



HAL
open science

Enhancing Visuospatial Attention Performance with Brain-Computer Interfaces

Romain Trachel, Thomas Brochier, Maureen Clerc

► **To cite this version:**

Romain Trachel, Thomas Brochier, Maureen Clerc. Enhancing Visuospatial Attention Performance with Brain-Computer Interfaces. ACM SIGCHI Conference on Human Factors in Computing Systems, Apr 2013, Paris, France. pp.ACM 978-1-4503-1952-2. hal-00817799

HAL Id: hal-00817799

<https://hal.science/hal-00817799v1>

Submitted on 25 Apr 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Enhancing visuospatial attention performance with Brain-Computer Interfaces

Romain Trachel^{1,2}
romain.trachel@inria.fr

Maureen Clerc²
maureen.clerc@inria.fr

¹ Aix-Marseille Université
CNRS, INT UMR 7289
Marseille, France

Thomas Brochier¹
thomas.brochier@univ-amu.fr

² Athena Project-Team
INRIA Sophia Antipolis -
Méditerranée
Sophia Antipolis, France

Abstract

Visuospatial attention is often investigated with features related to the head or the gaze during Human-Computer Interaction (HCI). However the focus of attention can be dissociated from overt responses such as eye movements, and impossible to detect from behavioral data. Actually, Electroencephalography (EEG) can also provide valuable information about covert aspects of spatial attention. Therefore we propose an innovative approach in view of developing a Brain-Computer Interface (BCI) to enhance human reaction speed and accuracy. This poster presents an offline evaluation of the approach based on physiological data recorded in a visuospatial attention experiment. Finally we discuss about the future interface that could enhance HCI by displaying visual information at the focus of attention.

Author Keywords

Brain-Computer Interfaces, Visual Attention, Single-Trial Classification, Electroencephalography

ACM Classification Keywords

H.1.2 [User/Machine Systems]: Human information processing; H.5.2 [User Interfaces]: Evaluation/methodology, Input devices and strategies (e.g., mouse, touchscreen)

General Terms

Experimentation, Human Factor, Measurement

Introduction

Visuospatial attention is a selection process that allows for continuously allocating cognitive resources to specific locations in the environment. Visual processing at the attended location is enhanced while reactions at competing locations are actively impaired. The focus of attention can be shifted to visual targets either overtly by a saccade or covertly without eye movements [5]. In Human-Computer Interaction (HCI), the tracking of overt changes in the head or gaze orientation is often used to evaluate shifts in visuospatial attention. However in complex visual displays, the spatial focus of attention can be dissociated from overt responses (e.g., eye movement) and impossible to capture with behavioral measures. In this situation, Electroencephalography (EEG) can be used to provide valuable information about the covert modulations of spatial attention. Therefore we explore the possibility of enhancing human reaction speed and accuracy in a visuospatial attention task by detecting covert shifts of visual attention.

This poster presents an offline analysis of physiological data recorded in ten human subjects in view of designing a Brain-Computer Interface (BCI). Using EEG features to classify covert shifts of attention in two opposite directions (e.g., left vs right), the proposed BCI is evaluated by selecting trials in which the focus of attention coincides with a visual target location. The BCI is calibrated using trials in which the target location can be predicted from spatial cues and validated with ambiguous cues providing incoherent spatial information. In perspective of this work, we propose the development of interactive systems that could adapt the visual display at the focus of attention.

Related Work

Recent research in neuroscience is expanding the potential of BCI technology from medical to non-medical domains. An important challenge in HCI is to develop applications for enhanced human performance, using physiological data to provide to the computer real-time information about the user's cognitive state [9]. Nowadays, improvements in BCI usability and system integration make this challenge possible. Recent studies showed that modulation of the posterior EEG activity is a reliable marker of voluntary shifts of attention that can be used for discrete [7, 6] or continuous [1] real-time control of a two-dimensional cursor. Our work aims at investigating how visuospatial modulations of attention could be used to enhance the behavioral performance in complex visual environments.

Experiment and Methods

The experimental protocol was designed using the Matlab PsychToolbox¹ and the Data Acquisition toolbox. Physiological data were analyzed using the MNE² and the sklearn³ Python toolbox.

