

---

# Proceedings of the 6th International Brain-Computer Interface Conference 2014

The Future of Brain-Computer Interaction: Basics, Shortcomings, Users

September 16-19 2014  
Graz University of Technology, Austria

---

Edited by

Gernot Müller-Putz, Günther Bauernfeind, Clemens Brunner,  
David Steyrl, Selina Wriessnegger, Reinhold Scherer

---

© 2014  
Verlag der Technischen Universität Graz  
<http://ub.tugraz.at/Verlag>



ISSN 2311-0422  
ISBN 978-3-85125-378-8  
DOI [10.3217/978-3-85125-378-8](https://doi.org/10.3217/978-3-85125-378-8)

## Preface

After the positive responses to the last event, first time held as a conference, we decided to keep the style and expand the event also with satellite events, since we want to provide a platform for scientific exchange within the Brain-Computer Interface research community. Here, researchers can present their own work either in the form of a talk or as a poster. For this purpose, we encouraged participants to submit papers, which were peer-reviewed and are published in this ebook. The BCI conferences held in Graz, Austria, may be considered to be a European initiative in the field of EEG-based Brain-Computer Interfaces that contributes to a stronger orientation towards scientific cooperation.

We are lucky that outstanding experts in the field, Dr. Scott Makeig (Swartz Center for Computational Neuroscience, Institute for Neural Computation, University of California San Diego), Prof. Nick Ramsey (UMC Utrecht, Rudolf Magnus Institute, Department of Neurology and Neurosurgery), and Manfred Halver (FFG - Austrian Research Promotion Agency, European and International Programmes) were able to accept our invitation to present keynote addresses at the conference.

During preparation of this conference many people have been involved. Here we want to acknowledge the work of the reviewers who contributed with their expertise and knowledge:

Allison Brendan	Bauernfeind Günther	Billinger Martin	Blankertz Benjamin
Brouwer Anne-Marie	Brunner Clemens	Brunner Peter	Chavarriaga Ricardo
Cichocki Andrzej	Cincotti Febo	Congedo Marco	Coyle Damien
Cuntai Guan	Daly Ian	Darvas Felix	Faller Josef
Farquhar Jason	Friedrich Elisabeth	Gao Shangkai	Grosse-Wentrup Moritz
Guger Christoph	Halder Sebastian	Horki Petar	Jin Jing
Kanoh Shin'ichiro	Kleih Sonja	Kothe Christian	Kreilinger Alex
Kübler Andrea	Lee Adrian Kc	Leeb Robert	Lotte Fabien
Mattia Donatella	McCane Lynn	Mehring Carsten	Müller Klaus-Robert
Müller-Putz Gernot	Nam Chang S.	Noirhomme Quentin	Ofner Patrick
Ojemann Jeffrey G.	Pinegger Andreas	Ramsey Nick	Ron Angevin Ricardo
Rupp Rüdiger	Ryan David	Scherer Reinhold	Seeber Martin
Silvoni Stefano	Sitaram Ranganatha	Solis Escalante Teo	Steyrl David
Tangermann Michael	Wagner Johanna	Wander Jeremiah	Wriessnegger Selina
Zander Thorsten	Zoltan Bart		

We gratefully acknowledge the support of the Graz University of Technology for providing the facilities and thank the staff of the Institute of Knowledge Discovery, BCI Lab, for their dedicated assistance.

### The Editorial Board

Gernot Müller-Putz, Günther Bauernfeind, Clemens Brunner, David Steyrl, Selina Wriessnegger, Reinhold Scherer



# An optimized auditory P300 BCI based on spatially distributed sound in different voices

Jing Jin<sup>1\*</sup>, Ian Daly<sup>2†</sup>, Minqiang Huang<sup>1</sup>, Yu Zhang<sup>1</sup> and Xingyu Wang<sup>1</sup>

<sup>1</sup>Key Laboratory of Advanced Control and Optimization for Chemical Processes, Ministry of Education, East China University of Science and Technology, Shanghai, China

<sup>2</sup>Brain embodiment lab, School of Systems Engineering, University of Reading, Reading, UK  
jinjingat@gmail.com, i.daly@reading.ac.uk

## Abstract

In this paper, a new paradigm is presented, to improve the performance of audio-based P300 Brain-computer interfaces (BCIs), by using spatially distributed natural sound stimuli. The new paradigm was compared to a conventional paradigm using spatially distributed sound to demonstrate the performance of this new paradigm. The results show that the new paradigm enlarged the N200 and P300 components, and yielded significantly better BCI performance than the conventional paradigm.

## 1 Introduction

Visual-based P300 brain-computer interfaces (BCIs) requires users' to have control over their gaze direction and, therefore, are not useable by blind patients or other patients who cannot maintain eye gaze [1]. In order to enlarge the group of users who may benefit from use of P300 BCIs, the audio-based P300 BCI was developed [2-3]. It has been shown that an audio paradigm using spatially distributed stimuli is better than using different pitches from one speaker [4]. Höhne et al (2011) presented a paradigm using high, medium, and low unnatural spatially distributed sounds from headphones [5]. However, these two studies did not show that this paradigm was better than a paradigm using same sound played by spatially distributed speakers. In our study, a new audio-based P300 BCI was designed by using three different natural sounds. The three different natural sounds were female, male, and child voices, which were played by six spatially distributed speakers. This paradigm was called the "123 pattern". Another paradigm using unnatural sounds (a "beep"), randomly played by six spatially distributed speakers, was used to evaluate the efficacy of our proposed pattern, this was called the "beep pattern".

## 2 Methods and Materials

### 2.1 Subjects

Twelve healthy right-handed subjects (7 male, 5 female, aged 21 to 25 years, mean age 23.2±1.0) participated in this study. All subjects signed a written consent form prior to this experiment and were paid 50 RMB for their participation. The local ethics committee approved the consent form and

---

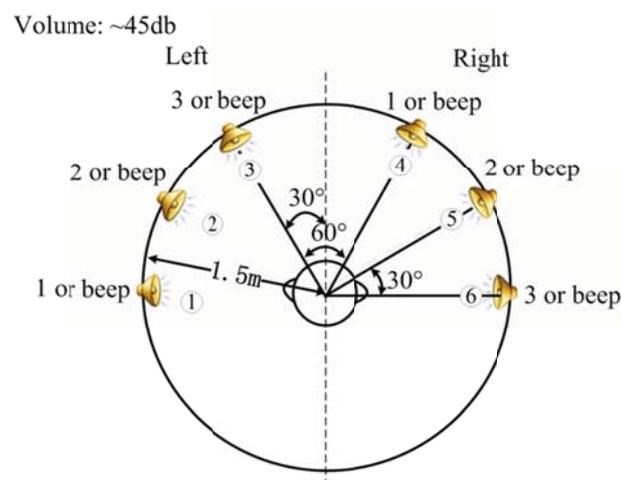
\* Corresponding author

† Corresponding author

experimental procedure before any subjects participated. All subjects' native language was Mandarin Chinese. Two subjects (labeled S7 and S9) had used a visual-based P300 BCI before this study.

## 2.2 Stimuli and flash patterns

Six speakers, placed in a semicircle around the participant, were used to present auditory stimuli (see figure 1). The distance between the participant and the speakers was 1.5 m and the distance between two adjacent speakers on the left and right side respectively was 30 degrees. The distance between the last speaker (speaker 3) on the left side and the first speaker (speaker 4) on the right side was 60 degrees (see figure 1). The labels of the six speakers were 1-6 from left to right. Speakers were calibrated to a common stimulus intensity of 45dB. In the experiment, participants were asked to close their eyes and pay attention to the target speaker and count the number of times that the speaker was played.



