^{\lambda}$, $i = 1, \dots, 4$ and $p = L$ or R , for left and right sides, respectively. Source pairs at position 3 are located, anatomically, at the FP1 and FP2 positions of the International 10–20 system.

from the sources delivered 110 MHz-modulated light to the forehead via 400 μm -diameter optical fibers and returned to the photomultiplier tube detectors by 3-mm-diameter optical fibres. The returning light was demodulated at a cross-correlation, or heterodyning, frequency (CCF) of 5 kHz. Light sources were cyclically switched to avoid two sources being on simultaneously. For a complete data collection cycle, which consisted of one complete sequence through all sixteen sources, each source remained on for eight periods of the CCF (1.6 ms) separated by a two-period break (0.4 ms) to avoid any overlap. Thus the effective sampling rate was 31.25 Hz per full data collection cycle. A fast Fourier transform was applied to the average of each source-detector pair of the 16 waveforms to obtain the ac (relative amplitude at CCF), dc (relative amplitude at 0 Hz), and phase delay (phase delay of demodulated light compared to a reference) components. Here, only the dc component is used as it has been shown to contain similar information to the ac signals, however, with a higher signal-to-noise ratio [34].

The 16 source fibers were grouped in pairs such that each pair contained one source at each wavelength allowing a single location to be probed by both wavelengths simultaneously. Both right and left prefrontal cortices were probed by positioning four source pairs around one detector on each side of the participant’s forehead (see Fig. 1); emitters in position 3 on each side were approximately located over the FP1 and FP2 positions of the International 10–20 System. The four source pairs positioned on each side were located 2.12 cm apart from their respective detectors; previous research has suggested that such distances suffice for hemodynamic response probing in the prefrontal cortex [35], [36]. For the sake of notation, the raw dc temporal signals obtained from each source are represented as $x_{p,i}^{\lambda}(t)$ where i indicates the source position ($i = 1, \dots, 4$), p the (L)left or (R)right side, λ indicates the wavelength (690 or 830 nm), and t indicates discrete time.

2) *ANS Signals*: Four physiological signals were recorded simultaneously using a ProComp Infiniti multimodality encoder (Thought Technology, Montreal, QC, Canada) at a sampling frequency of 256 Hz. Electrodermal activity was measured from two 10-mm-diameter Ag-AgCl surface electrodes attached with adhesive collars on the medial phalanges of the

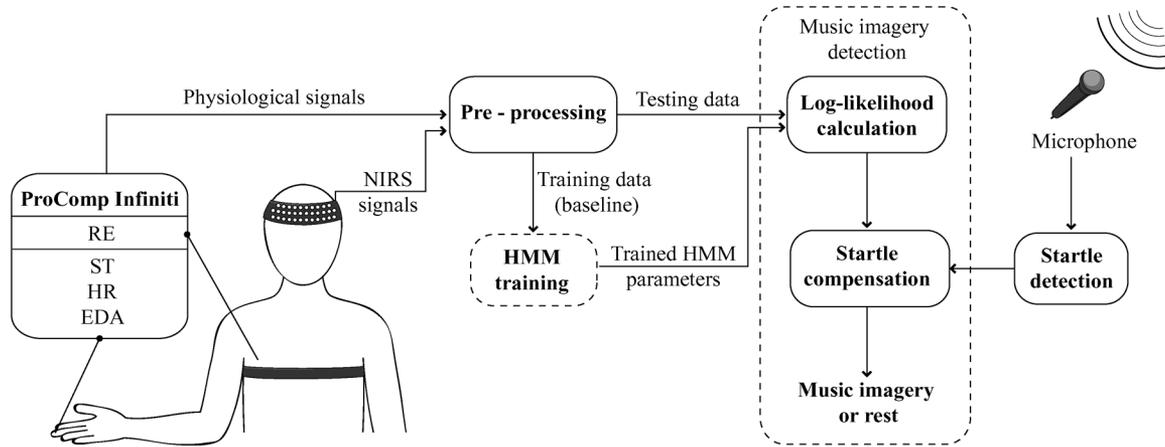


Fig. 2. Overall design of the proposed ANS-NIRS mental activity classification system. Hemodynamic responses in the prefrontal cortex were measured via NIRS in synchrony with four independent ANS-harnessed physiological signals, namely EDA, ST, blood volume pulse (subsequently converted to HR), and RE. Startle noise detection and compensation was performed to improve system performance under noisy conditions.

index and middle fingers. Medial phalanges were chosen as they represent a region of the skin containing a high density of sweat glands. A constant 0.5 V was applied between the two electrodes. Skin temperature was measured using a thermal sensor on the distal phalange of the fifth finger. Heart rate was computed from the interbeat intervals of the blood volume pressure waveform obtained with a photoplethysmograph sensor. Electrodermal, temperature, and blood volume pressure sensors were all secured to the nondominant hand of the participant. Lastly, respiration effort was measured by positioning a piezoelectric belt around the thoracic area; stretching due to expansion and contraction of the chest was converted into voltage changes.

C. Protocol

This study involved ten sessions on two separate days. Four sessions consisted of baseline trials, each 130 s in duration, with participants sitting at rest performing no mental task. Participants were instructed to focus on their breathing and to clear their minds. Half of these sessions were performed in a controlled silent environment and the other half in the presence of five startle stimuli presented in random order at 20-, 45-, 65-, 90-, and 110-s time instances. Noise stimuli were all approximately 800 ms in duration and their respective loudness intensities in decibel (dB) were: dog barking (80 dB), glass breaking (91 dB), door slamming (83 dB), person coughing (79 dB), and sneezing (82 dB).

The remaining six sessions, each 220 s in duration, consisted of the participants alternating between baseline (i.e., rest) and performing music imagery tasks. Each of these sessions began and ended with a rest interval. All intervals (rest and music imagery) were 20 s long. The participant was cued to start and stop music imagery with a light tap on the arm. Twenty-second activation windows were chosen to take into account the task initiation delay, which we have found in preliminary experiments to be as long as 5 s, as well as the inherent physiological latency of the hemodynamic responses, which can be in the order of 5–10 s [6]. For similar reasons, 20-s rest intervals were chosen. Such choice of stimulus pattern, however, can reinforce slow

third-order blood pressure waves and its subharmonics (circa 0.05–0.1 Hz), thus a wavelet-based denoising algorithm was used, as described in Section III-A.

Before each session, participants were instructed to select a minimum of two songs of the same emotional valence to be used throughout the session. Of these six sessions, two were performed in a controlled silent environment, two in the presence of continuous background noise (e.g., “humming” of an air conditioner), and two in the presence of the five aforementioned startle stimuli. In the latter case, the time instances in which the startle stimuli were presented was also randomly chosen between four possible options:

- 1) 31, 88, 111, 149, and 191 s;
- 2) 28, 82, 91, 128, and 151 s;
- 3) 10, 31, 72, 150, and 189 s;
- 4) 14, 89, 111, 170, and 190 s.

For all six sessions, participants sat comfortably at a desk and were equipped with a custom-built NIRS headband made from polyurethane and the four physiological signal sensors described in Section II-B2.