Experimental Procedure

Ten subjects (males, aged between 22 and 44) performed a visuospatial attention task in a Posner paradigm [5]. The experiment was organised in 3 parts including a training session (5 min), two adaptive procedures (40 trials each) and the recording session (1 training block + 8 blocs of 52 trials). Brain activity was recorded at 1024 kHz on 64 EEG channels mounted on an elastic cap using the BioSemi ActiveTwo system⁴. Eye movements were recorded by Electrooculography (EOG) using bipolar

¹<http://psychtoolbox.org>

²<http://martinos.org/mne/mne-python.html>

³<http://scikit-learn.org>

⁴<http://www.biosemi.com>

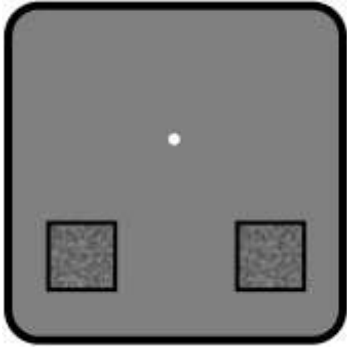


Figure 1: Screen view

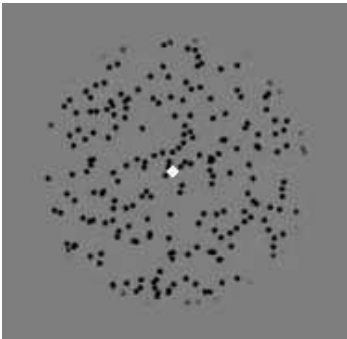


Figure 2: Spatial Cue

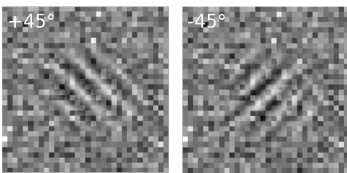


Figure 3: Target stimuli tilted at $\pm 45^\circ$

electrodes placed on each side of the forehead (horizontal component) and between the right eye and a frontal electrode (vertical component).

During the experiment, subjects were comfortably seated at 60 cm distance of a 20" CRT monitor (resolution 800×600 at 144 Hz). The goal of the task was to react as accurately and as fast as possible to visual targets displayed in the lower left or right location of the screen while gazing at a central dot (Figure 1). For each trial, a spatial cue (Figure 2) made of a cloud of moving dots was presented for 250 ms, followed by a variable delay of 1.5 to 2.5 seconds before a target stimulus (Figure 3) was displayed at cued location or at the opposite location for 70 ms. Subjects were instructed to first identify the dominant direction of the dots motion and to subsequently focus their attention towards the location indicated by this direction. The trial was correct if the subject pressed the button with the thumb of the left (resp. right) hand when the target was tilted at $+45^\circ$ (resp. -45°). Following the response, subjects received either a positive feedback or an error message and reported if the target appeared at the location to which they shifted attention. The behavioral performance in terms of Reaction Time (RT) and Error Rate (ER) was evaluated in two different conditions. In the predictive condition (75% of trials), the spatial cue provides clear information about the target location, inducing shifts of attention in the correct direction most of the time. In the ambiguous condition (25% of trials), the cue does not provide any coherent information, reducing behavioral performances because by probability, attention is shifted to the inappropriate location in half of the trials. Predictive and ambiguous trials were interleaved and presented in a pseudorandom order to ensure subjects could not associate the cue with its predictability.

Adaptive Psychological Procedure

Before the recording session, each subject performed two adaptive procedures [8] that determined the cue coherence (inverse of the variance of the distribution of dots direction) and the stimulus contrast (intensity). The first adaptive procedure ensures the subject can extract the cued location in at least 80% of trials over the two conditions [5]. The coherence threshold was adapted to individual subject's perception of coherent dots motion and applied in the predictive condition. The coherence is set to zero in the ambiguous condition inducing a shift of attention in the wrong direction for half of the trials. The second adaptive procedure is used to set the level of performance in the target identification task when attending at the correct location. In each trial, the contrast of the stimulus was varied independently of the noise, displayed at fixed contrast (35%). The contrast and coherence threshold were estimated so that the subject could achieve 90% correct responses. On average across subjects, thresholds were 1.29 and 17.80%.

