Symbiotic Mind Computer Interaction for Information Seeking

ICT-611570

Deliverable 2.3

Report on physiology-informed user model for information seeking

Authors	Markus Wenzel, Patrik Pluchino, Benjamin Blankertz
Reviewer	–
Version	f1.1 (final version)
Date	30.11.2016
Classification	Public
Contract Start Date	01.10.2013
Duration	36 months
Project Co-ordinator	University of Helsinki
File Name	MindSee_D2.3_f1.1_TUBerlin

Project funded by the
European Community under
the Seventh Framework Programme

Consisting of:

No	PARTICIPANT NAME	S.N.	COUNTRY
1	University of Helsinki	UH	FI
2	Aalto University	AALTO	FI
3	University of Padova	UNIPD	IT
4	i2media research ltd	i2media	UK
5	TU Berlin	TUB	DE

Responsible of the document: TU Berlin

Defined Contributors to the document: Markus Wenzel, Patrik Pluchino, Benjamin Blankertz

Other Expected Contributors: –

Document History

VERSION	ISSUE DATE	BY	CONTENT AND CHANGES
v1.0 Draft	17/10/2016	TU Berlin	First draft
v1.1 Draft	26/10/2016	TU Berlin	Improved descriptions
f1.0 Final	28/10/2016	TU Berlin	Final document
f1.1 Final	30/11/2016	TU Berlin	Final document – improved the appendix (respectively references to the appendix in the text)

Executive Summary

PROJECT SUMMARY

The MindSee (Symbiotic Mind Computer Interaction for Information Seeking) project's main objective is to exemplify the fruitful symbiosis of modern BCI technology with a recent real-world HCI application to obtain a cutting-edge information retrieval system that outperforms state-of-the art tools by more than doubling the performance of information seeking in realistic tasks.

More specifically MindSee will develop techniques to co-adapt user and computing systems utilizing implicit information such as EEG phenomena coupled with peripheral physiology during search to detect comprehensive user responses ranging from preconscious perception, cognition and emotion. Such implicit information feeds into a user model in reinforcement learning allowing to co-adapt exploration and exploitation in information seeking in relevance of information, complexity and aesthetic aspects. MindSee will capitalize on recent advances in BCI combined with peripheral physiological signals (EDR, facial EMG, eye gaze and pupillometry) for an unobtrusive registration of user states at a highly resolved level with respect to perception, cognition and emotions. Combined with machine learning based co-adaptation approaches this will allow predicting user intention for an unprecedented man-machine symbiosis.

DELIVERABLE SUMMARY

In this deliverable, we report in particular about the results of Task 2.5 of work package 2: "Physiology-informed optimization of user models for information seeking" (cf. the title of the deliverable presented here). Our approach to tackle this task can be summarized as follows. We determine in real-time which screen content attracts the interest of the individual user, based on EEG data guided by an eye tracker. The resulting implicit information about the attention allocation of the user allows to estimate the parameters of a model of the user interest, e.g. as part of a software application for information seeking. The task has been accomplished by addressing key questions in different experimental investigations, which resulted, so far, in three international, peer-reviewed journal publications.

In addition, we report about the results from Task 2.6 "The Preconscious Pupil", as announced in the second project periodic report ("PPR II").

Table of Contents

1 Progress for the Year and Individual Efforts

1.1 Progress

The progress during year 3 is presented by comparing the pursued and the achieved objectives of work package 2 and by describing the different investigations carried out.

1.1.1 Pursued objectives

The objectives of work package 2 are the following, according to the project proposal:

1) Extension of psychophysical studies of perception and cognition toward free view tasks.

2) Identification of EEG-features as suitable input for real-life applications such as information seeking.

3) Multi-modal integration of neurophysiology, peripheral physiology and behavioural data.

4) Demonstration of usage of the physiology-informed approach for information seeking.

1.1.2 Achieved objectives

1) Several experimental studies with unrestricted eye movements ("free viewing") were conducted, while electroencephalography (EEG) and eye tracking were recorded in order to infer the perception and cognition of the participants.

2) The data, recorded during the studies, were analysed in detail and EEG-features were identified that are suited for real-life applications in the context of human-computer interaction, in particular for information seeking.

