Multi-modal human-computer interfaces for handicapped user groups: Integrating mind and gaze

Andrea Finke
CoR-Lab
Bielefeld University
Bielefeld, Germany
afinke@cor-lab.uni-bielefeld.de

Hendrik Koesling
CITEC
Bielefeld University
Bielefeld, Germany
hendrik.koesling@uni-bielefeld.de

Helge Ritter
CoR-Lab and CITEC,
Bielefeld University
Bielefeld, Germany
helge@techfak.uni-bielefeld.de

ABSTRACT

Human-computer interfaces (HCIs) that are not uni-modal, but combine two or more input modalities, could facilitate interaction with technical devices – in particular for handicapped users. We implemented a multi-modal HCI that takes into account brain activity and eye movements, realizing a brain-eye computer interface for controlling a driving simulator. The system detects typical cortical signals from a motor-imagery task in EEG data and translates them into steering commands. The application also analyses eye fixations on a speed controller to set the driving speed. Both EEG and eye-movement data are processed on-line and allow for controlling the application in real time. The instantaneous feedback and the hands-free interaction by simply thinking and looking is a great advantage for users with immobilities. They can now interact with their environment via an interface that is adapted to their needs. The present implementation provides a starting point for other alternative multi-modal HCIs. These may include additional input modalities to account for users' special requirements.

Author Keywords

Human-computer interaction, multi-modal interface, brain-computer interface, EEG, eye tracking.

ACM Classification Keywords

H.1.2. Models and principles: User/machine systems: Human information processing.

General Terms

Algorithms, Experimentation, Measurement.

INTRODUCTION

The development and evaluation of multi-modal human-computer interfaces (HCIs) has been on the agenda of many researchers for some time. The idea of implementing "natural" ways of communicating with machines is certainly promising, even more so when combining various communication modalities in a way that humans normally do when interacting with each other. However, usability studies and evaluations of HCIs with able-bodied users have shown that this group is rather proficient and, in fact, efficient, in using established standard HCIs such as computer keyboards. Performance, e.g. text typing speed [15], and user satisfaction [10] are usually much higher for standard devices than for alternative HCIs, such as gaze-operated virtual keyboards - even after training.

For handicapped user groups, however, the situation is different. Handicapped people rely on alternative methods to interact or communicate with their environment and, specifically, with technical devices. For example, those with immobilities in their upper or lower limbs, are potential users of gaze interfaces. Another group, locked-in patients, can only communicate with their environment via a brain-computer interface (BCI)[11]. With many different such special groups that all require their own interaction interfaces or a combination thereof, specifically adapted to their special needs and individual capabilities, these users are amongst the main target groups for the development of novel multi-modal human-computer interfaces.

Most novel HCI methods were only developed recently and, to date, little effort has been made to combine these methods in multi-modal interfaces for special-needs groups. We propose that in particular the combination of such novel methods allows people with specific deficits to optimise the use of their (remaining) resources and make communication and interaction – and thus everyday life – easier for them.

The present paper focusses on the combination of brain- and gaze-controlled interfaces. It presents a starting point for the implementation and testing of other combinations of interaction methods, not necessarily restricted to combining only two at a time.

BRAIN-COMPUTER INTERFACES AND EEG

The advent of brain-computer interfaces (BCI) in the last two decades has opened a new pathway for providing input to a machine based solely on cortical activity without the need for physical motor control. Of course, the potential field of applications is much wider than the one sketched

UbiComp '10, Sep 26-Sep 29, 2010, Copenhagen, Denmark.

for locked-in patients. One outstanding feature is the non-invasive nature of BCIs based on EEG measurements. The current field of applications of EEG-based BCI systems ranges from wheelchair navigation [14] to neurofeedback training for children with attention deficit disorders (ADD)[7]. However, a major drawback of today's BCI systems is their limited number of control dimensions which are restricted by the number of discriminable activation patterns identified in a person's EEG. These in turn are limited due to the low signal-to-noise ratio of EEG signals and the spatial resolution of the measurements. A typical BCI provides between 2 and 4 control states.