III. ANS-NIRS SYSTEM DESIGN

Fig. 2 illustrates the overall design of the proposed ANS-NIRS mental activity classification system (henceforth termed “ANS-NIRS system”). Hemodynamic responses in the prefrontal cortex were measured via NIRS in synchrony with four independent ANS-harnessed physiological signals, namely electrodermal activity (EDA), skin temperature (ST), heart rate (HR), and respiration effort (RE). Signals were preprocessed (see Section III-A) and those collected during baseline trials (i.e., individual at rest, see Section II-C) were used to train hidden Markov models representative of the normative physiological and hemodynamic response of the individual at rest. A normalized log-likelihood measured was then used for classification; more specifically, for automated detection of music imagery events (using the music imagery test trials, see Section II-C), which in turn, could be used to control a binary BCI. Lastly, to compensate for startle noise effects, a microphone was used to detect startle events based

on techniques such as “environment sniffing” [33]; once startle noise events were detected, erroneous classifications could be suppressed. A detailed description of each processing module is described in the remainder of this section.

A. Signal Preprocessing and Feature Extraction

The raw physiological signals were downsampled to match the sampling frequency of the NIRS signals and then filtered using fifth order Butterworth lowpass filters with 0.2, 0.1, 1.2, and 0.3 Hz cutoff frequencies for electrodermal activity, skin temperature, heart rate, and respiration effort, respectively [37].

The raw dc NIRS temporal signals $x_{p,i}^\lambda(t)$, $t = 1, \dots, T$, where T is the signal length, $p = L, R$ and $i = 1, \dots, 4$ were filtered using wavelet-based filters to alleviate the effects of physiological noise (e.g., cardiac, 0.5–2 Hz, and respiration, 0.2–0.4 Hz) as well as the Mayer wave (~ 0.1 Hz) [9]. Wavelet-based filters have been shown previously to be effective for NIRS signal denoising [9]. Wavelet filters were based on a 12-level signal decomposition using the Daubechies wavelet. The resulting denoised signals $x_{p,i}^{\lambda*}(t)$ correspond to the reconstruction of the approximation wavelet coefficients and the last five detail coefficients.

Let \mathbf{O} denote the K -dimensional feature vector. For statistical models of resting hemodynamic responses and resting hemodynamic-ANS responses, the corresponding feature vectors were, respectively

$$\begin{aligned} \mathbf{O}_{\text{NIRS}} &= [\bar{x}_L^{690*} \bar{x}_R^{690*} \bar{x}_L^{830*} \bar{x}_R^{830*}] \\ \mathbf{O}_{\text{ANS-NIRS}} &= [s_{HRs} s_{RES} s_{EDA} s_{ST} | \mathbf{O}_{\text{NIRS}}] \end{aligned}$$

where the time dependence of each feature has been omitted for notational simplicity, s subscripted with HR , RE , EDA , and ST indicates the preprocessed signals for heart rate, respiration effort, electrodermal activity, and skin temperature, respectively, and “|” represents vector concatenation. Parameters $\bar{x}_p^{\lambda*}$ denote the position-averaged ($p = (L)$ eft or (R) ight) preprocessed dc signals given by

$$\bar{x}_p^{\lambda*} = \frac{1}{4} \sum_{i=1}^4 x_{p,i}^{\lambda*}. \quad (1)$$

B. Hidden Markov Models

A hidden Markov model (HMM) is a statistical model that represents a Markovian process wherein the observable outputs are dependent on the unobservable (hidden) states. HMMs have been used previously for NIRS- [6] and electroencephalography-based [38] classification of motor imagery tasks and are only discussed briefly here; the reader is referred to [39] for a more detailed discussion. Here, we use HMMs to capture the complex temporal signal interactions that arise once the user initiates the cognitive task.

Consider a Q -state HMM where the variable q_t indicates the HMM state at time t . As such, an HMM can be completely characterized by three entities $\boldsymbol{\lambda} = \{\boldsymbol{\pi}, \mathbf{A}, \mathbf{B}\}$, namely

- 1) the initial state distribution $\boldsymbol{\pi} = \{\pi_1, \dots, \pi_Q\}$, where $\pi_j = \Pr(q_0 = j)$;

- 2) the transition matrix ($\mathbf{A} = \{a_{i,j}\}$) where $a_{i,j} = \Pr(q_t = j | q_{t-1} = i)$ describes the probabilities of transitioning from state q_i to state q_j , $i \geq 1, j \leq Q$;
- 3) the observation probability distribution ($\mathbf{B} = \{b_j(\mathbf{O})\}$), $j = 1, \dots, Q$.

Commonly, M -component Gaussian mixture models (GMM) are used to model the output observation probabilities and are given by

$$b_j(\mathbf{O}) = \sum_{i=1}^M \alpha_{j,i} \mathcal{N}[\mathbf{O}, \boldsymbol{\mu}_{j,i}, \boldsymbol{\Sigma}_{j,i}] \quad (2)$$

where \mathbf{O} is the vector being modeled, $\alpha_{j,i} \geq 0$, $i = 1, \dots, M$ are the mixture weights for state q_j , with $\sum_{i=1}^M \alpha_{j,i} = 1$, and $\mathcal{N}[\mathbf{O}, \boldsymbol{\mu}_{j,i}, \boldsymbol{\Sigma}_{j,i}]$ are K -variate Gaussian densities with mean vector $\boldsymbol{\mu}_{j,i}$ and covariance matrix $\boldsymbol{\Sigma}_{j,i}$.

HMM model parameters $\boldsymbol{\lambda}$ are iteratively estimated with the expectation-maximization algorithm [39], [40] using training data vectors; the k -means algorithm [41] is used for parameter initialization. During training, the number of parameters that need to be estimated depends on the number of HMM states Q , type of HMM (e.g., fully connected, left–right), number of Gaussian components M , data dimensionality K , and GMM covariance matrix type (i.e., diagonal or full). More specifically, for fully-connected HMMs, the number of parameters, N , that need to be estimated during training is given by

$$N_{\text{full}} = Q \left(1 + Q + \frac{M}{2}(K^2 + 3K + 2) \right) \quad (3)$$

$$N_{\text{diag}} = Q(1 + Q + M(2K + 1)) \quad (4)$$

where the subscript indicates either full- or diagonal-covariance GMMs, respectively.

In order to avoid overfitting during training, the number of states, number of Gaussian components per state, and GMM covariance matrix type parameters are controlled such that the training ratio remains above a given threshold

$$\text{Training ratio} = \frac{N_{\text{train}}}{N_{\text{type}}} > \text{Threshold} \quad (5)$$

where N_{train} indicates the number of training vectors available for training. As a rule-of-thumb, the threshold is commonly set to be threshold = 10 [42]. In this study, full covariance GMMs were used to explore correlations between NIRS signals measured from neighboring channels. Hence, to guard against overfitting, the number of HMM states and Gaussian components per state were limited to 2–4 and 1–4, respectively; such values were consistent with previous HMM-based BCI studies (e.g., [6], [38]).

The NIRS and combined ANS-NIRS signals collected during the baseline trials *performed in silence* (see Section II-C) were used to train NIRS and ANS-NIRS HMMs, respectively, for each participant. Due to a sensor malfunction, the EDA signal was not collected for two participants. Different HMM configurations were explored given the known variations in NIRS signal levels and spatial distributions across individuals. These variations are due, in part, to inter-subject differences in mental alertness, familiarity with the procedure, and emotional intensity of user-selected songs [43].