**Figure 1:** The location of speakers and sounds that were played by the speakers.

## 2.3 Experimental set up, offline and online protocols

EEG signals were recorded with a g.USBamp and a g.EEGcap (Guger Technologies, Graz, Austria) with a sensitivity of  $100\mu\text{V}$ , band pass filtered between 0.1Hz and 100Hz, and sampled at 256Hz. Data were recorded and analyzed using the BCI platform software package developed through East China University of Science and Technology. We recorded from 24 EEG electrode positions based on the extended International 10-20 system. The electrodes were F3, Fz, F4, T7, C5, C3, C1, Cz, C2, C4, C6, T8, CP5, CP3, CP1, CPz, CP2, CP4, CP6, P3, P1, Pz, P2, and P4. The right mastoid electrode was used as the reference, and the front electrode (FPz) was used as the ground. These electrode positions are more convenient for patients who are lying on a bed and, therefore, not able to easily move their head. Thus, the setup of the BCI was designed to allow rapid adoption by patient groups.

There are two conditions called “123 pattern” and “beep pattern” respectively. The order of the conditions was counterbalanced. In the 123 pattern, auditory presentation of the number “1” was played by speaker 1 in a female voice, auditory presentation of the number “2” was played by speaker 2 in a male voice, and auditory presentation of the number “3” was played by speaker 3 in a child’s voice. Speakers 1-3 were located on the left side. On the right side, auditory presentation of “1” was

played by speaker 4 in a female voice, auditory presentation of “2” was played by speaker 5 in a male voice, and auditory presentation of “3” was played by speaker 6 in a child’s voice. The auditory presentation of numbers was in Chinese. In the beep pattern, only a beep would be played by each of the speakers.

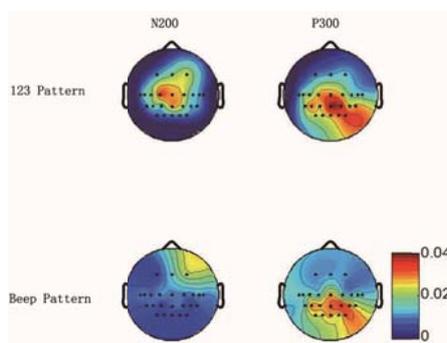
The stimulus on time was 200 ms and the stimulus off time was 440 ms. The inter-stimulus interval was 640 ms. In the experiment, each of the six speakers was played randomly and separately, and each speaker would be played once in one trial. In the offline experiment, there were twelve trials of stimuli in each run, which were needed for one character selection. Fifteen character selections were performed in the offline experiment for each pattern. In the online experiment, there were five trials of stimuli in each run, which were needed for one character selection. Twenty character selections were attempted in the online experiment for each pattern.

In the beginning of each run, there was an auditory cue (in Chinese) from the target speaker to guide the subject to locate the target speaker. The task of the subjects in the experiment was to count the number of times that the target speaker was played.

### 2.4 Feature extraction procedure

A third order Butterworth band pass filter was used to filter the EEG between 0.1 Hz to 30 Hz. The EEG was down-sampled from 256 Hz to 64 Hz by selecting every fourth sample from the filtered EEG. The first 1000 ms of EEG after presentation of a single stimulus was used to extract the feature.

## 3 Results



**Figure 2:** The topographic map of r-squared values for the 123 and beep patterns. The peak point value was used to calculate the r-squared value, which was obtained between 150-300 ms for the N200 and between 300-450 ms for the P300.

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	Avg
123-P(%)	75	90	80	80	90	10	80	80	70	90	60	95	75±22.7
B-P(%)	75	15	95	65	70	10	60	55	90	15	50	30	53±29.2

**Table 1:** Performance (classification accuracy) from online feedback runs across all 12 participants who participated in online runs. “123-P” denotes the 123 pattern and “B-P” the beep pattern.

Figure 2 shows the topographic map of r-squared values during presentation of the 123 pattern and the beep pattern. It shows that the 123 pattern obtained higher r-squared values than that of the beep pattern in the central region for the N200, and in the central and parietal regions for the P300.

Table 1 shows the online classification accuracy using five trials for constructing the average ERP waveform. Paired samples tests were used to show the difference between the two patterns. This shows that the classification accuracy of the 123 pattern is significantly higher than that of the beep pattern ( $t=2.3$ ,  $p<0.05$ ).

## 4 Discussion and conclusion

The goal of this study was to prove that an audio-based P300 BCI, using different pitches and spatially distributed speakers, could obtain higher classification accuracy than a comparable BCI using audio stimuli of the same pitch from spatially distributed speakers. Table 1 shows that classification accuracy of the 123 pattern was significantly higher than that of the beep pattern ( $t=2.3$ ,  $p<0.05$ ). The new paradigm used natural sounds, which would lead to high user acceptance and the new paradigm used spatially distributed sounds in difference voices, which allowed users to locate the target easily. However, a longer target to target interval was used in the audio-based BCI compared to the visual-based BCI, which would limit the communication speed of the audio-based BCI.

Neville and Lawson, (1987) explored ERPs from patients in response to spatially distributed tones [6]. Kübler et al., (2009) also tested the audio-based BCI on patients for consciousness assessment and communication [78]. In future, we will also focus on applying this system to patients who cannot control their eye gaze direction and, therefore, validate the performance of this audio-BCI system when it is used by target patient groups.

## 5 Acknowledgements

This work was supported in part by the Grant National Natural Science Foundation of China, under Grant Numbers 61074113, 61105122, 61203127 and 61305028 and was also supported by the Fundamental Research Funds for the Central Universities (WH1114038, WH1314023).

## References

- [1] Klobassa D S, Vaughan T M, Brunner P, Schwartz N E, Wolpaw J R, Neuper C, Sellers E W. (2009) Toward a high-throughput auditory P300-based brain-computer interface, *Clin. Neurophysiol.* **120**, 1252-1261
- [2] Hill N J, Lal T M, Bierig K, Birbaumer N, Schölkopf B (2004) An auditory paradigm for brain-computer interface NIPS, *books.nips.cc/papers/files/nips17/NIPS2004\_0503.ps.gz*
- [3] Furdea A, Halder S, Krusienski D J, Bross D, Nijboer F, Birbaumer N, Kübler A 2009 An auditory oddball (P300) spelling system for brain-computer interfaces, *Psychophysiology* **46**, 617-625.
- [4] Belitski A, Farquhar, Desain P 2011 P300 audio-visual speller *J. Neural Eng.* **8**, 025022
- [5] Höhne J, Schreuder M, Blankertz B, Tangermann M 2011, A novel 9-class auditory ERP paradigm driving a predictive text entry system *Front Neurosci* **22**,5:99
- [6] Beville H J, Lawson D 1987 Attention to central and peripheral visual space in a movement detection task III. Separate effects of auditory deprivation and acquisition of a visual language *Brain Res* **405**, 284-294
- [7] Kübler A, Furdea A, Halder S, Hammer E M, Nijboer F, Kotchoubey B 2009 A brain-computer interface controlled auditory event-related potential (p300) spelling system for locked-in patients *Ann N Y Acad Sci* **1157**, 90-100.

# A similitude-based BCI system for Communication

Anna Lisa Mangia<sup>1\*</sup> and Angelo Cappello<sup>1†</sup>

<sup>1</sup> Department of Electrical, Electronic and Information Engineering (DEI), University of Bologna, Cesena, Italy.

annalisa.mangia2@unibo.it, angelo.cappello@unibo.it

## Abstract

This work describes a procedure to design a similitude-based brain computer interface system for communication. Five healthy subjects and two patients with disorders of consciousness took part in the study. A support vector machine classifier applied to EEG data was used to detect answers to simple yes/no questions, while reducing the number of required electrodes. Just using ten electrodes we obtained a mean classification accuracy of 83.5% (SD 12%) for healthy subjects and 90% (SD 14.1%) for patients.

## 1 Introduction

One of the major concerns in recent studies is the discrimination between vegetative and minimally conscious state (Georgiopoulos, et al., 2010). The correct discrimination between these two conditions has major implication in subsequent rehabilitation of patients. In particular, establishing a communication with them would be advantageous and desirable. Several techniques, including functional magnetic resonance imaging (fMRI), cognitive event-related potentials (ERPs) and quantitative EEG analysis (qEEG), are currently developed to assess patients correctly and to attempt the communication with them (Monti, et al., 2010) (Cruse, et al., 2010) (John, et al., 2011). This work describes a procedure to investigate a subject's pattern of activation during mental imagery tasks. It aims to design a brain computer interface system for communication. Healthy subjects and patients with different levels of disorders of consciousness underwent EEG recording during yes/no personal questions. The first aim of the study was to develop a procedure of features selection in order to reduce the number of electrodes required. The second aim was to design a classifier that, after the training with two questions with known answers, is able to forecast the third unknown answer.