GAZE-BASED INTERFACES AND EYE TRACKING

Various types of hands-free human-computer interfaces use oculo-motor parameters from eye-movement recordings. Users can guide mouse cursors with their eyes [9], select keys from virtual keyboards by eye fixations or blinks [10] and choose web page contents [1] by visual attention. As oculomotor control is generally not affected by motor-control deficits concerning manual operations, many groups of patients can use gaze-controlled interfaces.

A MIND AND GAZE CONTROLLED BI-MODAL INTERFACE

We have developed and implemented a bi-modal human-computer interface based on the simultaneous measurement of EEG and eye movements. The integral parts of this interface are firstly, a modular, on-line BCI system, and secondly, an eye-tracking system that allows for real-time data acquisition and processing of eye movements. In our lab, we use a fully mobile, 16-channel EEG amplifier (gUSBamp, Guger Technologies) for EEG measurement and an SR Research EyeLink II eye tracker for the registration of eye movements.

The BCI part uses motor imagery, a particular neural signature that is correlated with the movement of a limb (e.g., a hand) without the execution of an actual movement. The somatotopic organization of the sensory-motor cortex enables a discrimination of signals referring to different limbs. We chose this component for the interface, because in contrast to other EEG components common in brain interfaces like the P300 potential [5] or steady-state visually evoked potentials (SSVEP) [12] it is to some degree "continuous" and independent of a timed stimulus presentation. A disadvantage is that motor imagery is a "mental skill" a new user needs to train for successive sessions before he/she can use a BCI efficiently. However, state-of-the-art feature extraction and classification algorithms are able to cope with data from only little trained subjects [2]. We usually train potential users in about three sessions to gain a reasonable degree of volitional control over their sensory-motor cortex activity. For training, a simple feedback program is used, where the subject has to move a virtual ball on a screen to either the left or right side by imagining a movement of the corresponding hand.

Human visual patterns usually consist of sequences of saccades and fixations. Saccades rapidly shift the centre of high visual acuity (fovea) to the focus of visual attention and do not allow for visual processing (saccade duration typically between 20-100 ms [4], mainly depending on saccade amplitude). Visual information is processed during fixations where the eye remains stable for a certain length of time, usually between 120-300 ms [17] and depending on the stimulus and task. Eye movements are to some extent controlled bottom-up by visual saliency features of the scene such as colour, contrast or intensity [8], but they are also driven by top-down processes such as the given task and are thus largely subject to the observer's volitional control [19]. We exploit this fact by utilizing these volitionally controlled fixations to control elements of a user interface.

Figure 1. A participant testing the novel bi-modal interface. EEG and eye gaze data are acquired and processed on-line for controlling the driving simulator.

A COMPUTER GAME FOR INTERFACE BENCHMARKING

In order to create an appealing benchmark application/system for evaluating the novel bi-modal interface, we implemented a 3-D car driving simulator based on the open-source game platform TORCS (http://torcs.sourceforge.net). Computer games are well suited to test novel HCIs because they require fast and often complex input. The control of the game is fully accomplished by mind and gaze alone. The fusion of both modalities is implemented in the game controller. By doing so, we ensure that the system architecture is fully modular and thus easily extendable to a multi-modal interface by adding further data sources.

In a car driving simulator, the control of the virtual car splits up into two components: steering to the left and right and speed adjustment. In order to implement these controls with the brain-eye interface, we use mind control for steering and gaze control for acceleration or deceleration, respectively. The left-right steering commands are mapped from the user's imagination of left or right hand movements. Speed control is realized by an accelerator slider bar that was added to the original control elements. This slider bar is operated gaze-contingently and uses the vertical component of the gaze position for adjusting the driving speed.