C. Automatic Mental Task Classification

In order to take NIRS-BCIs outside the lab and into everyday clinical settings, a system should be able to accurately classify the two mental states, namely, rest and music imagery, both in silent and in noisy conditions. Future developments will investigate the use of automatically detected music imagery events as control signals for a binary BCI. In order to detect music imagery events, features are extracted over time (namely, $\mathbf{O}(l)_{\text{test,NIRS}}$ or $\mathbf{O}(l)_{\text{test,ANS-NIRS}}$) from the imagery test trials described in Section II-C, with a running window of length L , and scored against the reference HMMs using a normalized likelihood measure

$$\begin{aligned} LL(l) &= \Pr(\mathbf{O}(l)_{\text{test}}|\boldsymbol{\lambda}) \\ &= \frac{1}{L} \sum_q \pi_{q_0} \prod_{t=l}^{l+L-1} a_{q_{t-1},q_t} b_{q_t}(\mathbf{O}(t)_{\text{test}}) \end{aligned} \quad (6)$$

where $l = 0, \dots, T - L + 1$, window lengths ranging from $L = 1-15$ s are explored, and a window overlap of 0.5 s is used (i.e., LL is computed at a 2 Hz “sample rate”). For numerical stability, the log-likelihood version of (6) is used in the simulations described in Section IV. Different window lengths are explored to account for the user-specific latencies inherent in both the NIRS signals, which can range from 4–8 s [6], and the physiological signals, which can range from 1–15 s [44].

Higher (log-)likelihood values suggest hemodynamic and/or combined hemodynamic-ANS responses akin to those observed during the baseline trials (i.e., rest) used to train the HMMs. Lower likelihood values, in turn, indicate responses different from rest and suggest music imagery events. As such, a decrease in the likelihood function is expected during imagery periods and either an increase or a constant value is expected during the rest periods. The plot depicted in Fig. 3(a) shows a representative log-likelihood NIRS temporal series clearly illustrating the expected increases and decreases during rest (unshaded) and imagery (shaded) intervals, respectively.

In order to automatically classify music imagery events, a simple slope-change based classifier is used where sustained changes in slope (from positive to negative) are used to identify imagery events. To prevent the effects of mind wandering and motion artifacts which may produce a momentary positive-to-negative slope change, an imagery event is detected ($I(l) = 1$) if at time “ l ” the slope of the LL changes from positive to negative and remains negative for at least 5 s post-time “ l ”

$$I(l) = \begin{cases} 1, & \text{if } \begin{cases} \text{sgn}(LL(l-1) - LL(l-2)) > 0 \\ \text{sgn}(LL(l) - LL(l-1)) < 0 \text{ and} \\ (LL(l+P) - LL(l+P-1)) < 0 \\ \text{where } P = 1, \dots, 10 \end{cases} \\ 0, & \text{otherwise.} \end{cases} \quad (7)$$

Note that the proposed classification strategy, once incorporated into a BCI, will impose an upper bound on the system information transfer rate of 12 bits/min, which is somewhat lower than existing EEG-based BCIs [4]. For comparison purposes, the 20-s imagery window used in this experiment imposes a lower bound of 3 bits/min.

D. Environmental Noise Effects on Classification Performance

Previous studies have suggested that the performance of mental tasks can be severely affected by environmental noise [18]–[22]. As a consequence, environmental distractions are expected to affect the normative behaviour of NIRS-driven signals, thus hampering the performance of NIRS-based BCI technologies based on prefrontal cortex activation (e.g., via music imagery). Fig. 3(b), for example, illustrates the detrimental effects of continuous background noise on the normalized likelihood measure described by (6) for music imagery test data collected in noisy environments (see Section II-C). As shown in this plot, an increased number of slope changes are observed (relative to music imagery test data collected in silence, see Fig. 3(a) thus resulting in erroneous classifications. Moreover, for the example in Fig. 3(b), the timing of the mental task was also affected, as witnessed by the delay and advancement in the onset of music imagery observed in the last two imagery intervals, respectively.

Additionally, the prefrontal cortex has also been shown to be involved in the processing of auditory information, particularly of distracting (startling) auditory stimuli [23], [24]. Fig. 3(c), for example, illustrates the detrimental effects of intermittent startle noises (represented by dashed vertical lines in the plot) on the log-likelihood function for test (imagery) NIRS data. As observed, environmental effects would have detrimental consequences on the performance of an NIRS-based classification system. In order to develop robust technologies that can be taken beyond controlled experimental settings and into a user’s typical surroundings, such effects need to be taken into account and compensated for. In this study, we propose to incorporate ANS-harnessed physiological signals to improve classification performance in noisy environments.

As will be shown in Section IV, however, the gains attained by incorporating ANS-harnessed signals are more modest under startle noise conditions, since physiological signals are also sensitive to such distracting auditory stimuli [45]. This behaviour can be illustrated by the log-likelihood function computed for training (baseline) ANS-NIRS data under startle noise conditions, as shown in Fig. 3(d). The sustained decreases in log-likelihood slopes observed in the baseline training sequence following the “dog barking,” “glass breaking,” and “door slamming” startle events would suffice to generate erroneous classifications. Hence, to improve system performance, a startle noise detection and compensation scheme is proposed to suppress misclassifications resultant from involuntary startle-reflex responses.

E. Startle Noise Detection and Compensation

In order to compensate for detrimental acoustic startle noise effects (assumed to be additive in nature), a simple energy thresholding algorithm is employed similar to that used for temporal clipping detection in voice over internet [46]. A complete description of the method is beyond the scope of this paper and the reader is referred to [46] for further details; here, only a brief summary is presented. An airborne microphone is placed near the user (see Fig. 2) and the instantaneous log-energy is computed continuously. A candidate startle event is detected once the energy differential signal exceeds a predetermined