3) In the different experimental studies, neurophysiological, peripheral physiological and behavioural data were integrated.

4) The conducted studies show how implicit information contained in physiological data can be used to infer the interest of the individual user in real-time, and how the gained information can be exploited during information seeking.

In particular, we report the results of Task 2.5 of work package 2: "Physiology-informed optimization of user models for information seeking" (cf. the title of the deliverable presented here). In addition, we report about the results from Task 2.6 "The Preconscious Pupil", as announced in the second project periodic report ("PPR II").

1.1.3 Physiology-informed optimization of user models for information seeking

Our approach to tackle the task "physiology-informed optimization of user models for information seeking" can be summarized as follows. We determine in real-time which words or images displayed on the screen attract the interest of the individual user, based on EEG data guided by an eye tracker. The gained implicit information about the attention allocation of the user allows for estimating the parameters of a model of the user interest. The resulting dynamic user model can be exploited by a software application, e.g. for information seeking. The task has been accomplished by addressing key questions in several experimental investigations.

1.1.4 Is Neural Activity Detected by ERP-based Brain-Computer Interfaces Task Specific?

It is known from previous research about brain-computer interfacing (BCI) that screen content which attracts the interest of a person results in a particular neural response that can be detected in the electroencephalogram (EEG). We aim to transfer this inference technique to relevance detection in human-computer interaction (HCI).

The screen content in brain-computer interfacing typically consists in sequential stimuli. For instance, different geometrical shapes represent different letters of the alphabet and are repeatedly flashed in the centre of the screen [e.g. Treder et al., 2011]. The BCI user mentally selects a letter and silently counts the appearance of the corresponding geometrical shape in order to focus the attention. The shape of interest evokes a neural response, which can be detected in the EEG. The approach allows inferring the mentally selected letter from the EEG and thereby writing with the "power of thought". It is conceivable that the technique can be expanded beyond a writing application and that any screen content that attracts the attention, because it is relevant to the objectives of the user, can be detected with a BCI.

For the proposed generalization to HCI, it is required that the BCI inference technique can serve as a general relevance detector, and is not closely linked to a specific task. While it is legitimate to direct the attention of a BCI user (who is paralysed and has no alternative means of communication) with a counting task, this is not the case in an HCI scenario. Computer users do not perform a specific task, like silent counting, each time a word or an image is recognized as relevant. Instead, some screen content simply attracts the attention of the users because it is considered as relevant to their respective objectives. Physiology-informed relevance detection in HCI can only enhance the use of mouse and keyboard in a meaningful way, if the information is *implicitly* contained in the sensor data – because a simple selection of the items of interest with the computer mouse might make more sense than wearing an EEG cap and marking the items of interest with a specific task such as silent counting.

Accordingly, it was checked if the EEG-based prediction of relevant screen content relies on a specific task, or if it captures rather the attention allocation and can therefore generalize to different scenarios. Result of the experimental study is that the neural activity used for the relevance prediction is not strictly task specific. The patterns detected in the neural data likely reflect the attention allocation or the workload and not merely neural processes related to a specific task. The result may hold promise for a generalization of BCI methods to implicit relevance detection in HCI.

Full details of the experiment and of the results are published in

Wenzel MA, Almeida I and Blankertz B (2016). Is Neural Activity Detected by ERP-based Brain-Computer Interfaces Task Specific? *PLoS ONE* **11**(10): e0165556. http://dx.doi.org/10.1371/journal.pone.0165556

1.1.5 From sequential stimuli to eye tracker guided EEG analysis

The screen content in brain-computer interfacing typically consists in sequences of single stimuli, as mentioned in the previous section. In contrast, computer users, operating a web browser or other software, are confronted with many words, pictures and pictograms that are displayed side by side at the same time. The single items are not flashed one by one but are fixated sequentially with the eye gaze.

Therefore, guidance by an eye tracker is needed, in order to determine which elements of the screen content attracted the interest of the user. First, it is necessary to know the position of the eye gaze to relate the ongoing neural processes to the respective words or pictures looked at. Second, BCIs based on event-related potentials rely on time markers of reference for the feature extraction from the continuous EEG signal. The landing of the eye gaze on a word or picture can serve for this purpose.

