Coming in touch with computing systems: Implicit interaction in human-machine systems

Matthias Rötting, Sandra Trösterer, Thorsten Zander und Jeronimo Dzaack

Keywords: eye-movements, EEG, multimodal interaction, human-machine systems

Zusammenfassung

Der Beitrag beschreibt einen Forschungsansatz zur Integration von impliziten Informationen des Menschen in die Mensch-Maschine-Interaktion. In einem ersten Schritt wird der neue Aspekt der impliziten Interaktion beschrieben und mit bestehenden Forschungsansätzen in Bezug gesetzt. Darauf aufbauend führen wir zwei unterschiedliche Formen der impliziten Interaktion ein: (1) blickbasierte Interaktion (gaze-based interaction) und (2) auf Gehirnaktivitäten basierende Interaktion (brain-computer interaction). Beide Interaktionsformen werden anschließend kurz diskutiert.

Abstract

This article describes our approach to integrate implicit information into human-machine interaction. In a first step we define the new aspect of implicit interaction and relate it to existing approaches of human-machine interaction. Building on this we introduce in a second step two different ways of interaction with technical systems: (1) gaze-based interaction and (2) brain-computer interaction. Both kinds of interaction are discussed shortly in the end.

Introduction

Imagine you click on a file on your computer by mistake. The computer processes the information and starts to open the corresponding application. But this takes some time. You immediately recognize your mistake and prepare to close the application right after it opens to continue your intended task. You feel distracted and helpless and your feelings are accompanied by facial expressions and inner thoughts. How would it be if the technical system could understand your mistake by analyzing selective information of you, the user? Like humans do in face-to-face communication, the system would recognize your mistake almost as soon as you did and adapt accordingly.

Human-machine interaction

Modern human-machine systems are not able to detect to what extend an action actually concurs to the users' wishes because the system does not have any means to get a real-time estimation of the situation and the user. To interrupt or end an erroneous action the user can only issue a new command to the system. Current human-machine systems lack the ability to detect implicit information as humans normally do in human face-to-face interaction.

Human-human interaction contains both explicit and implicit information. Humans use both kinds of information to analyze a situation and in this way two-way adaptation of behavior is supported, information is exchanged and the behavior in a group is regulated. In this context we define explicit interaction as a conscious action to exchange information, for example language and script. Implicit interaction can be defined as an unconscious action that is integrated in another action, for example mimic and gesture.

We believe that human-machine systems should be enhanced by channels providing the technical system with psycho-physiological and behavioral data of the users (e.g., brain activity, eye movements, and peripheral muscles). This would enable the system to access information that is implicitly generated by the human. On the one hand this integrates implicit aspects into human-machine interaction leading to a more human way of interaction. On the other hand this allows technical systems to extend their classical way of directive interaction by the interpretation of the selective implicit aspects. Both aspects can help (1) to meet the expectations of users of how interaction should be designed and how it should take place and (2) to reform the way how technical systems react on external manipulations by humans (e.g., automation, adaption, interpretation).

Implicit gaze-based interaction

In gaze-based interaction eye movements of a human are recorded and processed in real-time in order to control a computer system. The majority of existing gaze-based interaction systems is still designed to replace the computer mouse with gaze information. The location of the mouse cursor is replaced by the coordinates of the current fixation and the mouse click is replaced by fixations longer than a certain threshold (the so called dwell time). So whenever a certain element is fixated for a long enough period, a certain action is carried out. These systems are mostly designed for handicapped users, who can only use their gaze in order to communicate.

How implicit is the gaze-based interaction of such existing systems? Obviously, the degree of implicit gaze-based interaction depends on the application area. For example, when using the gaze to enter text on a screen, this has much to do with intention. The easiest way to enter text using only the gaze is to display a visual keyboard on a computer screen. A normal QUERTY keyboard is shown on the monitor and a letter can be entered by fixating it for a short interval (e.g. Istance et al., 1996). When using any such system, the user will move the eyes very intentionally to spell the word he or she has in mind. An example for implicit use of eye movements is the iDict system (Hyrskykari, 2006). iDict is a gaze-aware reading aid that helps to understand foreign words. The system is based on the assumption that a user fixates an unknown word longer than a known word. The system recognizes the change in fixation duration and then automatically displays a translation proposal.