Signal Processing and Feature Extraction

Raw signals were pre-filtered to 1-40 Hz and single trials were epoched 1.5 seconds before the target display. Epochs contaminated by eye movements were rejected if the EOG peaks exceeded $50 \mu\text{V}$. EOG features were computed with the mean signal amplitude for each epoch and classified to see if eye movements provided information about the focus of attention. EEG channels were referenced to the common average and band-pass filtered in all frequency bands between 5 to 35 Hz for each band width with 2 Hz shift. This processing step resulted in a set of pre-filtered signals $X(f_1, f_2) \in \mathbb{R}^{C \times T}$ with f_1 and f_2 the low and high cut off frequencies, C the number of electrodes and T the number of time samples for each epoch.

A regularized version of the Common Spatial Pattern (CSP) algorithm [4] was applied in order to extract band power features for left and right shifts of attention to be classified with the EEG signals [2]. The CSP is a supervised projection method that maximizes the variance of spatially filtered signals $X_{csp} \in \mathbb{R}^{C \times T}$ for one class of trials while minimizing it for the other. This method computes a set of spatial filters $W = \{w_j \in \mathbb{R}^C\}_{j=1, \dots, C}$ by solving the generalized eigenvalue problem $\Sigma_l W = \lambda \Sigma_r W$, with Σ_l (resp. Σ_r) the covariance of the filtered signals $X(f_1, f_2)$, for left (resp. right) trials. Following Lotte and Guan [4], we estimated the covariance matrix of the signals with an algorithm that computes the optimal L2-norm regularization parameter (see OAS in the sklearn.covariance package). Then, the band-pass filtered signals were projected from the original sensor space to the surrogate sensor space by $X_{csp}(f_1, f_2) = W^T X(f_1, f_2)$. The CSP patterns were computed by $A = (W^T)^{-1}$ for each subject. Finally, band power features were computed as the log variance of these spatially filtered signals for each trial.

Single-Trial Classification and Selection

The frequency band and the number of CSP features are calibrated using a 5-folds Cross Validation (CV) procedure applied to predictive trials. For each filtered signal, the set of 312 trials is split into a training set (4/5 of trials) and a testing set (1/5 of trials). The training set is used to extract the CSP features, train the linear Support Vector Machine (SVM) classifier and apply a Recursive Feature Elimination (RFE) procedure [3]. The RFE procedure recursively selects 90% and prunes 10% of the features, given weights assigned by the SVM classifier. RFE is repeated in an inner 5-fold CV to rank each feature set by the SVM classification error rate. The optimal L2 regularization parameter is also optimized by a F1 score

computed with an inner CV grid search in the train set. The test set is used to compute the SVM accuracy as a function of the corresponding features iteratively selected in the RFE ranking. Finally, the classification performance is averaged across 5 repetitions of the calibration for every permutation of the folds. The frequency band and the number of features are optimized by minimization of the average classification error. The classification accuracy in ambiguous trials is evaluated in a validation procedure for each subject. In this procedure, predictive trials are pooled in the training set to extract the CSP features, train the SVM and compute its optimal regularization parameter. Then each ambiguous trial is classified with the corresponding features and selected if the SVM decision (right or left shift of attention) matches with the target location (left or right box). Behavioral performance enhancement is evaluated by comparing trials selected with CSP or EOG features, and trials selected by a random procedure which is repeated 1000 times. The statistical significance (p-value) of the result is computed for each subject as the probability of producing lower RT and ER with the random selection.

Results

The results were computed on average across subjects (mean \pm standard deviation). The significance was computed with a Wilcoxon rank-sum test (z) for paired comparisons and Spearman's rank coefficient (r) for correlation analyses.