The problem that arises here is that the timing of the neural processes with respect to this time marker of reference can be variable. It is conceivable that salient screen content can be already detected in peripheral vision, i.e. before the landing of the eye gaze on the item, while less salient screen content may be recognized as relevant or irrelevant only after the landing. This potential jitter of the neural response with respect to the eye movements can be a problem for the inference from the EEG with state-of-the-art BCI methods that had been developed for sequences of flashed stimuli.

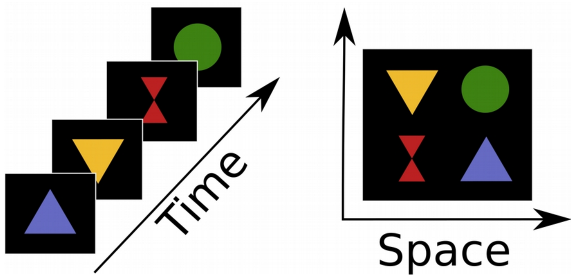

Figure 1. Time sequential stimuli versus stimuli displayed side by side.

Accordingly, it was investigated in an experimental study if the BCI paradigm can be adapted from time sequential stimuli to an eye tracker guided EEG analysis (cf. Figure 1). EEG features, following the onsets of the eye fixations, could be singled out that are informative about the task relevance of the fixated items. Furthermore, it was possible to cope with the temporal variability of the neural response due to the variable salience. Besides, we could show that neural processes were captured despite of the unrestricted eye movements that can potentially interfere with the EEG signals. The final results of the study are published in

Wenzel MA, Golenia JE and Blankertz B (2016). Classification of eye fixation related potentials for variable stimulus saliency. *Frontiers in neuroscience* **10**. http://dx.doi.org/10.3389/fnins.2016.00023

Graphics in computer software are dynamic with moving items and with new visual content fading in slowly or popping-up suddenly. Different dynamics can result in a temporal variability of the neural processes with respect to the eye movements, which serve as time markers of reference for the EEG analysis. In an experimental study, graphics with different visual dynamics were presented while EEG and eye tracking data were recorded. Relevant screen content could be inferred from the recorded data with analysis techniques that can cope with the mentioned temporal variability. The final results are published in

Ušćumlić M and Blankertz B (2016). Active visual search in non-stationary scenes: coping with temporal variability and uncertainty. *Journal of neural engineering* **13**(1), 016015. http://dx.doi.org/10.1088/1741-2560/13/1/016015

(Preliminary results of both studies were introduced in deliverable 2.1.)

1.1.6 Towards realistic screen content

Simple geometric shapes serve typically as stimuli in ERP-based brain-computer interfacing [e.g. Treder et al., 2011] and therefore also in the previously introduced studies. However, when transferring techniques from BCI to HCI, it has to be considered that the screen content in real software applications, e.g. a search engine running in a web browser, is not limited to geometrical shapes but can contain pictures and words. Accordingly, the investigations were expanded towards realistic screen content in two experimental studies.

In the first study, the relevance of photographies was inferred in real-time from the EEG guided by an eye tracker. An image search engine, such as *google image search* or *Flickr*, was imitated where several images are displayed side by side in a mosaic on the screen. The participants were asked to seek pictures belonging to a certain category (one of two possible interpretations of an ambiguous search term). It was estimated in real-time which images were relevant for the user and which images were irrelevant, based on electroencephalography and eye tracking. The information gained per picture was accumulated in order to determine which of the two possible interpretations of the ambiguous search term was of particular interest. The result constitutes a basic, bimodal, dynamic model of the interest of each individual user. The outcome of the inference of this online demonstrator application, serving for a proof-of-concept, was presented as feedback to the user.

Details of an early version of the study were introduced for the first time in milestone MS4 "Demonstrator of real-time EEG-input for information seeking".

Subsequently, technical details were published in a conference contribution.

Golenia JE, Wenzel MA and Blankertz B (2015). Live Demonstrator of EEG and Eye-Tracking Input for Disambiguation of Image Search Results. In *Symbiotic Interaction* Springer. http://dx.doi.org/10.1007/978-3-319-24917-9_8

The short preliminary conference contribution was expanded by recording the data of more participants and by analysing the data in much more depth. The results of the study were summarized in a manuscript and submitted to the Journal of Neural Engineering. The journal editor did not accept the manuscript for publication but encouraged the authors to resubmit a revised version, which addresses the concerns of the reviewer (work in progress).