These examples show that in gaze-based interaction implicit as well as explicit information can be derived from the analysis of eye movements. The examples given so far look primarily at information derived from single fixations. Whenever an object or element is fixated, it can be concluded that it is of certain interest for the user and that it is the current focus of his or her attention. Unfortunately, we cannot derive from fixations alone of what kind this interest is, although parameters like fixation frequency or fixation duration provide information about importance, difficulty of extraction, and interpretation of information (Jacob & Karn, 2003). More and different information can be derived from analyzing successions of saccades (i.e. very fast ballistic movements that direct the eye to an object) and fixations, the scan path. From the scan path, assumptions about the intention of the observer can be derived. The classic experiments by Yarbus (1967) show that eye movements are strongly influenced by the tasks a subject is instructed to do. Kahnemann (1973) refers to this as “task-relevant looking”. The level of workload the user experiences can be judged by the degree of structure of the eye movement behavior and changes in fixation durations. For example, fixation duration shortens with increasing complexity of aircraft (Gerathewohl et al., 1978) or car piloting-tasks (Unema, Rötting & Luczak, 1988), and eye movement behavior is more influenced by environmental saliencies when workload is increased (Trösterer & Dzaack, 2007).

What does this mean for gaze-based human-machine interaction? If information about eye movement behavior is provided to the machine, it could use this information in order to react or adapt better to the user. Prerequisite is of course that the machine is able to interpret the eye movement data in relation to the elements of the graphical user interface. For example, if an application icon is fixated, the corresponding application could be started in the background, assuming that the interest in the icon will be followed by a click to start the application. If the machine recognizes a special gaze pattern, the intention of the user could be inferred and the appropriate action for that case could be carried out. It might also be possible that the machine proposes help or a break if it recognizes that the eye movements of the user are getting less structured, indicating an increased level of user workload.

Enhancing Human-Machine Systems with implicit input from passive Brain-Computer Interfaces

For the development of new implicit information channels between human and machine, accessing the human brain seems to be a very promising approach. Methods from the field of Brain-Computer Interfaces (BCI) based on statistical machine learning have proven to be very successful for achieving this goal. A BCI allows gaining information about the cognitive state of the user without the necessity of the user to produce any activity outside the central nervous system. The first approach of using brain activity directly for communication comes from Vidal in the early 70s (Vidal 1973). First steps into the field of application were realized by Birbaumer and Wolpaw in the 80s and 90s (Wolpaw et al. 2002) by building brain activity-based support systems for patients suffering from amyotrophic lateral sclerosis (ALS). The application of methods from statistical machine learning (Blankertz, Curio & Mueller 2002) for BCI had a deep impact on the classification accuracy. Subsequently, it was possible to transfer the learning effort from the human being to the machine. Hence, it minimized the effort that is needed to develop the skills to control a BCI system. When the focus is laid on the applicability of BCI while interacting in typical human-machine system environments we suggest to use the term Brain-Computer Interaction instead of Brain-Computer Interface. User states hardly inferable from exogenous factors, so called covert user states (CUS, Reissland et al., 2009), are estimated by the BCI, interpreted and fed back into the technical part of the system. The information about the user states can be handled as explicit or implicit commands.

The methods derived from BCI research can be categorized as active and reactive. The term active BCI (aBCI) denotes BCIs which utilize brain activity of direct correlates of intended actions as input. A reactive BCI (rBCI) is still controlled via intended actions. In contrast to the aBCI, features are not derived from direct correlates to these actions, but from cognitive reactions on exogenous stimuli, as for example in the speller developed by Wolpaw et al. (2002). According to this line of thought, a third class of BCIs, passive BCI (pBCI), can be defined (Zander et al. 2008). Passive BCIs provides and feeds back information on CUS directly connected to current Human-Machine Interaction.

We could show that pBCIs could be applied in a very useful way in the field of HMS. It can be utilized to detect the CUS of the user perceiving a loss of control over a system (Jatzev et al., 2008), to enhance automated adaptation by detecting the CUS of the user is interpreting an error within the behaviour of the machine (Zander et al., 2008) or the CUS of bluffing in a game context (Reissland et al., 2009).