Behavioral Performance

On average across subjects, the focus of attention was shifted to the target location in $90.49 \pm 3.39\%$ of the trials in the predictive condition and $55.58 \pm 9.41\%$ in the ambiguous condition (z=-3.78; p=0.00016). When attention was shifted to the target location, averaged RT

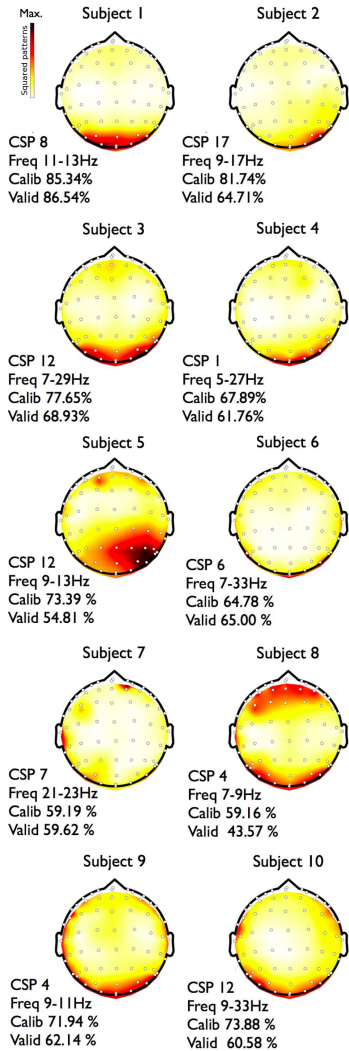


Figure 4: Topographic view of mean CSP patterns for each subjects

and ER were 0.725 ± 0.130 seconds and $6.42 \pm 5.72\%$ in the predictive condition, and 0.861 ± 0.178 seconds and $6.32 \pm 4.92\%$ in the ambiguous condition. For the opposed location, averaged RT and ER were 1.40 ± 0.402 seconds and $48.08 \pm 20.84\%$ in the predictive condition, and 1.28 ± 0.279 seconds and $33.04 \pm 15.68\%$ in the ambiguous condition. As expected, both RT and ER were enhanced when attention was shifted to the target location in the predictive ($z=-3.55$; $p=0.00038$ for RT and $z=-3.63$; $p=0.00028$ for ER) and the ambiguous condition ($z=-2.95$; $p=0.0032$ for RT and $z=-3.78$; $p=0.00016$ for ER). However, this attentional effect is larger in the predictive than in the ambiguous condition.

Classification Performance

Classification accuracy obtained for left vs. right shifts of attention with EOG features is near chance level with $52.01 \pm 2.69\%$ accuracy for predictive trials and $49.87 \pm 4.93\%$ for ambiguous trials. The accuracy obtained with CSP features is $71.49 \pm 8.85\%$ for predictive trials and $63.16 \pm 10.11\%$ for ambiguous trials. The accuracy obtained for ambiguous trials is higher with CSP than EOG features ($z=3.25$; $p=0.001$). As expected 49% of the ambiguous trials are selected by CSP and 51% by EOG in the validation procedure. The mean CSP patterns, number of feature and frequency band extracted in predictive trials are displayed for each subject in Figure 4 with corresponding classification accuracy obtained for the calibration (Calib) and the validation (Valid) procedure. The mean CSP features of predictive trials is correlated with the classification accuracy obtained for calibration ($r=0.63$; $p=0.047$) and validation ($r=0.66$; $p=0.037$). Moreover, the mean CSP patterns averaged over the posterior sensors is also correlated with the classification accuracy ($r=0.91$; $p=0.0002$ for calibration and $r=0.56$; $p=0.09$ for validation). These results show that

covert shifts of attention in two opposite directions can be classified with features extracted in the oscillatory activity of EEG signals recorded over the posterior brain regions.

Behavioral Enhancement

The enhancement of RT and ER obtained by CSP selection are shown in Table 1 for each subject. RT and ER obtained by CSP selection are on average lower than the mean random selection by 56.06ms and 4.44%. Compared to EOG features, the enhancement would be weaker with RT enhanced by 9.68ms and ER enhanced by 1.25%. As expected from the validation procedure, subjects with higher classification accuracy in ambiguous trials present the largest enhancement with CSP features ($r=-0.72$; $p=0.019$ for RT and $r=-0.55$; $p=0.098$ for ER).

