Using functional Near-Infrared Spectroscopy in HCI: Toward evaluation methods and adaptive interfaces

Audrey Girouard, Leanne M. Hirshfield, Erin Solovey, Robert J.K. Jacob
Computer Science Department
Tufts University
Medford, MA 02155
{audrey.girouard, leanne.hirshfield, erin.solovey, robert.jacob}@tufts.edu

ABSTRACT

Functional Near-Infrared Spectroscopy (fNIRS) is a new brain imaging tool that shows potential for use in the field of human computer interaction (HCI) because of its lightweight, non-invasive qualities. fNIRS could become an additional input to interfaces, by recording the user's mental state through the measure of blood flow in the brain. However, before we are able to use the tool at its full potential, we must test its feasibility in HCI, and develop methods to accurately analyze the output. This paper will introduce fNIRS, and briefly discuss a feasibility study conducted to explore the measurement of different levels of workload. Finally, we will present future research directions that follow from this work, such as evaluating new interaction styles according to the measured mental workload, adaptive interfaces with fNIRS, and combining fNIRS with EEG.

Author Keywords

brain computer interfaces (BCI), workload, fNIRS, interaction styles, EEG

ACM Classification Keywords

H.5.2 [Information Interfaces and Presentation]: User Interfaces; H.1.2 [Models and Principles]: User/Machine Systems — human factors;

INTRODUCTION

Acquiring measurements about the mental state of a computer user would be valuable in human-computer interaction (HCI), both for evaluation of interfaces and for real time input to computer systems. Although we can accurately measure task completion time and accuracy, factors such as mental workload, frustration and distraction are typically limited to qualitative observations or subjective surveys. These surveys are often taken after the completion of a task, potentially missing valuable insight into the user's changing experience throughout the task. New evaluation techniques that monitor user experiences while working with computers are increasingly necessary. To address these evaluation issues, much current research focuses on developing objective techniques to measure user states such as workload, emotion, and fatigue in real time. Although this ongoing research has advanced user experience measurements in the HCI field, finding accurate, non-invasive tools to measure computer users' states in real working conditions remains a challenge. In addition to aiding in the evaluation of interfaces, fNIRS output has potential as an additional parallel, lightweight input channel for users. This additional information from the brain could be used to improve the efficiency or intuitiveness of the user's interaction with the machine by adapting the interface accordingly. It also could provide new interaction methods for disabled users.

We investigate functional near-infrared spectroscopy (fNIRS) [1], a relatively new technology for brain activity measurement, which we combine with the use of machine learning to analyze the resulting data. The emerging fNIRS tool is safe, portable, non-invasive, and can be implemented wirelessly, allowing for use in real world environments, making naturalistic HCI possible.

This position paper briefly describes the new fNIRS technology, presents our first experiment with the fNIRS tool, and demonstrates its feasibility and potential for HCI settings. We then describe three avenues of research investigated at Tufts University: using fNIRS as a more objective measure for evaluating emerging interaction styles, as a non-invasive device for use with adaptable interfaces, and as a tool that can be combined with EEG, providing complementary measurements.

FUNCTIONAL NEAR-INFRARED SPECTROSCOPY

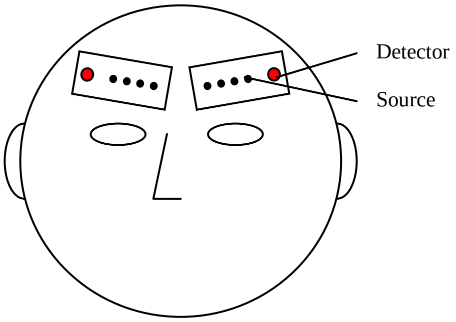

Figure 1: A schematic diagram of two detectors and their set of near-infrared light sources.

A new non-invasive brain imaging technique, functional near infrared spectroscopy (fNIRS), has been introduced [5, 7] to complement, and in some cases overcome technical and practical limitations, of EEG and other brain monitoring techniques. While EEG measures electrical activity, fNIRS measures blood oxygenation levels in the brain, providing a different, possibly complementary, source of information about brain functioning.

