Limitations of NIRS-Based BCI for Realistic Applications in Human-Computer Interaction

M. Strait[1], C. Canning[1], M. Scheutz[1]
[1] *Tufts University, Medford, MA, USA*

Correspondence: Megan Strait, Tufts University, 161 College Avenue, Medford, MA, USA. E-mail: mstrait@cs.tufts.edu

Abstract. Recent work in human-computer interaction (HCI) has endorsed NIRS-based BCI as suitable for use as a realtime input modality in *realistic* interaction settings. However, several discrepancies between the state-of-the-art NIRS signal processing and the purported applications in HCI suggest otherwise. To investigate such discrepancies and their effects on signal reliability, we conducted a large-scale evaluation of three factors – (1) unrestricted subject movement, (2) variable task duration, and (3) realtime signal processing – factors central to realistic settings. Our findings show high signal unreliability with NIRS-based BCI subject to these three factors, suggesting NIRS-based systems are premature for realistic applications in HCI.

Keywords: Human-Computer Interaction, NIRS

1. Introduction

Within the HCI community, NIRS-based BCI has gained much attention due to recent demonstrations of its use as a passive interaction modality to improve user performance [Solovey et al., 2012]. This work argues NIRS-based systems are better suited for realistic settings (in comparison to other neuroimaging techniques), reporting NIRS as being cheaper, more portable, and more robust to noise. However, state-of-the-art neuroimaging delineates these features still problematic for the use of NIRS-based systems (as well as other techniques) in realistic settings. Such issues (in addition to numerous physical limitations such as portability, ambient lighting, and probe placement) include a multitude of signal processing challenges, importantly: (1) motion artifact detection and removal, (2) separation of task-related and unrelated activation beyond standard 20–60 s task durations, and (3) realtime/single-trial classification. In this paper, we present an investigation of these three key challenges and their implications for realistic NIRS-based applications in HCI. Although our results show high unreliability of the NIRS-based systems and signal processing used currently in HCI, we hope the construction of a such a large, publicly available data set will facilitate steps towards more realistic and more reliable applications.

2. Material and Methods

We conducted a 3-part, 40-subject investigation to address how (1) unrestricted motion, (2) variable task duration, and (3) realtime/single-trial constraints affect signal reliability. These factors were selected for their importance in realistic HCI applications. Current NIRS-based BCI require participants to minimize movement (e.g. using a chin rest); however, standard HCI investigations do not obstruct participant movement. Furthermore, standard HCI tasks can range from 30 s to 30 min in duration, whereas standard BCI tasks range from 20 s to 60 s and it remains to be demonstrated that NIRS-based BCI can be employed in longer tasks. Finally, as passive NIRS-based BCIs are intended to operate in realtime to improve user performance, single-trial processing must be reliable. Based on the primary region of interest to HCI, we sampled only the anterior prefrontal cortex (aPFC).

2.1. Equipment and Analysis

A 2-probe, ISS OxiplexTS[1] tissue oximeter was used to record changes in hemoglobin at 3cm depth in the subject's aPFC at a temporal resolution of 6.25 Hz. We applied three serial centered moving averages to reduce the effects of systemic artifacts (cardiac pulsation, respiration, and Mayer waves) followed by a folding average over trial repetitions, based on [Fekete et al., 2011; Sassaroli et al., 2009; Solovey et al., 2012]. To restricted motion (RM), the subject was seated in the Caravan Zero Gravity chair[2] and was instructed not to move. In the unrestricted setup (UM), the subject was seated in a generic office chair and told he/she may sit comfortably.

[1] http://www.iss.com/biomedical/instruments/oxiplexTS.html
[2] http://www.amazon.com/gp/product/B0032UY0BK/ref

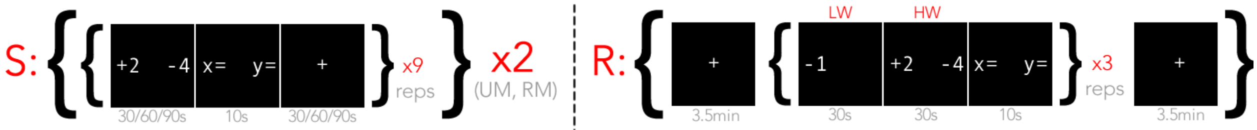

Figure 1: Experimental design. Standard block design (S, left) and realtime/single-trial trial design (R, right).