Subject	RT (ms)	p-value	ER (%)	p-value
1	-206.49	0.000	-18.58	0.000
2	-23.71	0.392	-2.80	0.185
3	-137.45	0.001	-8.78	0.000
4	-37.08	0.180	-3.74	0.047
5	7.11	0.602	-1.26	0.315
6	-64.37	0.106	-6.83	0.003
7	-69.92	0.095	-0.42	0.361
8	-1.38	0.502	-4.77	0.015
9	-84.41	0.113	2.63	0.652
10	57.01	0.880	0.11	0.420

Table 1: Enhanced RT and ER with CSP selection. Significant results in **bold** ($p \leq 0.005$, Bonferroni corrected).

Discussion and Future Work

This poster is a proof of concept which explores the use of BCIs to enhance the reaction speed and accuracy in a complex visuospatial attention task. The results obtained are encouraging in view of setting up a real-time BCI in

which the visual display could be adapted to the subject's shifts of attention. Therefore, we are developing a new experiment including a similar offline calibration procedure and two online validation protocols. The BCI is calibrated from data recorded with predictive spatial cues while the evaluation of the behavioral enhancement is validated in an ambiguous environment. In the first validation procedure, the visual target can be displayed at the focused location according to the classified shift of attention. Based on the accuracy results obtained offline, we can predict that the target will be displayed to the attended location more frequently than at random. In this situation, it is expected that the subject will readily anticipate the target location even with ambiguous spatial cues. Therefore, the BCI performance should be improved online from the subject motivation that may modulate the shifts of attention and the classifier accuracy. The second validation procedure is designed for a more ecological situation in which the target display cannot be manipulated but known to the system. When the shift of attention is not classified to the target location, we propose to flicker this location a few milliseconds before target presentation. This approach would capture the subject's focus of attention and enhance the visual processing of stimulations displayed at an unattended location. In perspective of this work, interactive systems that can adapt the visual displays to the user's focus of attention could be developed in HCI research.

Acknowledgements

The authors gratefully acknowledge the subjects for participating in the experiment. This work was supported by the French ministry of defence (DGA) and the CoAdapt ANR project (ANR-09-EMER-002).

References

- [1] Bahramisharif, A., van Gerven, M., Heskes, T., and Jensen, O. Covert attention allows for continuous control of brain-computer interfaces. *Eur. J. Neurosci.* 31, 8 (2010), 1501–8.
- [2] Fujisawa, J., Touyama, H., and Hirose, M. Extracting alpha band modulation during visual spatial attention without flickering stimuli using common spatial pattern. *IEEE Conference, EMBS (2008)*, 620–623.
- [3] Guyon, I., Weston, J., Barnhill, S., and Vapnik, V. Gene selection for cancer classification using support vector machines. *Mach. Learn.* 46 (2002), 389–422.
- [4] Lotte, F., and Guan, G. Regularizing Common Spatial Patterns to Improve BCI Designs: Unified Theory and New Algorithms. *IEEE Trans. Biomed. Eng.* 58, 2 (2011), 355–362.
- [5] Posner, M. Orienting of attention. *Quart. J. Exp. Psych.* 32, 1 (1980), 3–25.
- [6] Treder, M., Bahramisharif, A., Schmidt, N., van Gerven, M., and Blankertz, B. Brain-computer interfacing using modulations of alpha activity induced by covert shifts of attention. *J. NeuroEng. Rehab.* 8, 1 (2011), 24.
- [7] van Gerven, M., and Jensen, O. Attention modulations of posterior alpha as a control signal for two-dimensional braincomputer interfaces. *J. Neuro. Meth.* 179, 1 (2009), 78 – 84.
- [8] Watson, A., and Pelli, D. QUEST: a Bayesian adaptive psychometric method. *Percept. Psychophys.* 33, 2 (Feb. 1983), 113–120.
- [9] Zander, T., Kothe, C., Jatzev, S., and Gaertner, M. Enhancing human-computer interaction with input from active and passive brain-computer interfaces. In *Brain-Computer Interfaces, Human-Computer Interaction Series*. London, 2010, 181–199.