Golenia JE, Wenzel MA and Blankertz B (2016). Disambiguation of image search queries with EEG and eye tracking. *In preparation*.

[Available on request.]

In a second experimental study, it was tested if the relevance inference from EEG and eye tracking, previously tested for geometric shapes and pictures, can be expanded to word reading. A demonstrator application, serving for a proof-of-concept, was developed that mimicked (cf. Figure 2) the graphical user interface of a search engine for scientific literature ["SciNet"; Glowacka et al., 2013]. The developed application is able to estimate the relevance of single words to a semantic category of interest, in real-time based on EEG and eye tracking. The underlying software system was developed in a modular fashion in view of an integration with other software of the "MindSee" project. The fifteen participants of the study read various words displayed side by side on the screen and looked for words belonging to one out of five semantic categories. The words were dynamically replaced by new words fading in when they had been fixated with the eye gaze. It was estimated which words belonged to the chosen category of interest, based on implicit information contained in the EEG and eye tracking data. The relevance estimates for the single words were summarized per semantic category, which resulted in a five dimensional model of the user interest that was updated as soon as a new word had been read. The model was visualised online in this demonstrator application. For the visualisation, the category names were listed at the side of the screen. The font size and the hue of the category names represent the parameter values of the five dimensional user interest model and were updated online. A paper manuscript is being prepared (cf. Appendix):

Wenzel MA, Bogojeski M and Blankertz B (2017). Real-time inference of word relevance from electroencephalogram and eye gaze. *Submitted*.

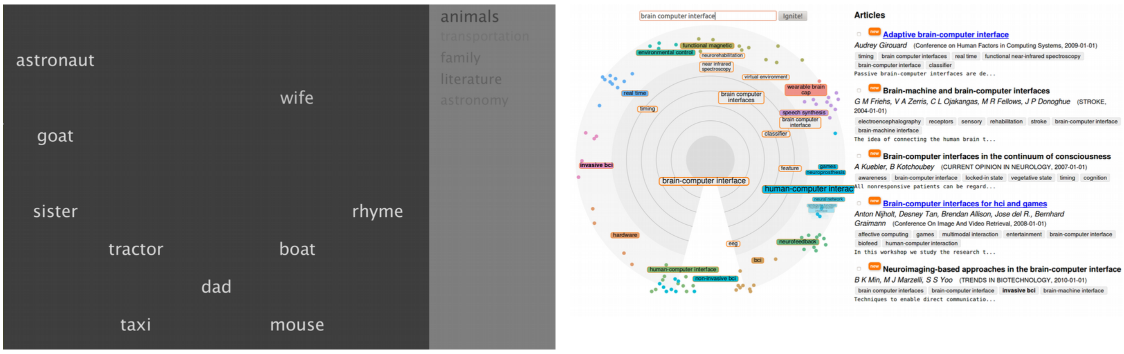

Figure 2. ***Left:*** **Looking for words related to one out of five semantic categories.** ***Right:*** **The interface of the search engine** ***SciNet*** **was mimicked.**

1.1.7 Summary

We accomplished the task "physiology-informed optimization of user models for information seeking" by building a machinery that is able to estimate in real-time which words or pictures attract the interest of the individual user, based on EEG data guided by an eye tracker. With this approach, we learn the attention allocation and the interest of the user. The resulting physiology-informed user model can allow software to adapt dynamically to the inferred interest of the user. A search engine could, for instance, refine the retrieval procedure based on the user model and present particularly relevant search results on the top of the list.

The approach exploits situations that allow integrating implicit information across several single items (images or words). Reading composite key words or texts may entail a number of new interesting challenges for the data analysis.

1.1.8 Semantic priming and early modulations in pupil behaviour

In this deliverable, we also report about the results from Task 2.6 "The Preconscious Pupil", as announced in the second project periodic report ("PPR II").

