Understanding How to Design Complex Brain-Controlled Applications

Krzysztof Z. Gajos
Intelligent Interactive Systems Group
Harvard School of Engineering and Applied Sciences
33 Oxford St. Rm 251
Cambridge, MA 02138, USA
kgajos@eecs.harvard.edu

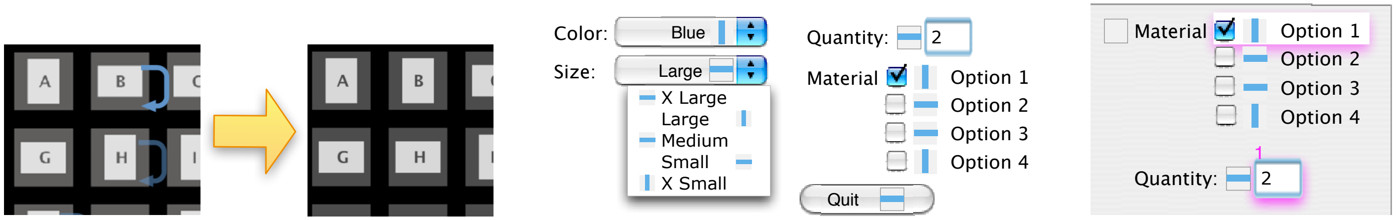
Bio: Krzysztof Gajos is an assistant professor of Computer Science at the Harvard School of Engineering and Applied Sciences. Krzysztof is primarily interested in intelligent interactive systems, an area that spans human-computer interaction, artificial intelligence, and machine learning. Krzysztof received his B.Sc. and M.Eng. degrees in Computer Science from MIT. Subsequently he was a research scientist at the MIT Artificial Intelligence Laboratory, where he managed The Intelligent Room Project. In 2008, he received his Ph.D. in Computer Science from the University of Washington in Seattle. Before coming to Harvard, he spent a year as a post-doctoral researcher in the Adaptive Systems and Interaction group at Microsoft Research.

CHI 2010, April 10–15, 2010. Atlanta, Georgia, USA
ACM 978-1-60558-930-5/10/04.

Introduction

Most of the effort in brain-computer interface (BCI) research so far has been directed at developing better sensors and better ways of extracting useful information from the brain signal, while little effort has been directed at systematically understanding the unique strengths and limitations of this input modality and their implications for interaction design. Past proof of concept brain-controlled applications either involved very simple interaction or they augmented existing complex applications with external widgets to enable limited control.

This relative lack of research directed at developing applications and interaction methods specifically for brain-computer interfaces is a concern for several reasons. First, some BCI technologies are mature enough to be used soon by large numbers of paralyzed users. Lack of compelling brain-controlled applications or the tools and techniques for building such applications will significantly limit the impact of these technologies. Second, recent work on ability-based user interfaces (e.g., [4, 5]) has demonstrated that large gains in efficiency of interaction and user satisfaction can be achieved if user interfaces are designed with a user's specific abilities and devices in mind. The results of our studies showed that adapting user interfaces to the unique abilities of people with a range of motor impairments helped close the performance gap between those users and able-bodied people by over 60% [4]. For BCI users, who need up to several tens of seconds to perform a single UI operation, efficiency of in-



(a) Our implementation of the P300 speller uses shape rotation rather than flashing as stimulus to avoid attention capture and to minimize interference among targets.

(b) Prototypes of basic GUI widgets adapted for use with the P300-based BCI.

(c) Afterglow effects may aid user's short term memory compensating for the cognitive load imposed by the operation of the BCI.

Figure 1: Examples of P300-based brain-controlled interfaces.

teraction will have an even larger impact and will likely determine whether an application is usable in practice. Lastly, a lot of the current research effort in BCI is directed at improving various aspects of the sensing and signal extraction technology. Good understanding of the user interaction requirements of brain-controlled applications will help inform and direct those efforts to maximize their impact on the user experience.

Motivated by these observations, we are starting a project to explore the properties and limitations of one particularly promising BCI paradigm as an input modality and to develop methods and tools for designing user interfaces for complex brain-controlled applications.