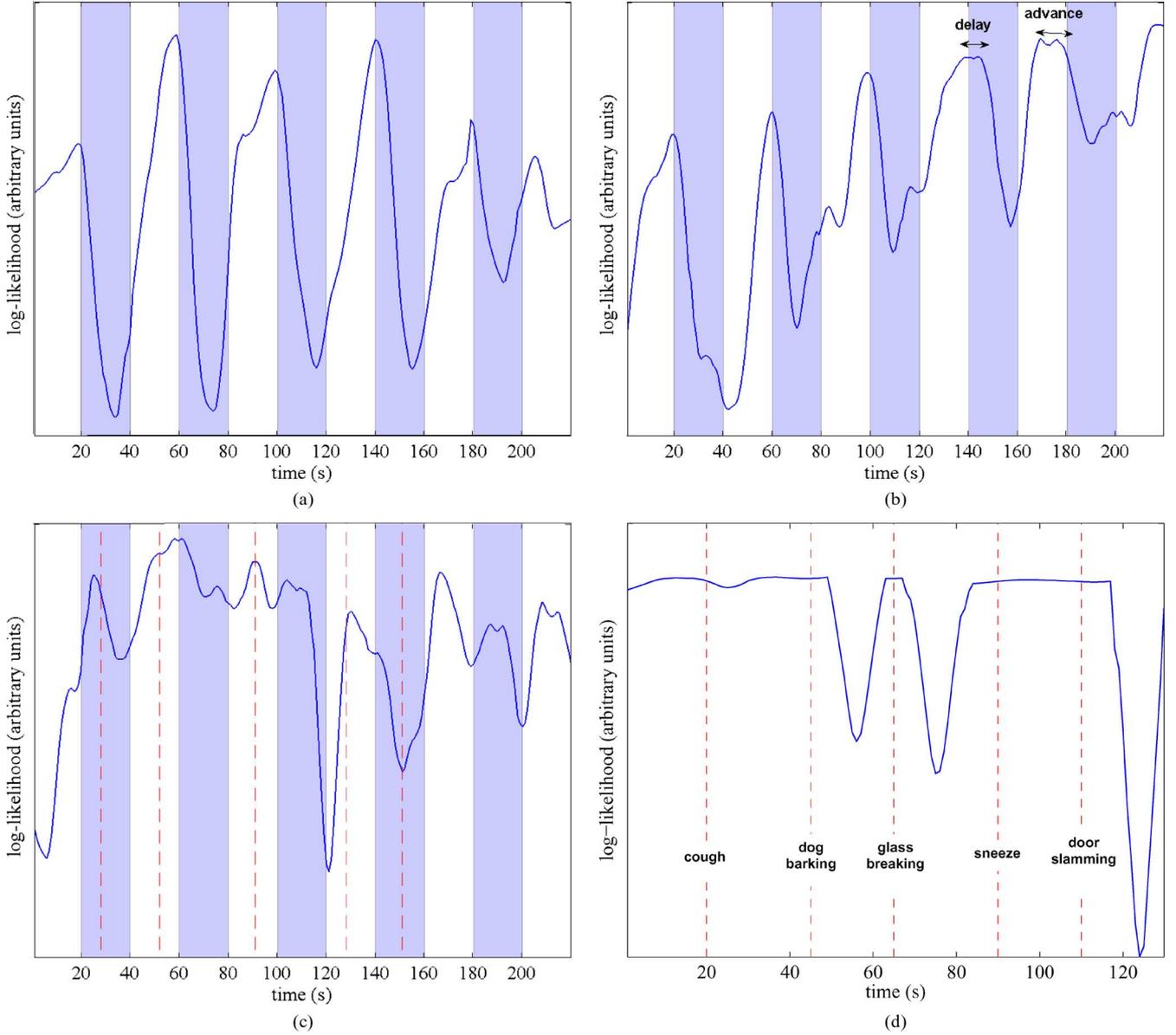


Fig. 3. Representative normalized log-likelihood temporal series computed for NIRS imagery test data recorded in (a) silence, and in noisy conditions with (b) constant background noise, and (c) intermittent startle noises. Shaded regions indicate music imagery and vertical dashed lines indicate startle event times. Subplot (d) depicts a representative normalized likelihood function for ANS-NIRS baseline data (*without* music imagery) in the presence of startle noises.

threshold, here empirically set to 50 dB. Fig. 4(a)–(c) depicts the waveform consisting of the five concatenated startle stimuli (a) used in this study, as well as the instantaneous energy (b), and the energy differential (c) with the empirically-set threshold depicted by the horizontal dotted line. As observed, all startle noises are correctly detected. It is important to emphasize that an adaptive threshold may allow for more accurate startle event detection as the user migrates from different physical environments, such as home and hospital; such investigation is left for future study.

Once acoustic startle noise events are detected, erroneous mental task classifications resultant from involuntary startle-reflex responses need to be suppressed. As observed from Fig. 3(d), however, the onset of the startle reflex response can occur approximately 1–7 s poststartle stimulus. As a consequence, startle noise compensation consists of suppressing

music imagery classification events that start within 5 s (empirically set) post-startle detection. On the other hand, if a startle is detected after a sustained decrease in the log-likelihood function has been observed for at least 2.5 s, suppression does not occur as it is assumed that the user was already in the process of music imagery while the startle event occurred. More specifically, let T_s denote the time of a detected startle event; hence, startle noise compensation consists of updating music imagery classification using the following rules:

$$\tilde{I}(l) = \begin{cases} 0, & \text{if } I(l) = 0 \\ 0, & \text{if } I(l) = 1, l < T_s + 5, \\ 1, & \text{if } I(l) = 1, l < T_s - 2.5 \\ 1, & \text{if } I(l) = 1, l > T_s + 5 \end{cases} \quad (8)$$

where \tilde{I} denotes the final imagery-versus-rest classification.

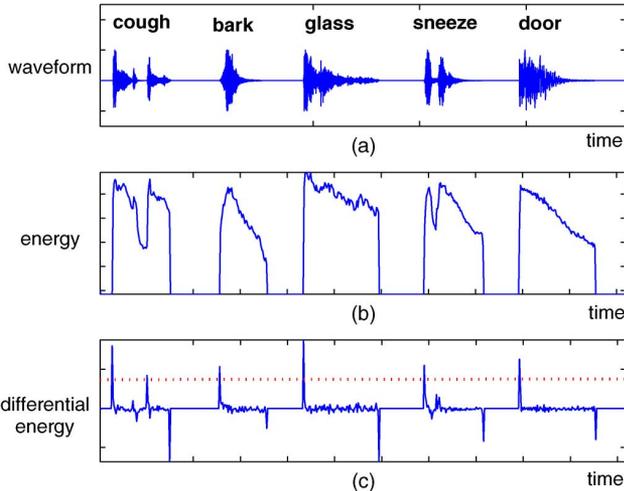


Fig. 4. Overview of startle noise detection. Subplots correspond to (a) temporal waveform, (b) instantaneous energy, and (c) energy differential. Startle noise detection is based on thresholding (horizontal dotted line) of the differential energy; here the threshold is empirically set to 50 dB.

IV. EXPERIMENTAL RESULTS

A. Performance Metrics

To objectively quantify system performance, three commonly used performance metrics are used, namely, classifier accuracy, sensitivity, and specificity. The last two measures are given by

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100\% \quad (9)$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \times 100\% \quad (10)$$

where TP and TN denote number of true positives and true negatives, respectively; FP and FN, in turn, represent the number of false positives and false negatives. The sensitivity measure relates to the percentage of correctly classified imagery events, whereas specificity relates to the percentage of correctly classified rest intervals.

For the developed system, a correctly classified imagery event (TP) occurs when a sustained 5-s slope decrease in the log-likelihood is observed within the first 13 s of a 20-s imagery interval. This window compensates for the latency in the hemodynamics-driven signals, which can extend to up to 8 s [6]; only one TP event is possible during a 20-s imagery period. If no activation is detected within this time period, an incorrectly classified rest interval (FN) is observed. Moreover, imagery events detected during expected rest intervals are deemed as false positives (FP) while expected rest intervals classified as so are termed true negatives (TN) with only one possible TN occurring per 20-s rest interval. It is important to emphasize that given the number of possible TPs and TNs available in our dataset, performance metrics need to surpass 63% to be considered significantly greater than chance ($p < 0.05$) [47].