## 2 Methods

### 2.1 Subjects

Five healthy subjects (HS) (age 26 to 37) and two patients (P) took part in the study. The first patient was a 21-year-old female and her level of cognitive functioning (LCF) was 7. The second patient was a 28-year-old female and her LCF was 3. The etiology of the injury was traumatic for both patients.

---

\* Collected and analysed the data and created the first version of this document

† Analyzed the data and supervised this document

## 2.2 Protocol

The experiment consisted of a Communication Trial. Simple yes/no personal questions, with known answers, were asked to the subjects (e.g. “Are you married?”). Subjects were instructed to imagine for 30 seconds a movement of the right hand for an affirmative answer and a movement of the right foot for a negative answer. The Trial comprised six questions which were repeated six times for HS and twice for P. The experiment was repeated on two consecutive days (sessions) for HS only. No feedback was provided to the subjects.

## 2.3 EEG recording and signal processing

The EEG was recorded from 31 electrodes positioned according to the international 10-20 layout using a Neurowave System (Khymeia, Italy). EEG signals were band-pass filtered between 3 Hz and 60 Hz and underwent manual identification and rejection of artefactual segments. For each section, the epochs after the fourth second were eligible for the classification process. Power spectral density (PSD) was extracted from two seconds epochs without overlap. A modified periodogram method, based on FFT-algorithm and Blackman Harris window, was used. Subsequently, we averaged 5 values of the extracted PSD with a six seconds overlap, thus obtaining one PSD for every 10 seconds. The power in four frequency bands was extracted: theta (4-8 Hz), alpha (8-13 Hz), beta (13-25 Hz) and gamma (25-40 Hz). For each subject, each session and each answer, 31 (electrodes)  $\times$  4 (bands)  $\times$   $n$  (repetitions) sets were collected. Each value of the variable described above was labelled with the corresponding imagery task. The first aim of the study was to choose the Best 10 common Electrodes (BE) for all HS using a similitude criterion between equal answers. After the BE selection procedure we listed the subject-specific most significant features, in terms of Band-Electrodes Couples (BEC) for each HS and each P. The feature selection procedure used the same similitude criterion of the BE search. A certain number of BEC will be used in the classification process. For each HS, each P and each session we divided the dataset into two parts. The first one includes the 31 (electrodes)  $\times$  4 (bands)  $\times$   $n$  (repetitions) sets of the first half of the Communication Trial and it was used to select the BE and the BEC. The second includes the 10 (BE)  $\times$  4 (bands)  $\times$   $n$  (repetitions) sets of the remaining half of the Trial and it was used to classify the answers of HS and P.

## 2.4 Search of the Best Electrodes and Band-Electrode Couples

For each subject, on the basis of the given answers, all the sequences of three answers, with the first and the second one different (one hand and one foot movement imagery), were identified in the first part of the dataset. We will call these sequences of three answers sub-sessions. The number of sub-sessions was different from subject to subject. Considering the data of each HS and each sub-session, we computed a similitude index  $s_{i,j,k}$ :

$$s_{i,j,k} = \frac{P3_{i,j,k} - P1_{i,j,k}}{\sum_k P2_{i,j,k} - P1_{i,j,k}}$$

where  $i$  is the electrode index,  $j$  the band index,  $k$  the repetition index,  $n$  the number of repetitions, and  $P1$ ,  $P2$  and  $P3$  are the power of the first, second and third answer of the triple, respectively. If the third answer is the same as the first one,  $s$  tends to zero, while if the third one is equal to the second,  $s$  tends to 1. Using  $s$  we calculated the similitude between equal answers and we selected the 10 BE that optimize  $s$  for all HS. The classification of the third answer is performed using the following conditions:

$$\begin{cases} -0.5 - sd(s_{i,j}) < mean(s_{i,j}) < 0.5 - sd(s_{i,j}) & \& \quad sd(s_{i,j}) < 3|mean(s_{i,j})| \quad P3 = P1 \\ 0.5 + sd(s_{i,j}) < mean(s_{i,j}) < 1.5 + sd(s_{i,j}) & \& \quad sd(s_{i,j}) < 3|mean(s_{i,j})| \quad P3 = P2 \end{cases}$$

The aim of this preliminary selection on the HS was to reduce the problem dimensionality. After this selection, we applied again the same criterion for each HS and each P separately, with the aim to find a list of BEC according to the similitude criterion.

## 2.5 Classification Performance

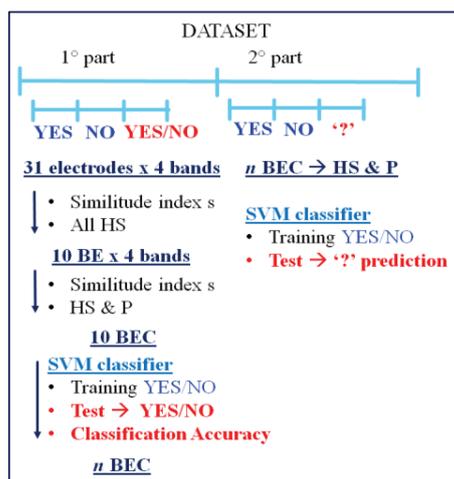


Figure 1: Selection and Classification previously selected. The first and the second answers were used to train the SVMc and the third one was used to test it (Figure 1). Each answer having a different number of repetitions, the class of attribution was decided by counting yes/no responses.

The second aim of the study was to establish a mean of communicating with the subject by detecting his/her answer to simple yes/no questions. We designed a classifier trained with two questions with known answers that was able to forecast the third unknown answer. After each sub-session of three questions, the classifier will be retrained with new data. We used the first part of the dataset to select the number of BEC to be considered. For each HS, each P and each session, a linear SVM classifier (SVMc) was used to train and to test all the triples of the first half of the dataset using a variable number, from 1 to 10, of BEC from the ordered list. The classification accuracy was computed and the number of BEC with the higher accuracy was selected. Subsequently we considered the second half of the dataset. For each HS, each P and each session we trained and tested all the triples using the number of BEC

## 3 Results and Discussion

### 3.1 Search of the Best Electrodes and Band-Electrode Couples

Considering all HS and all sessions, the ordered list of the ten BE was: PO3, Fc2, C3, O1, Fc1, Cz, Fz, P3, PO4 and T6. We found that the BE are mainly located in the fronto-central and parieto-occipital cortex. This confirms the results of previous studies demonstrating activation of motor cortex and parietal cortex during the execution of motor imagery task (Ishizu, Noguchi, Ito, Ayabe, & Kojima, 2009) (Lebon, Lotze, Stinear, & Byblow, 2012). Using the BE we searched, for HS and P, a subset of subject-specific and session-specific BEC optimizing the similitude index *s*. Table 1 lists the BECs selected for the classification process for each HS, each P and each session.

### 3.2 Classification Performance

The mean of the classification accuracy was 83.5% (SD 12%) for HS and 90% (SD 14.1%) for P. The random level of classification is 82.6%. Table 1 shows the results for each subject and each session. The table shows a predominance of the lower frequency band for P and a predominance of the higher frequency bands for HS (De Lange, Jensen, Bauer, & Toni, 2008). The preliminary selection of the ten electrodes in HS, proved suitable also for patients, thanks to the further selection of subject-specific subsets, and guaranteed a good accuracy in the classification of their answers. The proposed procedure allowed us to fix a robust common subset for all subjects (BE), but we also considered the inter and intra-subject variability by selecting a subject and session specific subset. In a

future practical application of our protocol, each communication session will be preceded by a brief configuration session in which the classification algorithm selects the optimum electrode subset from the fixed BE. Furthermore it will be necessary to train the classifier with two known questions. Even if the procedure is long and repetitive, it guarantees an high classification accuracy for the patients.