METHODS

Feature extraction and classification of EEG data

Feature extraction and classification of EEG data relies on supervised learning, making it necessary to acquire training data before the system can be used efficiently. We use common spatial patterns (CSP) for feature extraction, a spatial filtering and dimension reduction technique that has proven efficient for motor-imagery based BCIs [3]. CSP provides a projection matrix obtained by solving the generalized eigenvalue problem for the whitened covariance matrices of the two classes [16]: $S_1V = S_2V\lambda$ where $S_1 = P\Sigma_1P$ and $S_2 = P\Sigma_2P$. The filtering matrix is then obtained by $W = V^TP$. Covariance matrices are computed using a shrinkage estimate based on [18], who follow the Ledoit-Wolf theorem [13]. Only those dimensions of W (determined by the eigenvalues λ) are selected that represent the maximal separability of the classes with respect to the subsequent classification step. The final feature vectors are created by $\hat{x}_t = log(var(X_tW^T))$, X_t being a matrix of dimension $T \times C$, where T is the number of samples per segment and C is the number of channels. The vectors $\hat{x}_t$ usually have a dimension of 6 to 10, depending on the individual data set, i.e. the subject. Classification is done on a single-trial basis. We use FDA [6], a linear classifier where the result is obtained by projecting the feature vector onto a weight vector: $y_t = w^T\hat{x}_t + b$, b is the bias. As the computation of w requires again to estimate the covariance matrices we use a shrinkage estimate as before. Class labels can be found by simply taking the sign of y_t. However, we apply the real-valued y_t as inputs for the game controller.

Mapping classification results onto steering commands

The y_t are translated into "turn" commands for the virtual car in the game. The game platform allows to set the turning angle with the full range from $-180°$ to $+180°$. We, however, omit extreme angles due to the altered control modalities and fix the range from $-140°$ to $+140°$. The corresponding range y_{min} to y_{max} is determined from the training data. The mapping is done in a continuous manner, that is, we use the full turning interval instead of pre-defined turning steps.

Mapping gaze coordinates onto speed commands

We choose not to instantaneously set the slider bar in the speed adjustment area to the current vertical component of the gaze coordinate when the gaze moves into this area. This would lead to abrupt changes in the car's speed and makes smooth acceleration and deceleration difficult. Instead, adjustments to the current speed of the car can be achieved by looking above and below the actual slider bar location, i.e., at the new "target" speed, for acceleration and deceleration, respectively. As long as the gaze remains in the "interaction" areas, the car's speed as well as the slider bar location are adjusted accordingly. When the gaze leaves the interaction areas, the currently adjusted speed remains constant until gaze re-enters the interaction areas. The vigor of acceleration depends on the difference between current and target speeds and is adapted according to a linear function (hard acceleration for large speed differences, smooth for small differences). Under full acceleration, every 100 ms of fixating either of the interaction areas results in a speed change of 20 km/h while speed can be adjusted within a range of 0 to 180 km/h.

CONCLUSION

We tested the interface with five able-bodied participants in our lab. For comparison, we started each session by letting the user control the simulator the standard way, namely with the mouse and the keyboard. Afterwards, the game control was switched to the novel interface, such that the users could test the alternative control modalities. We asked them to complete once the given road track (preselected by the experimenter and identical for all participants). All five participants were able to complete the track using mind and gaze control at, however, significantly slower completion times than for manual control. For handicapped users, who depend on alternative input methods to interact with a machine at all, the relative slowness of the control scheme should not be an issue, as long as the control is stable and reliable.

The proposed bi-modal interface fuses two state-of-the-art distinct alternative HCI methods in one interface. By combining different modalities, like neural activity and eye gaze and possibly others to novel multi-modal human-computer interfaces, more flexible systems can be created that are suited to a wide range of applications and a larger group of users. In particular handicapped users depend on such alternative methods. Human-computer interaction has become an essential part of today's daily life. Providing powerful, alternative interfaces is an indispensable condition to enable handicapped people to gain a satisfactory level of autonomy in their everyday life.

ACKNOWLEDGEMENTS

The study was partially funded by the German Research Council (DFG), EXC 277. Andrea Finke gratefully acknowledges the financial support from Honda Research Institute Europe.