Objective. Pupil behaviour was evaluated in the context of a semantic categorization task. Early potential changes in the pupil size were of particular interest for the investigation. In the literature it has been shown that pupil diameter variations can reflect a variety of cognitive processes through different types of tasks (e.g., memory tasks, maths calculations, etc.). The pupil dilates as the level of cognitive load increases (Beatty & Kahneman, 1966; Van Orden, Jung, Makeig, 2000; Iqbal, Zheng, & Bailey, 2004; Bailey & Iqbal, 2008). These fluctuations in pupil's size are usually slow (around 1-1.5 seconds).

Approach. In an experimental study, the effects were examined of unconsciously presented word primes on the semantic categorization of word targets, while the pupils of the twenty participants were monitored. The words belonged either to the "living" or to the "non-living" semantic category. A priming paradigm was employed in order to estimate variations in the maximal dilation and constriction of the pupils and in their latencies. These indices were considered in an early temporal window, namely from the presentation of the target word until the moment in which the participants provided their responses.

Results. The latencies of maximal dilation and constriction of the pupils were shorter when the pair of target and prime was "living"-"living". The behavioural response reaction time was faster when the pair was "living"-"living" compared to all the other prime-target couples. The behavioural categorization accuracy was high for all prime-target combinations.

Significance. Pupillometry indices can potentially entail themselves as suitable means to augment the human-computer interaction. Symbiotic systems could assist users in accomplishing their tasks, by taking into consideration both the pupil dilation, in order to avoid cognitive overload, and the pupil constriction, in order to exploit also the cognitive and pre-cognitive processing in visual-short term memory.

Details of the study are presented in the appendix:

Patrik Pluchino (2016). Semantic priming and early modulations in pupil behaviour.

1.2 Individual Efforts

Task n.	Task title	Partner
2.5	Physiology-informed optimization of user models for information seeking	TUB
2.6	The Preconscious Pupil	UNIPD

2 References

Bailey BP & Iqbal ST (2008). Understanding changes in mental workload during execution of goal-directed tasks and its application for interruption management. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *14*(4), 21.

Beatty J & Kahneman D (1966). Pupillary changes in two memory tasks. *Psychonomic Science*, *5*(10), 371-372.

Glowacka D, Ruotsalo T, Konuyshkova K, Kaski S & Jacucci G (2013). Directing exploratory search: Reinforcement learning from user interactions with keywords. In *Proceedings of the 2013 international conference on Intelligent user interfaces* (pp. 117-128). ACM.

Iqbal ST, Zheng XS & Bailey BP (2004). Task-evoked pupillary response to mental workload in human-computer interaction. In *CHI'04 extended abstracts on Human factors in computing systems* (pp. 1477-1480). ACM.

Treder MS, Schmidt NM & Blankertz B (2011). Gaze-independent brain–computer interfaces based on covert attention and feature attention. *Journal of neural engineering*, *8*(6), 066003.

Van Orden KF, Jung TP & Makeig S (2000). Combined eye activity measures accurately estimate changes in sustained visual task performance. *Biological psychology*, *52*(3), 221-240.

Our contributions

Golenia JE, Wenzel MA and Blankertz B (2015). Live Demonstrator of EEG and Eye-Tracking Input for Disambiguation of Image Search Results. In *Symbiotic Interaction* Springer. http://dx.doi.org/10.1007/978-3-319-24917-9_8

Ušćumlić M and Blankertz B (2016). Active visual search in non-stationary scenes: coping with temporal variability and uncertainty. *Journal of neural engineering* **13**(1). 016015. http://dx.doi.org/10.1088/1741-2560/13/1/016015

Wenzel MA, Almeida I and Blankertz B (2016). Is Neural Activity Detected by ERP-based Brain-Computer Interfaces Task Specific? *PLoS ONE* ***11****(10): e0165556.* http://dx.doi.org/10.1371/journal.pone.0165556

Wenzel MA, Golenia JE and Blankertz B (2016). Classification of eye fixation related potentials for variable stimulus saliency. *Frontiers in neuroscience* **10.** http://dx.doi.org/10.3389/fnins.2016.00023

3 Appendix

The appendix contains the

- confidential unpublished paper manuscript *Wenzel MA, Bogojeski M and Blankertz B* (2017). *Real-time inference of word relevance from electroencephalogram and eye gaze. Submitted*.
- and details of the study on *Semantic priming and early modulations in pupil behaviour* by *Patrik Pluchino*.