**Table 1: The table shows, for each HS, each P and each session (S), the best couples electrode-band (BEC) selected using the similitude index  $s$  and the related classification accuracy (CA).**

	S	BEC				CA	Mean±SD
		$\theta$	$\alpha$	$\beta$	$\gamma$		
HS 1	1	O1	Cz-PO4	C3-Cz-PO4		90 %	<b>83.5±12%</b>
	2			P3-Fc1-T6	PO3-O1-T6	90 %	
HS 2	1	C3- T6-PO3	O1-PO3-P3		O1-PO4	80 %	
	2	P3-PO4		C3	Fz-T6-Fc1-P3-O1-PO3	80 %	
HS 3	1				Fz-C3	100 %	
	2	Fc1-Fc2	PO4		C3	75 %	
HS 4	1	Fc2		C3	Fc2-PO3	80 %	
	2	Cz			O1-PO3	60 %	
HS 5	1		C3	Fz		100 %	
	2	Fc2-Fz	O1-Cz-PO3-P3-PO4		O1	80%	
P 1	1	PO3-C3	Fc1-P3	Fc2	PO4	80%	<b>90±14.1%</b>
P 2	1	PO3-C3	C3-P3			100%	

## References

- Cruse, D., Chennu, S., Fernandez-Espejo, D., Payne, W. L., Young, G. B., & et al. (2010). Detecting Awareness in the Vegetative State: Electroencephalographic Evidence for Attempt Movements to Command. *PLoS ONE*, e49933.
- De Lange, F., Jensen, O., Bauer, M., & Toni, I. (2008). Interaction between posterior gamma and frontal alpha/beta oscillation during imagined actions. *Frontiers in Human Neuroscience*, 10.3389.
- Georgiopoulou, M., Katsakiori, P., Kefalopoul, Z., Elleul, J., Chroni, E., & al., e. (2010). Vegetative State and Minimally Conscious State: A Review of the therapeutic Interventions. *Stereotactic and Functional Neurosurgery*(88), 199-207.
- Ishizu, T., Noguchi, A., Ito, Y., Ayabe, T., & Kojima, S. (2009). Motor activity and imagery modulate the body-selective region in the occipital-temporal area: A near-infrared spectroscopy study. *Neuroscience Letter*, 85-89.
- John, E. R., Halper, J. P., Lowe, R. S., Merkin, H., Defina, P., & et al. (2011). Source imaging of QEEG as a method to detect awareness in a person in vegetative state. *Brain Injury*, 426-432.
- Lebon, F., Lotze, M., Stinear, C. M., & Byblow, W. D. (2012). Task-Dependent Interaction between Parietal and Contralateral Primary Motor Cortex during Explicit versus Implicit Motor Imagery. *PLoS ONE*, e37850.
- Monti, M. M., Vanhaudenhuyse, A., Coleman, M. R., Boly, M., Pickard, J. D., & et al. (2010). Willful modulation of brain activity in disorder of consciousness. *New England Journal of Medicine*, 579-589.

# Commanding a Robotic Wheelchair using High- or Low-Frequency SSVEP-BCI: A Comparative Study

Sandra M. T. Müller<sup>1</sup>, Pablo F. Diez<sup>2</sup>, Teodiano Freire Bastos-Filho<sup>1</sup>, Mário Sarcinelli-Filho<sup>1</sup>, Vicente Mut<sup>2</sup>, Eric Laciari<sup>2</sup> and Enrique Avila<sup>2</sup>

<sup>1</sup> Federal University of Espirito Santo, Vitória, Brazil

sandramuller@ceunes.ufes.br, teodiano@ele.ufes.br, mario.sarcinelli@ufes.br

<sup>2</sup> Universidad Nacional de San Juan, San Juan, Argentina

pdiez@gateme.unsj.edu.ar, laciari@gateme.unsj.edu.ar, vmut@inaut.unsj.edu.ar, eavila@inaut.unsj.edu.ar

## Abstract

This work presents a comparative study between low and high frequencies of visual stimulation used in a Brain-Computer Interface (BCI) based on Steady State Visual Evoked Potentials (SSVEP). This comparison has the goal of evaluating the visual tiredness produced by flickering visual stimuli in two distinct frequency ranges (low and high frequency). For this purpose, five volunteers with disabilities operated a wheelchair through a SSVEP-based BCI. In the experiments, each subject answered a questionnaire about performance and tiredness associated to the use of the BCI. Average ITR obtained for low- and high-frequency stimuli were 20.3 bits/min and 15.0 bits/min, respectively. Despite of a lower average ITR, it was found that high-frequency stimuli were more comfortable and could lead to a better performance in the accomplishment of the navigation tasks.

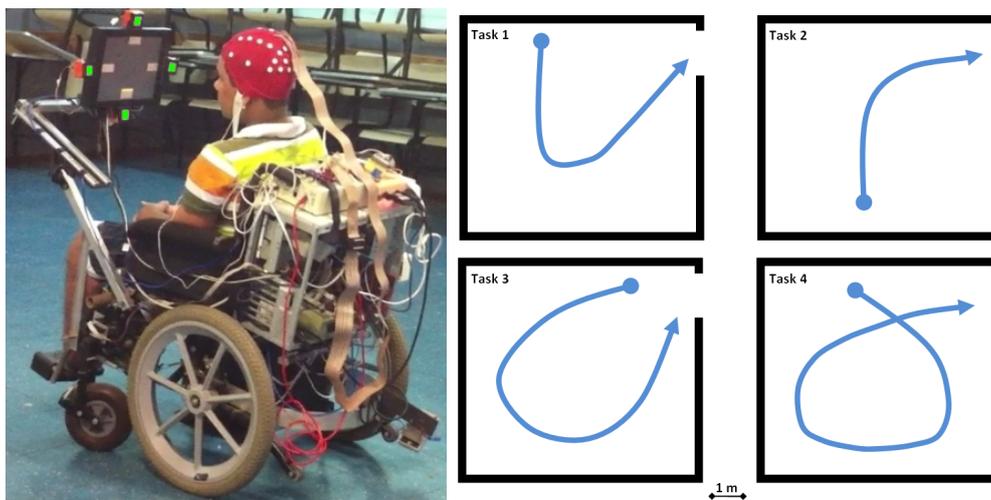
## 1 Introduction

Evoked potentials in electroencephalography (EEG) elicited by a train of stimuli are called Steady-State Visual Evoked Potentials (SSVEP). SSVEP can be used as a paradigm in BCI development [5]. Generally, SSVEP are stronger in low-frequency range (6 to 12 Hz) than in high-frequency range (more than 30 Hz) [6]. Therefore, the majority of SSVEP-BCI are based on low- and medium-frequency ranges [2]. However, high-frequency stimulation is less annoying and consequently produces a pronounced decrease of visual tiredness caused by flickering [6]. This statement is accepted in BCI bibliography, but there are insufficient studies focusing on it. This work presents a comparative study between two common stimulation systems used in SSVEP-based BCIs, trying to evaluate the visual tiredness of the user. More research has been conducted using computer screens than other stimulation source [7] and checkerboards is one of the basic choices. Hence, low-frequency stimulation was presented as checkerboards on a computer screen. However, high-frequency stimuli cannot be rendered on the screen, and because of this LEDs were used.

## 2 Materials

A BrainNet BNT-36 acquisition system was used to acquire EEG signals. Twelve EEG channels with the reference electrode at the left ear lobe and filtered between 0.1 and 100 Hz were digitized at 600 Hz. Using the extended international 10-20 system, the locations for the electrodes were P<sub>7</sub>, PO<sub>7</sub>, P<sub>5</sub>, PO<sub>3</sub>, PO<sub>Z</sub>, PO<sub>4</sub>, P<sub>6</sub>, PO<sub>8</sub>, P<sub>8</sub>, O<sub>1</sub>, O<sub>2</sub>, and O<sub>Z</sub>.