To measure blood oxygenation levels, functional near-infrared spectroscopy uses light sources placed on the scalp to send near-infrared light into the head. Biological tissues are relatively transparent to these wavelengths, so the light attenuation through tissues is sufficiently low to allow for tissue imaging at depths up to 2-3 centimeters. Deoxygenated and oxygenated hemoglobin are the main absorbers of near-infrared light in tissues, and they provide relevant markers of hemodynamic and metabolic changes associated with neural activity in the brain. Therefore, fNIRS researchers can estimate hemodynamic changes by using light detectors to monitor reflected light that has probed the brain cortex [3]. Figure 1 displays an example of the arrangement of detectors and light sources in an fNIRS device.

WORKLOAD AND INTERACTION STYLE EXPERIMENT

We have conducted a feasibility study using fNIRS. The goal of the study was to determine whether our approach could measure frontal lobe activity such as workload. We applied preprocessing and machine learning techniques to the brain data to classify this mental activity.

Four subjects completed thirty tasks in which they viewed the top and all sides of a rotating three dimensional shape comprised of eight small cubes. During each task, subjects counted the number of squares of each color displayed on the rotating shape in front of them. Figure 2 illustrates an example of a rotating shape. In the experiment, cubes could be colored with two, three, or four different colors, which we hypothesized would lead to different workload levels. A blank screen represented the baseline state (no colors).

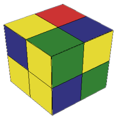

Figure 2. A cube made up of eight smaller cubes.

Experiment A: Graphical Blocks

The main goal of this experiment was to decide whether fNIRS data is sufficient for determining the workload level of users as they perform tasks. To accomplish this, a graphical interface displayed the rotating shapes.

Experiment B: Graphical versus Physical Blocks

We also wanted to determine whether there is a difference in mental workload when a user completes a spatial reasoning task on a graphical display versus completing the task using a physical object, such as a tangible user interface. Therefore, we included a physical cube with three colors, and the same rotation time and size as the graphical cube.

At the completion of each task, the subject was asked for his or her count of each color. He/she was then instructed to rest for thirty seconds, allowing his or her brain to return to a baseline state. After completing the tasks, the subject was presented with an additional example of each workload level and asked to fill out a NASA-Task Load Index [2], administered to compare our results with an established measure of workload. The NASA-Task Load Index results validated our hypothesis: an increased number of colors lead to a higher subjective workload level.

Data Analysis and Results

We preprocessed the data by normalizing it, applying a detrending algorithm and using a sliding window paradigm to generate average and slope features. Using a blocked cross-validation, we classified the data with a multilayer perceptron classifier. We tested distinguishing all five workload levels from each other, as well as distinguishing between two, three, and four different workload conditions (graphical). We also tried to classify between the graphical and physical tasks which had three colors.

Analysis of Graphical Blocks

We consider here one of the analyzed combinations of workloads. Comparing workload levels 0, 2, and 4 (no colors, 2 and 4 colors), classification accuracies ranged from 41.15% to 69.7%, depending on the subject. Considering that a random classifier would have 33.3% accuracy, the results are promising. It seems that we can predict, with relative confidence, whether the subject was experiencing no workload (level zero), low workload (level two), or high workload (level 4).

Analysis of Graphical versus Physical Blocks

We compared the results from the physical and the graphical workload level 3 (three colors). The average accuracy was 83%, with a range from 73% to 91%, depending on the subject (chance level 50%). These positive classification results are useful from a HCI perspective—there were distinguishable differences between displaying a cube in a graphical vs. physical user interface. Although we can accurately distinguish between the cognitive activities experienced in these two conditions, we cannot say for sure whether the difference is attributable to the workload of the interface, the workload of the task, or other variables affecting brain activity. However, these results encourage further exploration into cognitive workload associated with different interaction styles.