B. Classification Accuracy in Controlled Silent Environments

Table I reports per-participant classifier sensitivity and specificity values obtained by the NIRS- and ANS-NIRS systems

under silent experimental conditions. To facilitate a user-centered BCI design [4], optimal HMM configurations (i.e., Q and M parameters) as well as log-likelihood window size l are obtained on a per-participant basis; such parameter values are also reported in the table. The results reported for subjects 7 and 8 are based only on a subset of the ANS signals due to an EDA sensor malfunction during data collection. As observed, once ANS-harnessed physiological signals are incorporated into the design, average gains of approximately 14% and 10% are observed in sensitivity and specificity, respectively. Moreover, lower values of l were needed with ANS-NIRS signals, thus allowed for faster automatic decoding of functional intent. Interestingly, since the EDA signal was not available for subjects 7 and 8, somewhat higher l was needed. Additionally, the average classifier accuracy observed over the eight participants was of 74.25% for NIRS and 83% for ANS-NIRS, an approximate 12% gain.

For practical BCI applications, customizing the HMM parameters for each user would require a calibration session with known imagery and rest intervals. This optimization process could be performed with minimal intervention by an outside party, similar to the calibration session required by commercially available speech recognizers. The use of only resting baseline data for training, as opposed to both baseline and imagery data commonly required by existing BCIs, allows the calibration session to be relatively short (order of tens of seconds) and thus unlikely to pose a burden on the user.

C. Classification Accuracy Under Noisy Conditions

Table II reports per-participant classifier performance in constant background and startle noise environments. As can be seen, if NIRS signals were used alone, performances levels near chance were attained (see Section IV-A). Once peripheral autonomic signals were incorporated in the system design, however, performances significantly greater than chance were obtained. An approximate relative performance gain of over 40% in both sensitivity and specificity was observed for the constant background noise scenario; such gains were shown to be significant (t -test, $p < 0.03$). Moreover, average sensitivity and specificity levels were not significantly different from those observed in silent conditions ($p > 0.5$, see Table I). Lastly, average classifier accuracies of 63.2% and 81.5% were attained for NIRS and ANS-NIRS systems, respectively.

For intermittent startle noises, the performance gains obtained over conventional NIRS by incorporating ANS-harnessed signals were more modest than with continuous background noise; gains of approximately 20.3% and 15.5% were observed in system sensitivity and specificity, respectively. System specificity, was deemed significantly different from that obtained in silent conditions ($p = 0.04$). Average classifier accuracies for NIRS and ANS-NIRS were of 63.6% and 75.5%, respectively. With the proposed startle noise detection and compensation technique in place, the specificity values of the NIRS- and ANS-NIRS systems increased to 70.8% and 78.6%, respectively; sensitivity values remained unchanged. The obtained specificity levels were not significantly different ($p > 0.5$) from those observed in silent conditions. Such findings suggest that the proposed classification strategy, once

TABLE I
CLASSIFICATION ACCURACY COMPARISON BETWEEN NIRS AND ANS-NIRS SYSTEMS IN SILENT CONDITIONS. COLUMNS LABELED “SENS,” “SPEC,” AND “GAIN” INDICATE, RESPECTIVELY, CLASSIFIER SENSITIVITY, SPECIFICITY, AND RELATIVE PERFORMANCE GAINS OBTAINED OVER CONVENTIONAL NIRS BY USING ANS-NIRS SIGNALS. OPTIMAL HMM PARAMETERS (Q AND M) AND LOG-LIKELIHOOD WINDOW SIZES (l) ARE ALSO REPORTED. RESULTS FOR THE LAST TWO SUBJECTS USE ONLY A SUBSET OF THE ANS SIGNALS

Participant number	NIRS					ANS-NIRS						
	Sens (%)	Spec (%)	Q	M	l (s)	Sens (%)	Gain (%)	Spec (%)	Gain (%)	Q	M	l (s)
1	80.0	92.0	4	1	7	83.3	4.2	100.0	8.7	2	2	1
2	100.0	58.0	2	1	3	100.0	0.0	91.7	58.0	4	2	2
3	80.0	58.0	2	2	10	77.5	-3.1	66.7	14.9	4	3	6
4	50.0	100.0	4	1	10	80.0	60.0	91.7	-8.3	4	2	4
5	80.0	83.0	2	1	10	100.0	25.0	50.0	-39.8	4	3	6
6	90.0	75.0	4	1	5	100.0	11.1	100.0	33.3	4	2	4
7*	80.0	75.0	4	1	3	89.3	14.2	76.2	1.6	4	3	6
8*	80.0	58.0	4	1	5	100.0	11.1	66.7	15.0	4	2	8
Average	80.0	74.9	–	–	–	89.3	14.2	80.4	10.4	–	–	–

TABLE II
CLASSIFICATION ACCURACY COMPARISON BETWEEN NIRS AND ANS-NIRS SYSTEMS IN NOISY (CONSTANT BACKGROUND NOISE AND INTERMITTENT STARTLE NOISES) CONDITIONS. COLUMNS LABELED “SENS,” “SPEC,” AND “GAIN” INDICATE, RESPECTIVELY, CLASSIFIER SENSITIVITY, SPECIFICITY, AND RELATIVE PERFORMANCE GAINS OBTAINED OVER CONVENTIONAL NIRS BY USING ANS-NIRS SIGNALS. RESULTS FOR THE LAST TWO SUBJECTS USE ONLY A SUBSET OF THE ANS SIGNALS

Participant number	Constant background noise						Intermittent startle noise					
	NIRS (%)		ANS-NIRS (%)				NIRS (%)		ANS-NIRS (%)			
	Sens	Spec	Sens	Gain	Spec	Gain	Sens	Spec	Sens	Gain	Spec	Gain
1	50.0	75.0	100.0	100.0	70.0	-6.7	60.0	92.0	43.3	-27.8	91.7	-0.4
2	90.0	66.7	90.0	0.0	83.3	25.0	60.0	50.0	80.0	33.3	65.0	30.0
3	60.0	66.7	90.0	50.0	58.3	-12.5	67.0	57.0	100.0	49.3	58.3	2.3
4	40.0	50.0	62.5	56.3	66.7	33.3	75.0	50.0	63.3	-15.6	58.3	16.7
5	80.0	50.0	100.0	25.0	83.3	66.7	67.0	67.0	80.0	19.4	50.0	-25.4
6	80.0	41.7	100.0	25.0	100.0	140.0	60.0	83.0	80.0	33.3	66.7	-19.7
7*	66.7	66.7	100.0	49.9	75.0	12.4	60.0	57.0	80.0	33.3	83.3	46.1
8*	66.7	50.0	90.0	34.9	83.3	66.6	67.0	50.0	91.7	36.9	87.0	74.0
Average	66.7	58.3	91.6	42.6	77.5	40.6	64.5	63.3	77.3	20.3	70.0	15.5

incorporated into a BCI paradigm, would correctly suppress false activations resultant from involuntary startle reflexes, thus leading to improved performance in everyday settings.

V. DISCUSSION

A. Multimodal Mental State Classification

This study investigated the concurrent measurement of peripheral (i.e., electrodermal activity, skin temperature, respiratory effort, blood volume pulse) and central (i.e., cerebral oxygenation) nervous system responses to music imagery from able-bodied individuals for the purpose of accurate and robust mental activity classification. The prefrontal cortex has been implicated in the processing of emotional behavior [48]. Music imagery of self-selected songs, for example, has been shown to elicit the intense emotional responses needed to activate the prefrontal cortex [43], [49], [50]; such brain responses have been observed by functional magnetic resonance imaging [51], positron emission tomography [10], and NIRS [36], [52].