Two FPGA are used to produce stimuli in low- and high-frequencies. The stimuli in low-frequency range are shown in a 12" LCD display at a distance of 0.5 m from the user. They are composed of four black/white checkerboard stripes, flickering on the screen at 5.6 Hz (top), 6.4 Hz (right), 6.9 Hz (bottom) and 8.0 Hz (left), as illustrated in Figure 1 (a). On the other hand, stimuli in high frequency range are illuminated by high efficiency green LEDs flickering at 37 Hz (top), 38 Hz (right), 39 Hz (bottom) and 40 Hz (left) on the LCD sides, also shown in Figure 1 (a). Each checkerboard stripe or led corresponded to a movement of the wheelchair: forward, left, right and stop.



(a) Volunteer onboard the wheelchair under a low and high frequency stimulation, used at different times.

(b) The four navigation tasks.

Figure 1: Volunteer onboard a wheelchair and the four navigation tasks.

Five volunteers with different disabilities (quadriplegia, paraplegia and Duchenne dystrophy) participated on the experiments. They were informed about the experimental procedure and they (or their relatives) provided written consent to participate on. The experiments were carried out according to the rules of the Ethics Committee of UFES/Brazil (reg. number CEP-048/08).

### 3 Methods

Before operating the wheelchair, each volunteer performed a training session with the BCI. They were asked to follow the verbal cues to gaze at a stripe for 30 seconds. Visual feedback denoting the detected stripe was presented to the user. Then, the volunteers could operate the wheelchair. The four tasks are illustrated in Figure 1 (b). The room dimensions are 8.75m long by 7.07m wide. The goal of all tasks was to reach the area next to the door.

Finally, the volunteers answered a questionnaire with questions related to tiredness and comfort when using the BCI. This questionnaire allows evaluating the influence of tiredness and concentration on user performances qualitatively. The questions were:

(A) Are you tired?

(B) Did the screen oscillations interfere with your concentration? (This question is related to screen oscillations due to wheelchair movements, particularly when it begins or ends a movement).

(C) Was the stimuli colour annoying?

These questions should be answered according to the ranking: 1 - None; 2 - A little; 3 - Medium; 4 - Quite.

The EEG signal processing method is fully described in [3]. Basically, the EEG is filtered and then, the Power Spectral Density (PSD) was determined. Later, a Spectral F-Test (SFT) is applied in the feature extraction step [1]. A classifier based on a decision tree was implemented with attributes that maximize the discrimination among classes. The training step is unnecessary for this classifier because its operation is straightforward. Moreover, baseline or reference signal are unnecessary and supervisor intervention is not required. Thus, since the user sits on the wheelchair and wears the EEG cap, he will be ready to use the BCI. This BCI worked asynchronously and detections were performed at each second, accordingly a command is sent to the wheelchair every second. On one hand, the BCI considered the first three harmonics in SSVEP detection for low frequencies. On the other hand, only the first harmonic was considered for high frequencies.”

## 4 Results

Table presents the hit rate and the ITR obtained just before the wheelchair operation for low- and high-frequency stimuli. Table 1 presents the average detection accuracy (Acc) among the four classes and its respective average ITR, calculated according to [5]. The number of navigation tasks completed by each volunteer is presented as well.

Table 1: Results for low- and high-frequency stimulation.

Vol	Low Frequency			High Frequency		
	Acc $\pm$ SD	ITR	Completed Tasks	Acc	ITR	Completed Tasks
1	46% $\pm$ 6.05	8.9	4	60% $\pm$ 13.12	23.7	4
2	44% $\pm$ 7.5	7.4	1	40% $\pm$ 10.24	4.7	3
3	78% $\pm$ 5.35	53.5	2	63% $\pm$ 15.04	27.8	4
4	62% $\pm$ 9.88	26.4	2	51% $\pm$ 13.12	13.4	3
5	41% $\pm$ 11.81	5.3	1	41% $\pm$ 9.6	5.3	2
<b>Average</b>	<b>54% <math>\pm</math> 8.12</b>	<b>20.3</b>	<b>2</b>	<b>51% <math>\pm</math> 12.22</b>	<b>15.0</b>	<b>3</b>

Classification results were evaluated using a non-parametric statistical analysis, according to sample size [4]. Hence, the Wilcoxon signed paired test was used. Differences in the hit-rates obtained from low- and high-frequency stimulation range were not statistically significant ( $p=0.465$ ). At the end of the experiments, the volunteers answered the questionnaire, whose results are presented in Table 2. The median values for Question A was 'Medium' for low-frequency stimulation and 'A little' for high-frequency stimulation. This difference in tiredness was statistically significant according to Wilcoxon signed paired test ( $p=0.025$ ). On the other hand, differences for the questions related to display movement and colour tiredness were not statistically significant ( $p=0.317$  and  $p=0.461$ , respectively).

## 5 Discussions and Conclusions

Stimulation parameters are a very important issue for a SSVEP-based BCI implementation and can affect the system performance and the user comfort and safety. A statement issued in

Table 2: Questionnaire applied to the volunteers after the robotic wheelchair operation.

Question	Frequency Range	Vol1	Vol2	Vol3	Vol4	Vol5	Median
A	Low	Medium	Medium	A little	Medium	A little	Medium
	High	A little	A little	None	A little	None	A little
B	Low	None	Medium	Medium	None	A little	A little
	High	None	Medium	None	None	A little	None
C	Low	None	Quite	None	Quite	None	None
	High	A little	A little	A little	None	None	A little

BCI bibliography is that high-frequency stimulation produces less visual tiredness than lower frequency stimulation [2], [6]. However, is high-frequency stimulation less annoying than lower frequency stimulation? The current research tries to answer partially this question by evaluating the two commonly used stimulation system in SSVEP-based BCI, checkerboards in low-frequency and LEDs in high-frequency. Note that the different stimulus patterns could affect the SSVEP signals.

In low frequency stimulation, only volunteer Vol1 could execute the four navigation tasks. Volunteers reported less tiredness for high-frequency LED stimulation than for low-frequency checkerboard stimulation (except Vol1), which was statistically significant ( $p=0.025$ ). However, the other analyzed variables such as colour and display movement, do not seem to affect the volunteer performance, since those differences were non-significant ( $p>0.05$ ).

Although average ITR values in low-frequency (20.3 bits/min) and in high-frequency (15 bits/min) were different, the volunteers could perform more tasks with high-frequency stimuli. Then, less tiring stimuli are visually more comfortable and could lead to a better performance.

Although the SSVEP are stronger in low-frequency range (and consequently their detection is easier), the developed BCI could detect the visual evoked potentials in high-frequency range with good performance. Results show that the LED stimuli in high-frequency range produce lower visual tiredness on the users, compared with low-frequency checkerboard. Note that the small sample size limits the conclusions exposed on this paper.

## References

- [1] Antonio Mauricio F. L. Miranda de Sá, Humberto C. Thiengo, Ingrid S. Antunes, and David M. Simpson. Assessing time- and phase-locked changes in the eeg during sensory stimulation by means of spectral techniques. *Proc. IFMBE 2009*, 25:2136–2139, 2009.
- [2] Pablo F. Diez, Vicente Mut, Eric Laciár, and Enrique Avila. Asynchronous bci control using high-frequency ssvep. *Journal of Neuroengineering and Rehabilitation*, 8(39), 2011.
- [3] Sandra Mara Torres Müller, Antonio Mauricio F. L. Miranda de Sá, Teodiano Freire Bastos-Filho, and Mário Sarcinelli-Filho. Spectral techniques for incremental SSVEP analysis applied to a BCI implementation. *Proceedings of V Latin American Congress on Biomedical Engineering, CLAIB 2011, IFMBE Proceedings*, 33:4pp, 2011.
- [4] R. L. Scheaffer and J. T. McClave. *Probability and Statistics for Engineers*. Editorial Iberoamericana, Mexico, 1993.
- [5] Francois-Benoit Vialatte, Monique Maurice, Justin Dauwels, and Andrzej Cichocki. Steady-state visually evoked potentials: Focus on essential paradigms and future perspectives. *Progress in Neurobiology*, 90:418–438, 2010.
- [6] Y. Wang, R. Wang, Xiaorong Gao, B. Hong, and Shangkai Gao. A practical VEP based brain computer interface. *IEEE Trans. on Neural Syst. and Rehab. Eng.*, 14(2):234–239, 2006.
- [7] Danhua Zhu, Jordi Bieger, Gary Garcia Molina, and Ronald M. Aarts. A survey of stimulation methods used in SSVEP-based BCIs. *Computational Intelligence and Neuroscience*, 702357, 2010.