FUTURE DIRECTIONS AT TUFTS UNIVERSITY

We believe the results from these experiments are promising and demonstrate the feasibility of our approach. Current work is iterating on the machine learning and pre-processing algorithms in order to attain better classification results. The analysis and results described above demonstrated the need to explore data analysis techniques that are better suited to the fNIRs data. We will discuss these new techniques at the workshop. Unlike many other brain-computer interaction studies, our initial goal is improved interaction for all users, rather than only disabled users, for whom brain input is a viable alternative to otherwise unavailable inputs. We see potential in the application of these objective workload measurements to research on evaluating user interfaces, especially post-WIMP interfaces and on the development of adaptable user interfaces. We also hypothesize that adding EEG to our experiments will complement the output.

Evaluating emerging interfaces using fNIRS

Current methods of evaluation (performance, speed of execution, etc.) may be insufficient for emerging interface or interaction styles. In a CHI 2006 workshop entitled “What is the Next Generation of Human-Computer Interaction?” [4] participants discussed the need for new evaluation metrics for new interactions style. When it comes to emerging interaction styles such as virtual reality, tangible user interfaces and context aware systems, evaluating “intuitiveness,” enjoyment, mental workload or fatigue may be valuable.

Our feasibility study indicated promise in our ability to measure mental workload using fNIRS, and to classify the different workload levels using machine learning techniques. Given those results, it is possible to imagine using fNIRS to measure mental workload for different types of emerging interfaces while users accomplish a similar task, and to evaluate them according to their mental workload levels. Because a well designed interface should ‘melt away’, allowing the user to focus on the task at hand , an interface that causes lower mental workload would be preferable to its more difficult to use counterparts.

While brain measurement in HCI has typically been used to investigate overall task difficulty, we propose to dig more deeply by separating the total mental workload into two components. Specifically, we hypothesize that the overall mental effort required to perform a task using an interactive computer system is composed of a portion attributable to the difficulty of the task itself plus a portion attributable to the difficulty of operating the user interface of the interactive tool. In this regard, we follow the concept of Shneiderman's theory of syntactic/semantic components of a user interface [6]. The syntactic component includes interpreting the feedback it presents and formulating and inputting commands to the interface. An important goal in interface design is to reduce the amount of mental effort devoted to the interface-related or syntactic aspects so that more mental capacity can be devoted to the underlying task or semantic aspects. Acquiring objective measures of mental workload in users’ real working conditions can advance research on evaluating emerging interaction styles.

Adaptive interface using fNIRS

Objective, non-invasive measurements of user workload would be valuable as real time inputs to interactive systems, which can then adapt their behavior to current information measured from the brain. For example, given that we could measure and classify mental workload accurately, we could use this data and create an interface that would adapt to the user’s mental workload. An overloaded user would have the interface’s difficulty level reduced, while an underloaded user would see an increase in difficulty. Both scenarios need an interface adjustment because overloaded users tend to make errors due to stress and failure to accomplish all tasks, while underloaded users have a drop in productivity as well due to boredom and under challenging [8, 9].

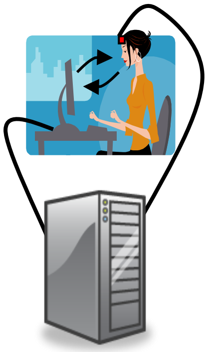

Figure 3. Adaptive user interface system design using fNIRS

The goal, thus, is to design user interfaces that treat the brain activity as an additional input channel, rather than as the primary input. For example, the user would operate a conventional interface with a mouse or other interaction method, and the interface would respond not only to the explicit mouse inputs but also to the information we can measure from the brain. In this case the challenge is to design a user interface that makes judicious use of brain input. Instead of using the brain input to, for example, directly drive the cursor, we want to use it in a much more subtle way.

The design challenges for such an unobtrusive, passive, real-time interface are considerable. There might be many different signals that can or could be measured by the fNIRS device. We intend to explore such possible signals: workload, emotions, fatigue, boredom, etc. Selecting the interface to be adapted and the adaptation method are another factor to take into account in our research. Finally, note that the difficulty level of the interface could be linked to multiple factors, such as the number or frequency of the tasks of the given interface, or the ergonomics of such an interface.