Similarly, decades of research have shown that signals harnessed from the peripheral autonomic nervous system can be volitionally controlled via mental tasks. The biofeedback literature, for example, suggests that fingertip temperature [26], [53] and blood pressure [27], [54] can be volitionally controlled

using mental imagery. Polygraphy studies have shown that electrodermal reactions can be controlled via mental relaxation exercises [55]. The mental stimulation literature, in turn, has shown that mental imagery tasks, such as motor or music imagery, can be used to control several ANS signals with the degree of observed responses being proportional to the mental effort exerted by the subjects [28], [30], [32], [56].

Given that both the central and peripheral nervous systems respond to music imagery, we pursued a multimodal paradigm, where prefrontal oxygenation and autonomic responses were monitored simultaneously. The underlying hypothesis was that multimodal classification would yield better mental state classification than looking at prefrontal oxygenation alone. While the signals harnessed from the two systems have different response times (e.g., 5–8 s for NIRS and 1–15 s for ANS signals), we have found that HMMs were effective at modeling the complex temporal interaction between the measured autonomic and NIRS signals during resting state. Dynamic time warping strategies [57] were investigated but led to no significant performance improvements. As argued in [28] and [56], mental imagery tasks can cause significant increases in skin conductance levels, heart and respiration rates, as well as a decrease in skin temperature. Recent work has also shown significant increases in NIRS raw signal amplitudes during mental imagery tasks [52]. The classification strategy proposed here uses a window

of 5 s post-log-likelihood slope changes (from positive to negative). It is believed that this choice of window length allows the majority of the signals to reach levels significantly different from those observed during rest, and thus can be accurately detected with HMMs without the need for more complex dynamic time warping strategies. The results shown in Table I for silent conditions suggest that indeed, once both modalities are incorporated, functional intent is better detected for most participants.

B. ANS Signal Contributions

Studies such as those reported in [28] and [56] suggest that all four ANS-harnessed signals undergo significant changes during mental imagery tasks. More recently, studies have gauged the contribution of each peripheral autonomic signal for mental task classification [58], [59]. It was observed, both previously and in this study, that skin conductance and skin temperature were the most labile, thus resonating with the findings reported in [30] and [32] and corroborating the somewhat lower results obtained for subjects 7 and 8 (see Table I). Alternately, respiration effort was shown to be the signal that least contributed to accurate mental state classification, a finding also reported in the polygraph literature (e.g., [60]).

C. Environmental Noise Effects

It is widely known that the performance of mental tasks can be severely affected when performed in a noisy environment [18]–[22] due to “mind wandering.” In the context of NIRS signals harnessed from the prefrontal cortex, further artifacts may arise from the processing of distracting auditory stimuli [23], [24]; such artifacts can be observed from the undulating baseline log-likelihood time series depicted by Fig. 3(d).

In terms of peripheral physiological signals, research has focused mostly on the effects of prolonged exposure to continuous ambient noise on human health, where indicators of stress and cardiovascular problems have been shown (e.g., [61]). A limited number of studies have explored the short term effects of exposure to continuous ambient noise. The study reported in [31], for example, suggested that continuous ambient noise had negligible effect on physiological signals. Startle stimuli, on the other hand, were shown to cause significant changes in cardiovascular blood pressure and volume, breathing, pulse rate, skin conductance, amongst other neural and corporeal activities [21], [30]–[32]. It is believed that the arousal responses due to startle stimuli are generated to increase alertness against potentially harmful events. For continuous ambient sounds, the human body habituates and arousal responses are suppressed as the repetitive noises signify nonthreatening conditions [31]. The results reported in Table II corroborate these findings. It was observed that for continuous ambient noise, inclusion of ANS-harnessed signals significantly improved classification performance over classification with NIRS signals alone. For intermittent startle noises, on the other hand, the gains were more modest. Interestingly, however, the gains observed for subjects 7 and 8 were significantly higher than those attained with the

remaining participants. This improvement is likely due to the lability of the EDA signal to startle stimuli which was not collected for these two participants.

D. Study Limitations and Future Investigations

Ultimately, the goal is to incorporate the proposed strategies into a BCI system aimed at improving the health and well-being of individuals with severe and multiple disabilities. In such cases, measuring the four physiological signals described herein may pose a challenge as, e.g., individuals with complete spinal cord injury may exhibit no electrodermal reactions below the level of injury [30] and individuals with late stage amyotrophic lateral sclerosis who are being artificially ventilated will not possess useful respiration signals. To this end, subject-specific physiological signal combinations may be needed where only a subset of the four investigated signals could be used, potentially leading to a decrease in performance, as observed with subjects 7 and 8 (see Table II). Additionally, for some individuals, alternative cognitive tasks may elicit stronger physiological reactions, thus leading to improved performance. Similarly, users may react differently, both in terms of reaction “magnitude” and reaction times, to various startle stimuli [see Fig. 3(d)]. As a consequence, further improvements may be obtained if startle noise classification (e.g., [62]) is incorporated into the compensation step.

VI. CONCLUSION

This paper has investigated the detrimental effects of environment noise on NIRS-based mental state classification. The first steps towards achieving robustness against environment noise were taken and two strategies were developed. First, physiological signals harnessed from the autonomic nervous system were incorporated as complementary information to cortical hemodynamics cues. The developed paradigm was shown to reliably discriminate between two cognitive states (rest and music imagery) which, in turn, could be used as BCI control signals in both silent and ambient noise conditions. Secondly, a startle noise detection and compensation algorithm was proposed. When in place, the classification accuracy obtained in noisy environments was comparable to that obtained in silent conditions. Such findings encourage further research into the design of BCIs that can be taken beyond controlled settings and into everyday environments.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their insightful comments and S. Bernstein for his assistance with Fig. 2.

REFERENCES

- [1] A. Sesin, M. Adjouadi, M. Ayala, M. Caberirzo, and A. Barreto, “Eyeing a real-time human-computer interface to assist those with motor disabilities,” *IEEE Potentials*, vol. 27, no. 3, pp. 19–25, 2008.
- [2] Y. Chen, “Application of tilt sensors in human-computer interface for people with disabilities,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 9, no. 3, pp. 289–294, Sep. 2001.