P300-Based BCI

We chose to work with EEG because it is well-studied, relatively inexpensive, and has been demonstrated to support reasonable (for BCI) information transfer rates. For the mental strategy, we chose the P300 event-related potential (ERP) [11] because of ease of detection and because it is involuntary so it requires no training.

Briefly, the P300 ERP can be elicited as follows: if a person is paying attention to a particular visual object and a sudden surprising change (stimulus) is applied to its appearance, for example it abruptly becomes much brighter, the person's brain will respond with a spike of activity roughly 300 milliseconds after the onset of the surprising stimulus. An effective application of the P300 ERP for intentional control of a computer application was demonstrated with the P300 speller [3], which we intend to generalize to arbitrary GUIs.

The P300-based brain-computer interface paradigm has several unique properties important for interaction design:

- this approach supports **discrete selection** from a fairly **large number of options** visually presented to the user (36 in our preliminary study);
- each selection takes a **considerable amount of time**—typically between 10 and 20 seconds are needed to robustly classify the person's intended selection;
- this input modality involves some amount of **uncertainty**—in a small number of situations, typically much below 10%, the user's intention may be misrecognized.

These properties make P300-based BCI substantially different from other well-studied input modalities for graphical user interfaces. Prior HCI research has considered noisy continuous (1D or 2D) input (e.g., [10, 5, 13]) or discrete input with very few choices, i.e., switches (e.g., [7, 2]). The most unique property of the P300-based BCI is that it allows selection from a large number of simultaneous options and that these options can be directly mapped to interface elements. What would a user interface look like if it were designed specifically for control with P300-based BCI?

Basic Operations

All actions that can currently be performed with a mouse click—with the possible exception of those relying on precise positioning of the mouse pointer—can be accomplished in the P300 paradigm by attaching a visual stimulus to that action. For example, Figure 1(b) shows several examples of simple GUI elements augmented with visual markers (currently ▤, though their design may change). To activate a button, to select a check box, or to select an item in a list, the person will have to focus their attention on the corresponding visual marker while a stimulus (rapid 90 degree rotation) is being applied to it. Selecting a marker associated with a text field, will bring up a virtual keyboard, possibly similar to the existing P300 speller.

We use the rotation of a geometric marker as a stimulus (following [6]) because, unlike the flashing stimulus used in most implementations of the P300 speller, the shape rotation does not cause attention capture [12] or interference between stimuli applied to neighboring symbols [8], opening the way for this approach to be used for controlling complex graphical user interfaces. To verify this, we have implemented a P300 speller using both flashing and shape rotation stimuli. In our experiments, this shape rotation stimulus (Figure 1(a)) resulted in more consistent P300 brain responses and, consequently, in more accurate classification of the user's intentions than the flashing stimulus.

Task Complexity and BCI Performance

Most prior work using the P300 paradigm has been conducted in laboratory settings and involved relatively simple tasks. Our goal is to enable control of complex applications, some of which will enable users to engage in cognitively challenging tasks. It has been observed, however, that cognitive load induced by the task itself may interfere with the performance of a brain-computer interface [9]. We intend to investigate this effect and its implications for the design of real brain-controlled applications. Specifically, we intend to address two questions. First, does the cognitive load induced by the primary task affect the performance of the brain-computer interface? Second, does the need to invest attention into the operation of the brain-computer interface affect the person's short term memory performance?

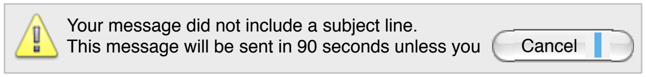
The answers to the first question will most likely affect the specifics of the classification algorithm's design and the number of repetitions of the stimuli before a selection decision is made. These results will help ensure that the BCI can be used reliably for controlling realistic applications.

The answers to the second question will affect the inclusion and the design of explicit short term memory aids in brain-controlled interfaces. This may be accomplished by providing an explicit history of recent actions performed with the user interface or by visualizing recent actions with an afterglow effect like the one proposed in Phosphor [1] (Figure 1(c)).