Combining fNIRS and EEG

Currently, we are studying the two directions mentioned above with the use of functional near-infrared spectroscopy. However, we plan on adding EEG to experiments, and combining the results from both tools to obtain an enhanced picture of the brain activity when the user is using a particular interface. The two types of sensors can be placed around the scalp in alternating patterns, with minimal interference. EEG measures surface electrical activity, and has poor spatial resolution because the measured brain activity can be attributed only to large region of the brain. fNIRS is complementary to EEG because while it has a lower spatial resolution compared to fMRI, its measurements are still more localized than in EEG. However, both tools have a high temporal resolution which allows for a good machine learning online analysis.

CONCLUSION

We presented fNIRS, a new noninvasive brain measurement tool that promises to be a useful input for research in HCI. We discussed a study testing the feasibility of the fNIRS device to detect levels of workload in HCI. Our experiment showed several workload comparisons with promising levels of classification accuracy. We developed classification techniques to interpret fNIRS data and demonstrated the use of fNIRS in HCI. Our next goals are to use this technology as a real time input to a user interface in a realistic setting as well as an evaluation technique for emerging interactions styles. We observe that our equipment places no unreasonable restrictions on a subject using an interactive system, and it can collect and transmit data in real time. This proves promising for the HCI community.

ACKNOWLEDGMENTS

We thank Kelly Moran, Hadar Rosenhand, and the rest of our colleagues in the HCI research group at Tufts; Sergio Fantini, Angelo Sassaroli and Yunjie Tong from the Biomedical Engineering Department at Tufts for their collaboration and help with fNIRS; Carla Brodley, Rachel Lomasky, Umaa Rebbapragada, and D. Sculley at Tufts for their help with machine learning aspects of our work; and Desney Tan at Microsoft Research for his helpful inputs and encouragement.

We thank the National Science Foundation for support of this research (NSF Grant Nos. IIS-0713506 and IIS-0414389). Any opinions, findings, and conclusions or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of the National Science Foundation. We also thank the Natural Sciences and Engineering Research Council of Canada for financial support.

REFERENCES

1. Chance, B., Anday, E., Nioka, S., Zhou, S., Hong, L., Worden, K., Li, C., Murray, T., Ovetsky, Y. and Thomas, R. A novel method for fast imaging of brain function, non-invasively, with light. *Optics Express*, *10* (2), 1988, 411-423.
2. Hart, S.G. and Staveland, L.E. Development of NASA-TLX (Task Load Index): Results of empirical and theorical research. in Hancock, P., Meshkati, N. ed. *Human Mental Workload*, Amsterdam, 1988, pp 139 - 183.
3. Izzetoglu, K., Bunce, S., Onaral, B., Pourrezaei, K. and Chance, B. Functional Optical Brain Imaging Using Near-Infrared During Cognitive Tasks. *International Journal of Human-Computer Interaction*, *17* (2), 2004, 211-231.
4. Jacob, R.J.K., Girouard, A., Hirshfield, L.M., Horn, M., Shaer, O., Solovey, E.T. and Zigelbaum, J. What Is the Next Generation of Human-Computer Interaction? *interactions*, 2007.
5. Meek, J., Elwell, C., Khan, M., Romaya, J., Wyatt, J., Delpy, D. and Zeki, S. Regional Changes in Cerebral Haemodynamics as a Result of a Visual Stimulus Measured by Near Infrared Spectroscopy. *Proc. Roy. Soc. London*, *261*, 1995, 351-356.
6. Shneiderman, B. and Plaisant, C. Designing the User Interface: Strategies for Effective Human-Computer Interaction, Fourth Edition, Addison-Wesley, Reading, Mass., 2005.
7. Villringer, A., Planck, J., Hock, C., Schleinkofer, L. and Dirnagl, U. Near Infrared Spectroscopy (NIRS): A New Tool to Study Hemodynamic Changes During Activation of Brain Function in Human Adults. *Neurosci. Lett.*, *154*, 1993, 101-104.
8. Wickens, C.D. and Hollands, J.G. *Engineering Psychology and Human Performance*. Prentice-Hall, Englewood Cliffs, NJ, 2000.
9. Wickens, C.D., Lee, J., Liu, Y.D. and Gordon-Becker, S. *Introduction to Human Factors Engineering (2nd Edition)*. Prentice-Hall, Inc, 2003.