- [3] L. Struijk, "An inductive tongue computer interface for control of computers and assistive devices," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 12, pt. 2, pp. 2594–2597, Dec. 2006.
- [4] M. van Gerven *et al.*, "The brain-computer interface cycle," *J. Neural Eng.*, vol. 6, no. 4, 2009.
- [5] J. Wolpaw *et al.*, "Brain-computer interface technology: A review of the first international meeting," *IEEE Trans. Rehabil. Eng.*, vol. 8, no. 2, pp. 164–173, Jun. 2007.
- [6] R. Sitaram *et al.*, "Temporal classification of multichannel near-infrared spectroscopy signals of motor imagery for developing a brain-computer interface," *NeuroImage*, vol. 34, no. 4, pp. 1416–1427, 2007.
- [7] T. Germon, P. Evans, N. Barnett, P. Wall, A. Manara, and R. Nelson, "Cerebral near infrared spectroscopy: Emitter-detector separation must be increased," *J. Anaesthesia*, vol. 82, no. 6, pp. 831–837, 1999.
- [8] S. Coyle, T. Ward, C. Markham, and G. McDarby, "On the suitability of near-infrared (NIR) systems for next-generation brain-computer interfaces," *Physiol. Meas.*, vol. 25, pp. 815–822, 2004.
- [9] F. Matthews, B. Pearlmutter, T. Ward, C. Soraghan, and C. Markham, "Hemodynamics for brain-computer interfaces," *IEEE Signal Process. Mag.*, vol. 25, no. 1, pp. 87–94, Jan. 2008.
- [10] A. J. Blood and R. J. Zatorre, "Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion," *Proc. Nat. Acad. Sci. USA*, vol. 98, pp. 11 818–11 823, 2001.
- [11] M. Naito, Y. Michioka, K. Ozawa, Y. Ito, M. Kiguchi, and T. Kanazawa, "A communication means for totally locked-in ALS patients based on changes in cerebral blood volume measured with near-infrared light," *IEICE Trans. Info. Syst.*, vol. 90, no. 7, pp. 1028–1037, 2007.
- [12] A. Villringer and B. Chance, "Non-invasive optical spectroscopy and imaging of human brain function," *Trends Neurosci.*, vol. 20, no. 10, pp. 431–433, Oct. 1997.
- [13] S. Luu and T. Chau, "Decoding subjective preference from single-trial near-infrared spectroscopy signals," *J. Neural Eng.*, vol. 6, no. 4, Feb. 2009.
- [14] S. Coyle, T. Ward, and C. Markham, "Physiological noise in near-infrared spectroscopy: Implications for optical brain computer interfacing," in *Proc. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2004, vol. 2, pp. 4540–4543.
- [15] Q. Zhang, E. Brown, and G. Strangman, "Adaptive filtering for global interference cancellation and real-time recovery of evoked brain activity: A Monte Carlo simulation study," *J. Biomed. Opt.*, vol. 12, no. 4, p. 044014, 2007.
- [16] C. Nam, "Environmental noise and P300-based brain-computer interface," in *Proc. Human Factors Ergonomics Soc. Annu. Meeting*, 2008, vol. 52, pp. 803–807.
- [17] T. Falk, K. Paton, S. Power, and T. Chau, "Improving the performance of NIRS-based brain-computer interfaces in the presence of background auditory distractions," in *Proc. Int. Conf. Acoustics, Speech, Signal Process.*, 2010, pp. 517–520.
- [18] S. Bambury and D. Berry, "Disruption of office-related tasks by speech and office noise," *Br. J. Psychol.*, vol. 89, no. 3, pp. 499–517, 1998.
- [19] V. Gumenyuk, O. Korzyukov, K. Alho, C. Escera, and R. Naatanen, "Effects of auditory distraction on electrophysiological brain activity and performance in children aged 8–13 years," *Psychophysiology*, vol. 41, no. 1, pp. 30–36, 2004.
- [20] G. Cassidy and R. MacDonald, "The effect of background music and background noise on the task performance of introverts and extraverts," *Psychol. Music*, vol. 35, no. 3, pp. 517–537, 2007.
- [21] M. Flaten, E. Nordmark, and A. Elden, "Effect of background noise on the human startle reflex and prepulse inhibition," *Psychophysiology*, vol. 42, no. 3, pp. 298–305, 2005.
- [22] A. Furnham and L. Strbac, "Music is as distracting as noise: The differential distraction of background music and noise on the cognitive test performance of introverts and extraverts," *Ergonomics*, vol. 45, no. 3, pp. 203–217, 2002.
- [23] A. Engelien *et al.*, "The neural correlates of 'deaf-hearing' in man: Conscious sensory awareness enabled by attentional modulation," *Brain*, vol. 123, no. 3, p. 532, 2000.
- [24] M. Boly *et al.*, "Auditory processing in severely brain injured patients: Differences between the minimally conscious state and the persistent vegetative state," *Arch. Neurol.*, vol. 61, no. 2, pp. 233–238, 2004.
- [25] S. Blain, A. Mihailidis, and T. Chau, "Conscious control of electrodermal activity: The potential of mental exercises," in *Proc. IEEE Int. Conf. Eng. Med. Biol. Soc.*, 2006, pp. 6561–6564.
- [26] R. Freedman and P. Ianni, "Self-control of digital temperature: Physiological factors and transfer effects," *Psychophysiology*, vol. 20, pp. 682–689, 1983.
- [27] D. Kristt and B. Engel, "Learned control of blood pressure in patients with high blood pressure," *Circulation*, vol. 51, no. 2, pp. 370–378, 1975.
- [28] C. Deschaumes-Molinari, A. Dittmar, and E. Vernet-Maury, "Autonomic nervous system response patterns correlate with mental imagery," *Physiol. Behav.*, vol. 51, no. 5, pp. 1021–1027, 1992.
- [29] B. Wilhelm, M. Jordan, and N. Birbaumer, "Communication in locked-in syndrome: Effects of imagery on salivary pH," *Neurology*, vol. 67, no. 3, pp. 534–535, 2006.
- [30] S. Blain, A. Mihailidis, and T. Chau, "Peripheral autonomic signals as access pathways for individuals with severe disabilities: A literature appraisal," *Open Rehabil. J.*, vol. 1, pp. 27–37, 2008.
- [31] K. Kryter, "Non-auditory effects of environmental noise," *Am. J. Public Health*, vol. 62, no. 3, pp. 389–398, 1972.
- [32] S. Blain, S. Power, E. Sejdic, A. Mihailidis, and T. Chau, "A cardiorespiratory classifier of voluntary and involuntary electrodermal activity," *BioMedical Eng. OnLine*, vol. 9, no. 11, 2010, 13 pages.
- [33] M. Akbacak and J. Hansen, "Environmental sniffing: Noise knowledge estimation for robust speech systems," *IEEE Trans. Audio, Speech, Language Process.*, vol. 15, no. 2, pp. 465–477, Feb. 2007.
- [34] G. Morren *et al.*, "Detection of fast neuronal signals in the motor cortex from functional near infrared spectroscopy measurements using independent component analysis," *Med. Biol. Eng. Comput.*, vol. 42, no. 1, pp. 92–99, 2004.
- [35] I. Sase, H. Eda, A. Seiyama, H. Tanabea, A. Takatsuke, and T. Yanagida, "Multi-channel optical mapping: Investigation of depth information," in *Proc. Int. Soc. Opt. Eng.*, 2001, vol. 4250, pp. 29–36.
- [36] K. Tai and T. Chau, "Single-trial classification of NIRS signals during emotional induction tasks: Towards a corporeal machine interface," *J. NeuroEng. Rehabil.*, vol. 6, no. 39, 2009.
- [37] J. Kim and E. Andre, "Emotion recognition based on physiological changes in music listening," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 12, pp. 2067–2083, Dec. 2008.
- [38] B. Obermaier, C. Guger, C. Neuper, and G. Pfurtscheller, "Hidden Markov models for online classification of single trial EEG data," *Pattern Recognit. Lett.*, vol. 22, no. 12, pp. 1299–1309, 2001.
- [39] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. IEEE*, vol. 77, no. 2, pp. 257–286, Feb. 1989.
- [40] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. R. Stat. Soc.*, pp. 1–38, 1977.
- [41] A. Gersho and R. Gray, *Vector Quantization and Signal Compression*. Norwell, MA: Kluwer, 1993.
- [42] T. Falk and W. Chan, "Feature mining for GMM-based speech quality measurement," in *Proc. Conf. Signals, Syst. Comput.*, 2004.
- [43] M. Boso, P. Politi, F. Barale, and E. Enzo, "Neurophysiology and neurobiology of the musical experience," *Funct. Neurol.*, vol. 21, no. 4, pp. 187–191, 2006.
- [44] A. Kistler, C. Mariauzouls, and K. von Berlepsch, "Fingertip temperature as an indicator for sympathetic responses," *Int. J. Psychophysiol.*, vol. 29, pp. 35–41, 1998.
- [45] E. Cook, L. Hawk, T. Davis, and V. Stevenson, "Affective individual differences and startle reflex modulation," *J. Abnormal Psychol.*, vol. 100, no. 1, pp. 5–13, 1991.
- [46] T. Falk and W. Chan, "Hybrid signal-and-link-parametric speech quality measurement for VoIP communications," *IEEE Trans. Audio, Speech Language Process.*, vol. 16, no. 8, pp. 1579–1589, Aug. 2008.
- [47] G. Muller-Putz, R. Scherer, C. Brunner, R. Leeb, and G. Pfurtscheller, "Better than random? A closer look on BCI results," *Int. J. Bioelectromagn.*, vol. 10, no. 1, pp. 52–55, 2008.
- [48] J. Fuster, "The prefrontal cortex—An update: Time is of the essence," *Neuron*, vol. 30, no. 2, pp. 319–333, 2001.
- [49] E. Altenmüller, K. Schürmann, V. Lim, and D. Parlitz, "Hits to the left, flops to the right: Different emotions during listening to music are reflected in cortical lateralization patterns," *Neuropsychologia*, vol. 40, no. 13, pp. 2242–2256, 2002.
- [50] C. Krumhansl, "An exploratory study of musical emotions and psychophysiology," *Canadian J. Exp. Psychol.*, vol. 51, no. 4, pp. 336–353, 1997.
- [51] B. Kleber, N. Birnbauer, R. Veit, T. Trevorrow, and M. Lotze, "Overt and imagined singing of an Italian aria," *NeuroImage*, vol. 36, pp. 889–900, 2007.
- [52] S. Power, T. 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, no. 2, 2010.