# Brain-computer music interfacing for continuous control of musical tempo

Ian Daly<sup>1</sup>, Duncan Williams<sup>2</sup>, Faustina Hwang<sup>1</sup>, Alexis Kirke<sup>2</sup>, Asad Malik<sup>1</sup>, Etienne Roesch<sup>1</sup>, James Weaver<sup>1</sup>, Eduardo Miranda<sup>2</sup> and Slawomir J. Nasuto<sup>1</sup>

<sup>1</sup> Brain Embodiment Lab, School of Systems Engineering, University of Reading, RG6 6AY, UK  
[i.daly@reading.ac.uk](mailto:i.daly@reading.ac.uk)

<sup>2</sup> Interdisciplinary Centre for Computer Music Research, University of Plymouth, PL4 8AA, UK

## Abstract

A Brain-computer music interface (BCMI) is developed to allow for continuous modification of the tempo of dynamically generated music. Six out of seven participants are able to control the BCMI at significant accuracies and their performance is observed to increase over time.

**Keywords** Brain-computer music interface, Music generation, EEG, Online continuous control

## 1 Introduction

Brain-computer music interfaces (BCMIs) are devices that allow control over, or interaction with, music via brain activity and without activation of the efferent nervous system [6]. BCMIs therefore allow individuals to interact with, create, or modify music in situations where this would not otherwise be possible. For example, individuals with severe movement restrictions may benefit from use of a BCMI for recreation or as a therapeutic device.

Previous studies have presented BCMIs for passive modification of musical properties, such as beat and loudness, via measurements of complexity of different frequency bands in the electroencephalogram (EEG) [5]. They have also demonstrated selection from a discrete set of two-tone bass frequency drone sounds via emotional imagery [3] and selection from a discrete set of musical scores via steady-state visual evoked potentials (SSVEPs) [4].

Thus, BCMIs have been shown to allow active control of music generation via selection from a discrete set of instructions or passive interaction with music. However, there is an opportunity to further investigate whether BCMI users can wilfully modify musical properties on a continuous scale during a period of online music control.

Therefore, a BCMI is constructed to allow users to modulate the tempo of a piece of music dynamically via intentional control. Specifically, users are able to increase the tempo of the music via kinaesthetic motor imagery (MI) [7] and decrease the tempo via relaxation. Music tempo is mapped to the strength of the users motor imagery allowing them to move the tempo continuously across a specified range.

## 2 Methods

### 2.1 Participants

Seven healthy right-handed individuals (median age=23; SD=2.9; 6 males) voluntarily participated in the experiment. All participants gave informed consent, and the study was approved as per the University of Reading guidelines for ethics.

## 2.2 Measurements

EEG was recorded from 19 channels positioned via the international 10/20 system, referenced to a central electrode at FCz, and a ground electrode at AFz.

EEG was sampled at a rate of 1,000 Hz via a Brain Products BrainAmp EEG amplifier (Brain Products, Germany). Impedances were kept below 5 k $\Omega$  for all participants.

## 2.3 Music generation

To sustain novelty and avoid complications caused by user familiarity, a generative algorithm was implemented to produce monophonic piano music, with real-time control over tempo.

The generative algorithm allows the user to specify three parameters from which it creates sequences of tone rows (string of pitch classes with no repeated notes) and a pool of rhythm data. In our system the generated row is used to supply the selection of notes from which the musical sequences are derived.

Our BCMI uses 6 notes in a tone row. The rhythm pool is generated according to a list of default values (eight quavers per sequence), with a number of variations introduced according to the starting parameters. The rhythm pool is used to select notes from the generated tone rows randomly, creating a large variation of possible musical sequences (musical streams created by combining notes from the tone row with rhythms from the pool) from a small amount of seed data. The starting note was randomly selected for each participant, and tempo was scaled between the range of 99 beats per minute (bpm) to 288 bpm.

## 2.4 Brain-computer interface

Musical tempo was mapped to the mean alpha (8-13 Hz) band-power recorded over electrodes F3, T3, C3, Cz, and P3 centred over the left motor cortex. Participants were instructed to increase the tempo of the music via kinaesthetically imagining squeezing a ball in their right hand or decrease the tempo by relaxing. Mean alpha band-power was inverted, scaled by a constant  $k$ , and mapped to music tempo. Thus, relaxing increases alpha and decreases tempo.

A single trial of the BCMI control paradigm first presented a fixation cross from -4 s to 0 s. Music playing began at a fixed tempo of 148 bpm with the appearance of the cross. An arrow cue (either up; increase tempo, or down; decrease tempo) was presented in the centre of the screen (visual angle  $\approx 7.5^\circ$ ) at second 0. The cue remained on screen for 12 seconds, while the tempo was mapped to the alpha band power recorded over the left motor cortex and updated every 100 ms. Upon disappearance of the cue the music generation ceased and a visual feedback stimuli was presented for 500 ms. This took the form of either a happy face or a sad face.

The paradigm was split into 9 runs. The first was a calibration run to train the parameter  $k$ . This run contained 30 trials in pairs of increase and decrease tempo trials. After each pair  $k$  was either increased by  $\alpha$  ( $\alpha = 10$  at the start of the run) if tempo was reduced in both trials, decreased if tempo increased in both trials, or held constant. If  $k$  was held  $\alpha$  was halved.

For all subsequent runs the term  $k$  was held constant at the value arrived at after the calibration run. Each run contained 18 trials and participants were given breaks between runs.

The BCMI was constructed using open standards developed in the TOBI framework [1].

## 2.5 Analysis

EEG was visually inspected for artefacts via a scorer blinded to the contents of the trials and results. Trials were rejected if they contained artefacts on channels F3, T3, C3, or P3.

Participant	Session								Avg.
	1	2	3	4	5	6	7	8	
1	0.444	0.450	0.444	0.500	0.556	0.500	0.550	0.500	0.50 (0.04)
2	0.357	0.527	0.518	0.300	0.548	0.341	<b>0.732</b>	0.571	0.52 (0.14)
3	<b>0.708</b>	<b>0.764</b>	<b>0.750</b>	<b>0.687</b>	0.625	0.556	0.500	0.375	<b>0.66 (0.13)</b>
4	<b>0.800</b>	<b>0.800</b>	<b>0.875</b>	<b>0.800</b>	<b>0.889</b>	<b>0.687</b>	<b>0.700</b>	<b>0.889</b>	<b>0.80 (0.07)</b>
5	0.500	0.500	0.500	0.512	<b>0.667</b>	<b>0.722</b>	0.556	0.587	0.53 (0.08)
6	<b>0.917</b>	<b>0.812</b>	<b>0.722</b>	0.632	0.611	<b>0.687</b>	-	-	<b>0.71 (0.12)</b>
7	0.300	0.300	0.312	0.625	<b>0.833</b>	<b>0.750</b>	0.417	0.333	0.38 (0.22)
Avg.	0.50	0.53	0.52	0.63	0.63	0.69	0.55	0.54	-
	(0.23)	(0.19)	(0.19)	(0.13)	(0.13)	(0.14)	(0.12)	(0.19)	

Table 1: Balanced accuracies over sessions for each participant. Statistically significant accuracies ( $p < 0.05$ ) are in bold. Bonferroni correction is applied on a per participant basis.

Event related (de)synchronisation (ERD/S) was identified in the EEG recorded over the left motor cortex by extracting band-power measures of alpha (8-13 Hz) frequency band activity relative to a baseline period, defined as the fixation cross presentation period (-4 to 0 s).

Balanced accuracy was used to evaluate each user’s ability to control the BCMI [2].

### 3 Results

A total of 154 trials were excluded due to artefact contamination (the presence of blinks, movement, and other sources of noise). Additionally, one participant (participant 6) did not complete the last two runs of the experiment. Thus, a total of 818 (84.16 %) of the trials were artefact free and may be used to evaluate user performance.