REFERENCES

1. R. Bates and H. Istance. Why are eye mice unpopular? a detailed comparison of head eye controlled assistive technology pointing devices. *Universal Access in the Information Society*, 2:280–290, 2003.
2. B. Blankertz, G. Dornhege, M. Krauledat, G. Curio, and K.-R. Mueller. The non-invasive berlin brain-computer interface: Fast acquisition of effective performance in untrained subjects. *NeuroImage*, 37(2):539–550, 2007.
3. B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K.-R. Mueller. Optimizing spatial filters for robust EEG single-trial analysis. In *IEEE Signal Processing Magazine*, pages 581–607, 2008.
4. J. Findlay. Programming of stimulus-elicited saccadic eye movements. In K. Rayner, editor, *Eye movements and visual cognition: Scene perception and reading*, pages 8–30, New York, 1992. Springer.
5. A. Finke, A. Lenhardt, and H. Ritter. The MindGame: A P300-based brain-computer interface game. *Neural Networks*, 22:1329–1333, 2009.
6. R. Fisher. The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, 7:179–188, 1936.

7. C. Gani, N. Birbaumer, and U. Strehl. Long term effects after feedback of slow coritcal potentials and of theta-beta-amplitudes in children with attention-deficit/hyperactivity disorder. *International Journal of Bioelectromagnetism*, 10(4):209–232, 2008.

8. L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11:1254–1259, 1998.

9. R. Jacob. The use of eye movements in human-computer interaction techniques: What you look at is what you get. *ACM Trans. Inf. Syst.*, 9:152–169, 1991.

10. H. Koesling, M. Zoellner, L. Sichelschmidt, and H. Ritter. With a flick of the eye: Assessing gaze-controlled human-computer interaction. In H. Ritter, G. Sagerer, R. Dillmann, and M. Buss, editors, *Cognitive Systems Monographs, Human Centered Robot Systems: Cognition, Interaction, Technology*, pages 83–92, Berlin, 2009. Springer.

11. A. Kuebler, B. Kotchoubey, J. Kaiser, J. R. Wolpaw, and N. Birbaumer. Brain-computer communication: Unlocking the locked in. *Psychological Bulletin*, 127(3):358–375, 2001.

12. E. C. Lalor, S. P. Kelly, C. Finucane, R. Burke, R. Smith, R. B. Reilly, and G. McDarby. Steady-state VEP-based brain-computer interface control in an immersive 3D gaming environment. *EURASIP Journal on Applied Signal Processing*, 19:3156–3164, 2005.

13. O. Ledoit and M. Wolf. A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, 88(2):365–411, February 2004.

14. R. Leeb, D. Friedman, G. R. Müller-Putz, R. Scherer, M. Slater, and G. Pfurtscheller. Self-paced (asynchronous) bci control of a wheelchair in virtual environments: a case study with a tetraplegic. *Intell. Neuroscience*, 2007(2):1–12, 2007.

15. P. Majaranta and K. J. R"aih"a. Text entry by gaze:utilizing eye tracking. In I. MacKenzie and K. Tanaka-Ishii, editors, *Text entry systems: Mobility, accessibility, universality*, pages 175–187, San Francisco, 2007. Morgan Kaufmann.

16. H. Ramoser, J. Mueller-Gerking, and G. Pfurtscheller. Optimal spatial fitering of single trial EEG during imagined hand movement. *IEEE Transactions on Rehabilitation Engineering*, 8(4):441–446, Dec 2000.

17. K. Rayner. Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124:372–422, 1998.

18. J. Schaefer and K. Strimmer. A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. *Statistical Applications in Genetics and Molecular Biology*, 4(1):Article 32, 2005.

19. G. Underwood, T. Foulsham, E. van Loon, L. Humphreys, and J. Bloyce. Eye movements during scene inspection: A test of the saliency map hypothesis. *European Journal of Cognitive Psychology*, 18:321–342, 2006.