- [53] J. Grabert, N. Bregman, and H. McAllister, "Skin temperature regulation: The effects of suggestion and feedback," *Int. J. Neurosci.*, vol. 10, pp. 217–221, 1980.
- [54] E. Bleecker and B. Engel, "Learned control of ventricular rate in patients with atrial fibrillation," *Psychosomatic Med.*, vol. 35, no. 2, pp. 161–175, 1973.
- [55] E. Elaad and G. Ben-Shakhar, "Effects of mental countermeasures on psychophysiological detection in guilty knowledge test," *Int. J. Psychophysiol.*, vol. 11, no. 2, pp. 99–108, 1991.
- [56] O. Bolliet, C. Collet, and A. Dittmar, "Autonomic nervous system activity during actual and mentally simulated preparation for movement," *Appl. Psychophysio. Biofeedback*, vol. 30, no. 1, pp. 11–20, 2005.
- [57] D. Berndt and J. Clifford, "Using dynamic time warping to find patterns in time series," in *Proc. AAAI Workshop Knowledge Discovery Databases*, 1994, pp. 229–248.
- [58] M. Guirgis, T. Falk, S. Power, S. Blain, and T. Chau, "Harnessing physiological responses to improve NIRS-based brain-computer interface performance," in *Proc. ISSNIP Biosignals Biorobotics Conf.*, 2010, pp. 59–62.
- [59] T. Falk, M. Guirgis, S. Power, S. Blain, and T. Chau, "On the use of peripheral autonomic signals for binary control of body-machine interfaces," *Physiological Measurement*, vol. 31, no. 10, pp. 1–11, 2010.
- [60] G. Ben-Shakhar and K. Dolev, "Psychophysiological detection through the guilty knowledge technique: Effects of mental countermeasures," *J. Appl. Psychol.*, vol. 81, no. 3, pp. 273–281, 1996.
- [61] W. Babisch, "Stress hormones in the research on cardiovascular effects of noise," *Noise and Health*, vol. 5, no. 15, pp. 1–11, 2003.
- [62] A. Abu-El-Quran, R. Goubran, and A. Chan, "Security monitoring using microphone arrays and audio classification," *IEEE Trans. Instrum. Meas.*, vol. 55, no. 4, pp. 1025–1032, Aug. 2006.



Tiago H. Falk (S'00–M'09) received the B.Sc. degree from the Federal University of Pernambuco, Brazil, in 2002, and the M.Sc. (Eng.) and Ph.D. degrees from Queen's University, Kingston, ON, Canada, in 2005 and 2008, respectively, all in electrical engineering.

From 2009 to 2010, he was with Holland Bloorview Kids Rehabilitation Hospital, affiliated with the University of Toronto, Toronto, ON, Canada. He is now an Assistant Professor at INRS-EMT, University of Quebec, Montreal, Canada. His research

interests include multimodal signal analysis, biomedical signal processing, rehabilitation engineering, and multimedia quality assessment and enhancement.

Dr. Falk is recipient of the IEEE Kingston Section Ph.D. Research Excellence Award (2008), the Best Student Paper Awards at ICASSP (2005) and IWAENC (2008), and the Newton Maia Young Scientist Award (2001).



Mirna Guirgis received the B.A.Sc. degree in engineering science in 2010 and is currently pursuing a M.A.Sc. in biomedical engineering at the University of Toronto, Toronto, ON, Canada.

Her research interests are in biomedical signal processing and in electrophysiology focusing on epilepsy.



Sarah Power received the B.Eng. degree in electrical engineering from the Memorial University of Newfoundland, St. John's, NL, Canada, in 2006, and the M.A.Sc. degree in electrical/biomedical engineering, in 2008, from the University of Toronto, Toronto, ON, Canada, where she is currently a doctoral candidate in biomedical engineering.

Her research interests involve the development of access pathways to enable individuals with severe and multiple disabilities to communicate by unconventional means.



Tom T. Chau (S'92–M'97–SM'03) received the B.A.Sc. degree in engineering science and the M.A.Sc. degree in electrical and computer engineering from the University of Toronto, Toronto, ON, Canada, in 1992 and 1994, respectively, and the Ph.D. degree in systems design engineering from the University of Waterloo, Waterloo, ON, Canada, in 1997.

He was previously with IBM Canada. Currently, he is a Senior Scientist at Holland Bloorview Kids Rehabilitation Hospital, Toronto, ON, Canada, and an Associate Professor in the Institute of Biomaterials and Biomedical Engineering, University of Toronto, where he also directs the Clinical Engineering Program. He has held a Canada Research Chair in Pediatric Rehabilitation Engineering since 2004. His research program focuses on body-machine interfaces for alternative access to communication. Chau leads the Infinity Centre for Access Innovations at Holland Bloorview, a state-of-the-art multi-investigator turnkey laboratory for technology-mediated access alternatives for children who are non-verbal and severely physically impaired.