2.2. Participants and Procedure

40 subjects (18 male), ages 18–45, were recruited via a university website. This study was approved by the Tufts IRB and all participants provided informed, written consent. Based on [Sassaroli et al., 2009], we induced aPFC activation using a simple arithmetic task (see Fig. 1). A standard trial consisted of a task period (30, 60 or 90 s) in which new arithmetic operations (addition or subtraction) were shown every 5 s, followed by an answer period (10 s), and then a rest period (of duration equal to the task period). In each trial, the subject was tasked to remember the starting values and the running sums of two variables.

Following an initial 5-minute baseline measurement, two experimental conditions (standard block design (S), and realistic continuous-task design (R), were administered in a randomized order (see Fig. 1). The standard condition was performed twice, once with restricted motion (RM) and once with unrestricted motion (UM), to test the effects of the latter on signal reliability. To investigate the effects of variable task duration, the participant completed three standard trials of each 30, 60, and 90 s task durations for a total of 9 trials (see Fig. 1, left). To evaluate continuous/realtime processing constraints (R), we adapted the trial described above to create a two-workload variant (see Fig. 1, right). Here, the task period consisted of a low-workload task (LW, *maintaining only one variable*), followed by a high-workload task (HW, *two variables*), followed by an answer period but *no interim rest period* (to emulate the two-workload, realtime task used in [Solovey et al., 2012]). This realtime trial was repeated three times in serial, between two rest periods (pre- and post-task).

3. Results and Discussion

Average task performance was 35.7 % (SD 24.5 %) correct answers, with an average response rate of 98.6 % (SD 4.7 %), indicating the arithmetic task used was difficult but that subjects were attentive. To confirm a task-evoked hemodynamic response, we averaged the first 30 s of all standard RM trials and compared it to baseline activity, finding significant differences (for 38/40 subjects. Thus we concluded the task elicited a significant hemodynamic response in the aPFC, as designed.

Regarding the two experimental conditions (S and R), we found the following: (1) Unrestricted motion significantly affects signal reliability: UM x RM was significantly different for 38/40 subjects for each time variation (30, 60, and 90 s). (2) Signal reliability degrades over time: in the standard 30 s restricted setup (RM, 30 s), the task signal was significantly different from rest. However, this was not the case for the longer trial durations (60 and 90 s; $p > 0.36$). (3) Realtime constraints severely reduces signal reliability. Single-trial analysis of the realtime (R) condition showed significant differences between repetitions of the same workload trials (e.g. LW_1 x LW_2 x LW_3) for 36/40 subjects. Furthermore, at no point during the 3.5 minute post-task rest period did the signal become non-significantly different from pre-task rest for 37/40 subjects. This implies detecting the switch from an active task to no task is not feasible (at least not within a 3.5 minute period), which more importantly, suggests low signal resolution.

The implications of our findings with regards to current NIRS-based BCI applications in HCI targeting the aPFC are the following: (1) motion should be restricted until further advancements in signal processing; (2) task durations should be restricted to the standard 20–60 s range or otherwise explicitly demonstrated for a given task that it is reliably classified for a longer duration; and (3) passive NIRS-based BCI is premature for any aPFC-based tasks. While such issues have been raised and discussed for fMRI and EEG techniques, until now they have not been sufficiently discussed in regards to NIRS. However, we hope that by making such a dataset publicly available, we can further investigate and address the challenges limiting its applications in HCI.

References

Fekete, T., Rubin, D., Carlson, J., and Mujica-Parodi, L. (2011). The NIRS analysis package: Noise reduction and statistical inference. *PloS one*.

Sassaroli, A., Zheng, F., Coutts, M., Hirshfield, L. M., Girouard, A., Solovey, E. T., Jacob, R. J. K., Tong, Y., Frederick, B., and Fantini, S. (2009). Application of near-infrared spectroscopy for discrimination of mental workloads. *SPIE Proceedings*, 7174.

Solovey, E., Schermerhorn, P., Scheutz, M., Sassaroli, A., Fantini, S., and Jacob, R. (2012). Brainput: enhancing interactive systems with streaming fNIRS brain input. In *Proceedings of CHI*, pages 2193–2202.