Discussion and Outlook

Our experiences with pBCIs show that they enable new channels of information within the interaction between human and machine. The passive BCI approach aligns with the idea of implicit interaction in general. Within this context, the utilization of information directly inferred from cognitive processes allows for completely new insights into and applications based on the interaction of human and machine.

Similarly, eye movements can be used to derive information about the interest, the intention and the workload of a person interacting in a human-machine system.

The combination of both techniques, BCI and eye movements, could result in even greater advantages compared to the use of one technology on its own. This form of multi-modal interaction using explicit and implicit exchange of information opens a new and wide range of applications. It might lead to new types of machines that allow an intuitive and more natural interaction between a user and the machine.

References

Blankertz, B., Curio, G. & Mueller, K.-R. (2002) Classifying single trial EEG: Towards brain computer interfacing. *In Advances in Neural Inf. Proc. Systems, vol. 14, pp. 157–164*

Gerathewohl, S.J., Brown, E.L., Burke, J.E., Kimball, K.A., Lowe, W.F. & Strackhouse, S.P. (1978). Inflight measurement of pilot workload: A panel discussion. *Aviation, Space, and Environmental Medicine*, 810-822.

Hyrskykaria, A. (2006). *Eyes in Attentive Interfaces: Experiences from Creating iDict, a Gaze-Aware Reading Aid.* PhD thesis, University Of Tampere, Finland. Verfügbar unter: http://acta.uta.fi/pdf/951-44-6643-8.pdf [09. Juli 2009].

Istance, H. O., Spinner, C. & Howarth, P.A. (1996). Eye-based Control of Standard GUI Software. In MA. Sasse, J. Cunningham & R.L. Winder (Hrsg.), *Proceedings of HCI on People and Computers XI* (pp. 141-158). London: Springer Verlag.

Jacob, R.J.K. & Karn, K.S. (2003). Eye Tracking in Human-Computer Interaction and Usability Research: Ready to Deliver the Promises (Section Commentary). In J. Hyöna, R. Radach & H. Deubel (Hrsg.), *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research* (pp. 573-605). Amsterdam, Elsevier Science.

Jatzev, S., Zander, T. O., DeFlilippis, M., Kothe, C., Welke, S., Roetting, M.(2008). Examining causes for non-stationarities: The loss of controllability is a factor which induces non-stationarities. Proceedings of *4th Int. BCI Workshop & Training Course*, TU Graz Publishing House.

Kahnemann, D. (1973). *Attention and effort.* Englewood Cliffs: Prentice Hall.

Reissland, J., Kothe, C., Jatzev, S., Welke, S., Gärtner, M. and Zander, T.O., (2009). BCI detection of complex covert user states - or: Can we detect bluffing in a game?. *Proceedings of* 8. Berliner Werkstatt MMS, Berlin, Germany.

Trösterer, S. & Dzaack, J. (2007). Cognitive workload as a means to distract top-down visual search in dynamically changing interfaces? In *ECEM 2007, 14th European Conference on Eye Movements, August 19-23, Potsdam, Germany – Conference Abstracts* (p. 44).

Unema, P., Rötting, M. & Luczak, H. (1988). Eye Movements and Mental Workload in Man-Vehicle Interaction. In A.S. Adams, R.R. Hall, B.J. McPhee & M.S. Oxenburgh (Hrsg.), *Ergonomics International 88, Proceedings of the Tenth Congress of the Inter-national Ergonomics Association* (S. 463-465). Sydney: Ergonomics Society of Australia Inc.

Vidal, J. J. (1973) Toward direct brain-computer communication. *Annual Reviews in Biophysics and Bioengineering, 2:157–180.*

Wolpaw, J.R., Birbaumer, N., McFarland, D.J., Pfurtscheller, G. & Vaughan, T.M. (2002) Brain-computer interfaces for communication and control. *Clinical neurophysiology, 113:767–7.*

Yarbus, A.L. (1967). *Eye movements and vision.* New York: Plenum Press.

Zander, T. O., Kothe, C., Welke, S. and Roetting, M. (2008). Enhancing human-machine systems with secondary input from passive brain-computer interfaces. Proceedings of *4th Int. BCI Workshop & Training Course*, TU Graz Publishing House.