Error Recovery and Actions With Side Effects

Because uncertainty in the classification of user intentions is inherent in brain-computer interfaces, and because user interface operations take a significant amount of time to perform, the undo and redo operations should be easily accessible in brain-controlled applications. This can be accomplished simply by presenting these functions at the top level of the interface rather than putting them inside menus.

Additional difficulties arise for actions with significant irreversible side effects such as printing, or making an on-line payment. These actions should only be performed when the system is certain of the user's intent. The fact that BCI input is noisy, makes this a particularly important requirement. A common approach in today's GUIs is to add an explicit confirmation step. The time and effort required to perform each user interface operation with BCI make this approach undesirable, however, because in typical situations, when the system correctly interprets the person's intent, this confirmation step will add costly overhead to an already very slow interaction. We believe that making irreversible actions *temporarily* reversible may address this apparent conflict between certainty and interaction efficiency: if a user is given certain amount of time to undo an action before it is actually executed, the confirmation is obtained implicitly and an explicit user intervention is needed only to undo an action:


Your message did not include a subject line.
This message will be sent in 90 seconds unless you

Conclusion

We are interested in understanding the properties and limitations of BCI as an input method for controlling complex applications. What would applications designed specifically for P300-based BCI look like? What are the practical limits to the complexity of such applications? How much will the operation of the BCI interfere with the primary task? We have briefly presented a few of the questions and potential design approaches we intend to pursue in our project.

References

[1] P. Baudisch, D. Tan, M. Collomb, D. Robbins, K. Hinckley, M. Agrawala, S. Zhao, and G. Ramos. Phosphor: explaining transitions in the user interface using afterglow effects. In *Proc. UIST '06*, pages 169–178, New York, NY, 2006. ACM.

[2] J. Borisoff, S. Mason, A. Bashashati, and G. Birch. Brain-computer interface design for asynchronous control applications: improvements to the LF-ASD asynchronous brain switch. *IEEE Transactions on Biomedical Engineering*, 51(6):985–992, 2004.

[3] L. A. Farwell and E. Donchin. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and Clinical Neurophysiology*, 70(6):510–523, Dec 1988.

[4] K. Z. Gajos, J. O. Wobbrock, and D. S. Weld. Improving the performance of motor-impaired users with automatically-generated, ability-based interfaces. In *Proc. CHI '08*, pages 1257–1266, New York, NY, USA, 2008. ACM.

[5] S. Harada, J. O. Wobbrock, and J. A. Landay. Voicedraw: A voice-driven hands-free drawing application. In *Proc. ASSETS'07*. ACM Press, 2007.

[6] J. Hill, J. Farquhar, S. Martens, F. Bießmann, and B. Schölkopf. Effects of Stimulus Type and of Error-Correcting Code Design on BCI Speller Performance. In *Proc. NIPS'09*, pages 665–672, 2009.

[7] J. Mankoff, A. Dey, U. Batra, and M. Moore. Web accessibility for low bandwidth input. In *Proc ASSETS'02*, pages 17–24, New York, NY, USA, 2002. ACM.

[8] S. M. M. Martens, N. J. Hill, J. Farquhar, and B. Scholkopf. Overlap and refractory effects in a brain-computer interface speller based on the visual p300 event-related potential. *J. Neural Eng.*, 6(2), Apr 2009.

[9] M. Moore. Real-world applications for brain-computer interface technology. *IEEE T Neur Sys Reh*, 11(2):162–165, Jan 2003.

[10] M. Moore, P. Kennedy, E. Mynatt, and J. Mankoff. Nudge and shove: frequency thresholding for navigation in direct brain-computer interfaces. In *Proc. CHI '01*, pages 361–362, New York, NY, USA, 2001. ACM.

[11] T. Picton. The P300 wave of the human event-related potential. *Journal of clinical neurophysiology*, 9(4):456, 1992.

[12] R. Rauschenberger. Attentional capture by auto- and allo-cues. *Psychonomic Bulletin & Review*, 10(4):814–842, Dec 2003.

[13] J. O. Wobbrock, B. A. Myers, and J. A. Kembel. EdgeWrite: a stylus-based text entry method designed for high accuracy and stability of motion. In *Proc. UIST '03*, pages 61–70, New York, NY, USA, 2003. ACM.