Table 1 reports accuracies achieved by each participant over the sessions. Figure 1 illustrates the median and standard deviation of accuracies over sessions achieved by participants.

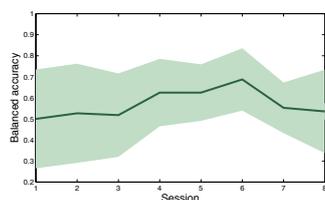


Figure 1: Balanced accuracy over sessions. The solid line indicates median accuracy and the shaded area  $\pm 1$  standard deviation.

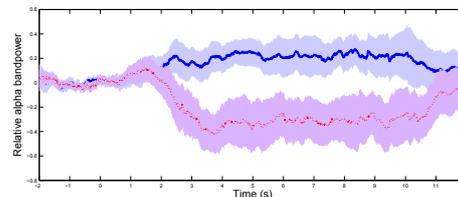


Figure 2: Relative mean alpha bandpower on C3, participant 4, all sessions. Dotted line, bandpower during the increase condition.

Figure 2 illustrates the median change in alpha band-power over the left motor cortex (channel C3) relative to the pre-cue baseline period. Note, for the reduce tempo (motor imagery) condition the user was able to produce a sustained event-related desynchronisation (ERD).

## 4 Discussion

Music can be actively controlled via 6 out of our 7 BCMI users with significant accuracy in one or more sessions. It may be argued that multiple comparison correction is needed. However, sessions are not independent, thus Bonferroni correction on a per session basis is not appropriate. Instead, Bonferroni correction is applied to mean accuracies per participant, revealing 3 participants to be able to control the BCMI at significant accuracies.

There is considerable variability and for some sessions participants were not able to control the BCMI at significant accuracies. Nonetheless, our results demonstrate that music tempo can be understood and utilised as a feedback mechanism by the majority of BCMI users.

These results open up the possibility of allowing users a greater level of control of music than previously explored [4, 3]. The performance increase over sessions visible in figure 1 is also encouraging as it suggests that users can, through increased experience, improve their ability to control tempo. We also note, however, that in the final sessions our users displayed a performance reduction. This may be due to fatigue, but further exploration is required.

Future work will seek to compare users' ability to learn to control properties of a piece of music with feedback modalities more traditionally used in BCI control (for example, visual feedback). We will also explore the effects of more advanced computational methods on the performance of the BCI. Additionally, the interaction of music and emotion is an interesting research angle and it would be very interesting to extend the work presented in [3] to allow continuous emotional imagery based control of musical features such as tempo, mode, or timbre. Therefore, future work will also look at extending our BCMI to continuously detect emotions.

## Acknowledgments

This work was supported by the EPSRC grants (EP/J003077/1 and EP/J002135/1).

## References

- [1] C Breitwieser, I Daly, C Neuper, and G Muller-Putz. Proposing a Standardized Protocol for Raw Biosignal Transmission. *IEEE trans. on bio-med. eng.*, 59(3):852–859, December 2012.
- [2] Kay Henning Brodersen, Cheng Soon Ong, Klaas Enno Stephan, and Joachim M. Buhmann. The Balanced Accuracy and Its Posterior Distribution. In *2010 20th Int. Conf. on Patt. Recon.*, pages 3121–3124. IEEE, August 2010.
- [3] S. Makeig, G. Leslie, T. Mullen, D. Sarma, N. Bigdely-Shamlo, and C. Kothe. First Demonstration of a Musical Emotion BCI. *Affecti. Comput. and Int. Interact. Lect. Notes in Comp. Sci.*, 6975:487–496, 2011.
- [4] E. R. Miranda, W. L. Magee, J. J. Wilson, J. Eaton, and R. Palaniappan. Brain-Computer Music Interfacing (BCMI): From Basic Research to the Real World of Special Needs. *Music and Medicine*, 3(3):134–140, March 2011.
- [5] Eduardo Reck Miranda. Brain-Computer music interface for composition and performance. *IntJ Dis Human Dev*, 5(2), 2006.
- [6] E.R. Miranda, S. Roberts, and M. Stokes. On Generating EEG for Controlling Musical Systems. *Biomedizinische Technik*, 49(1):75–76, 2004.
- [7] Christa Neuper, Reinhold Scherer, Miriam Reiner, and Gert Pfurtscheller. Imagery of motor actions: differential effects of kinesthetic and visual-motor mode of imagery in single-trial EEG. *Brain research. Cognitive brain research*, 25(3):668–77, 2005.

# Specific effects of slow cortical potentials neurofeedback training on attentional performance

Ruben GL Real, Andrea Kübler and Sonja C Kleih

Institute of Psychology  
University of Würzburg, Germany  
ruben.real@uni-wuerzburg.de

## Abstract

The association between successful regulation of slow cortical potentials and latencies in two reaction time tasks was investigated. At the end of a three-day training period successful regulators showed a stronger reduction in latencies than unsuccessful regulators. This was true for a task involving learning, but not if the task only required responding as fast as possible. Our results suggest that healthy participants can learn to regulate SCPs during the course of a short training, and that training success is associated with reduced response latencies. Possible applications of SCP-based biofeedback for rehabilitation after acquired brain injuries are discussed.

## 1 Introduction

Stroke is one of the leading causes of death and disability in western societies [1]. Cognitive impairments, such as attention deficits play a prominent role [2] and EEG biofeedback (or neurofeedback, NF) has been proposed as an effective treatment [3]. Negative slow cortical potentials (SCP) reflect summation of synchronized excitatory postsynaptic potentials from apical dendrites, whereas positive SCPs derive from reduced inflow to apical dendrites or inhibitory activity. In line with this view, studies show that an increase of negative SCP amplitudes is related to better task performance when attention and motor reactions are required [4, 5]. However, the knowledge about the relation between attention and slow cortical potentials [6] has not yet been exploited for improving attentional functioning in stroke [2].

In this study we sought to relate the progress in SCP control during NF training to latencies (as a measure of attention) in two reaction time (RT) tasks. In the *simple task* subjects had to monitor a series of 31 random letters for 8 occurrences of the letter "X" and respond to "X" as fast as possible via a button-click. In the *learning task* subjects had to monitor a series of 31 random letters for 6 occurrences of the two-letter sequence "A"—"X" and respond to "X" as fast as possible via a button-click. Optimal performance in this task required subjects to learn the contingency of "A" being always followed by "X". ITI was 690 ms, and both tasks were presented in blocks of nine sequences each.

Across training, we hypothesized a learning effect for SCP regulation, i.e. increasing separation of negativity and positivity trials, and for the learning tasks, i.e. reduced latencies. Our main interests were in the association of regulatory success and response latencies, hypothesizing that successful SCP regulation would correlate with reduced latencies in the learning but not in the simple task.

## 2 Methods

Each NF session included 250 trials in which cortical positivity had to be increased and 250 trials in which cortical negativity had to be increased. Trials lasted for 8 s (baseline: 0-2 s, active phase: 2-8 s). Online feedback, from electrode Cz referenced to linked mastoids, consisted of a circle whose size and color indicated whether subjects regulated successfully with regard to baseline activity. Trials were judged successful, and success indicated to the participant, if brain activation was regulated as required by the task (towards positivity or negativity, respectively). Vertical eye movements were corrected for using a regression procedure. After getting acquainted with the RT tasks, eleven healthy student subjects (mean age: 23.20, SD: 5.20, range 20-38) trained for three days during one week. On days one and three, training consisted of two sessions (A & B), and each session comprised one neurofeedback training and the two RT-tasks. On day two, the single training session consisted of the sequence RT-tasks–neurofeedback training–RT-tasks. Here, we evaluate associations between latencies and amplitudes in the EEG at days one and three.

Preprocessing steps for the offline analysis corresponded to the steps used during online training (10 Hz low-pass filtering, detrending to compensate for DC drift, epoching, alignment to pre-trial baseline, correction for ocular artifacts, exclusion of trials with absolute voltages exceeding 100  $\mu V$ ). Then, trials were averaged per regulatory condition, and the mean activity during the 2 s long interval starting from 4 s after the beginning of each trial exported for statistical analysis. One participant was excluded due to low quality of the EEG recording, and another participant did not participate in all training sessions and was therefore excluded. The total sample size available for analyzing the association between amplitude and latencies is, thus,  $n = 9$ .

To analyze the relation between latencies in the RT tasks and success during SCP regulation, we followed a two-step approach. First, we established that latencies in the RT tasks decreased, and that SCP regulatory success, i.e. the absolute difference between negativity and positivity trials, increased across training. Then, we analyzed the association between the change in latencies and the change in SCP regulatory success between day three and day one. Reported  $p$ -values are two-sided.

## 3 Results

Repeated measures ANOVA of latencies in the learning task revealed significant effects of training day ( $F_{1,9} = 3.51, p = .047$ ), session ( $F_{1,9} = 10.37, p = .011$ ), and their interaction ( $F_{1,9} = 7.01, p = .027$ ). Follow-up analysis indicated that latencies decreased across sessions at day one but remained stable thereafter (see Figure 1 A, Table 1). Latencies in the simple task did not change across training (all  $p > .30$ ). Mean number of errors was low (range 1 – 4%) and did not differ between conditions (all  $p > .16$ ).

Results for SCP regulation (see Figure 1 B) showed significant effects of regulation ( $F_{1,9} = 16.63, p = .003$ ), the interaction of regulation and training day ( $F_{1,9} = 11.63, p = .008$ ) and a marginal effect of the interaction of regulation and session within each day ( $F_{1,9} = 4.68, p = .059$ ; see Table 1).

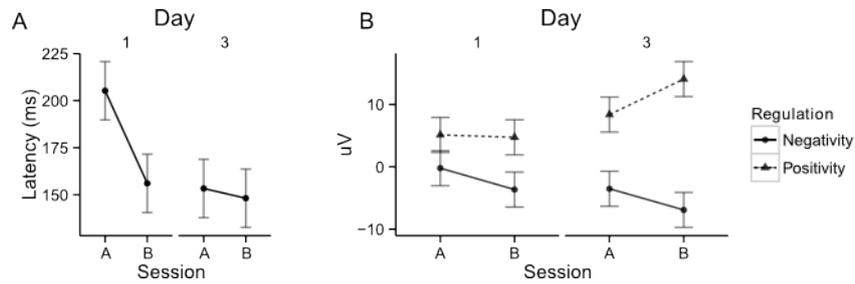


Figure 1: Results of RT-task and SCP training

A: Effects of training on latencies in the learning RT-task at Day 1 and Day 3. B: Results of SCP training at Day 1 and Day 3. Error bars show Fisher's least significant differences.

Day	Session	SCP ( $\mu V$ )		learning task (ms)		simple task (ms)	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
1	a	7.94	8.61	205.23	66.24	435.46	30.23
1	b	8.10	7.95	156.05	55.76	430.46	34.08
3	a	8.85	8.96	153.33	76.56	444.73	38.05
3	b	21.59	13.77	148.14	87.48	434.22	29.11

Table 1: SCP regulatory success and latencies across training

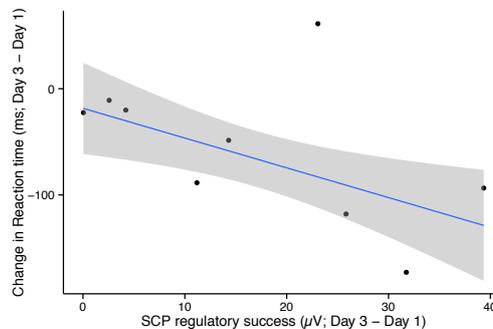


Figure 2: Association between SCP regulatory success and response latencies

Analysis of the association between learning in the learning task and learning of SCP regulation between day three and day one indicated a large –but marginal– negative association (Spearman's  $\rho = -.63, p = .076$ ). Figure 2 shows that subjects who were able to increase the difference in SCP amplitudes between days three and one, tended to show faster response latencies on day three in comparison to day one. In the simple task, changes in latencies between day three and day one were not associated with changes in SCP regulatory success ( $\rho = .017, p = .982$ ).

## 4 Discussion

To summarize, our data show that healthy participants can learn to regulate their SCPs during the course of a short training. Latencies in the learning task decreased across training, but latencies in the simple task remained stable. Further, learning of SCP regulation was marginally associated with reduced latencies in the learning, but not in the simple task. However, because we employed a correlative design it is unknown whether changes in SCP regulatory success play a causal role for latencies in the learning task.

Our results support previous findings that successful SCP regulation is associated with faster reaction times [5]. However, this association was found only for a task which involved learning, but not in a simple RT task. This may suggest that the choice of tasks used in neurofeedback studies may play a crucial role. We thus speculate that associations between SCP regulatory success and behavioral data should be strongest with tasks of medium difficulty involving executive functions [7].

Our results show that successful regulation of SCPs is possible even with limited training and that this change is associated with increased performance in a task requiring attention and motor reaction. Based on these encouraging results, studies are underway aiming to replicate these findings in a large sample of chronic stroke patients.

## Acknowledgments

This work is supported by the European ICT Programm (Project FP7-287320, contrast-project.eu). We thank Anna-Sophia Fritsch for her support during data collection.

## References

- [1] P. Heuschmann, O. Busse, M. Wagner, M. Endres, A. Villringer, J. Röther, P. Kolominsky-Rabas, and K. Berger. Schlaganfallhäufigkeit und versorgung von schlaganfallpatienten in deutschland. *Aktuelle Neurologie*, 37(07):333–340, 2010.
- [2] J. A. Michel and C. A. Mateer. Attention rehabilitation following stroke and traumatic brain injury. a review. *Europa medicophysica*, 42(1):59–67, 2006.
- [3] Lonnie Nelson. The role of biofeedback in stroke rehabilitation: Past and future directions. *Topics in Stroke Rehabilitation*, 14(4):59–66, 2007.
- [4] Werner Lutzenberger, Thomas Elbert, Brigitte Rockstroh, and Niels Birbaumer. The effects of self-regulation of slow cortical potentials on performance in a signal detection task. *International Journal of Neuroscience*, 9(3):175–183, 1979.
- [5] B. Rockstroh, T. Elbert, W. Lutzenberger, and N. Birbaumer. The effects of slow cortical potentials on response speed. *Psychophysiology*, 19(2):211–217, 1982.
- [6] N. Birbaumer, T. Elbert, A. G. Canavan, and B. Rockstroh. Slow potentials of the cerebral cortex and behavior. *Physiological reviews*, 70(1):1–41, 1990.
- [7] Aljoscha C. Neubauer and Andreas Fink. Intelligence and neural efficiency. *Neuroscience & Biobehavioral Reviews*, 33(7):1004–1023, 2009.

# Classification of error-related potentials in EEG during continuous feedback

Martin Spüler<sup>1</sup>, Christian Niethammer<sup>1</sup>, Wolfgang Rosenstiel<sup>1</sup> and Martin Bogdan<sup>1,2</sup>

<sup>1</sup> Wilhelm-Schickard-Institute for Computer Science, University of Tübingen, Tübingen, Germany

<sup>2</sup> Computer Engineering Department, University of Leipzig, Leipzig, Germany

## Abstract

When a Brain-Computer Interface (BCI) delivers erroneous feedback, an error-related potential (ErrP) can be measured as response of the user recognizing that error. Classification of ErrPs has been previously used in BCIs with time-discrete feedback to correct errors or to improve adaptation of the classifier for more robust BCI feedback. In this study we investigated if ErrPs can be measured in electroencephalography (EEG) recordings during continuous feedback and if ErrPs can be classified. We recorded EEG data from 10 subjects during a video game task and investigated two different types of error (execution error, due to inaccurate feedback; outcome error, due to not achieving the goal of an action). ErrPs could be measured in the EEG for both types of errors and we were able to classify both types of errors using a Support Vector Machine (SVM).

## 1 Introduction

If a subject makes or perceives an error, an error-related potential (ErrP) can be detected in the EEG due to the subject recognizing the error [1]. That an ErrP can also be detected when a Brain-Computer Interface (BCI) delivers erroneous feedback has been shown in several publications and it has further been shown that the detection of ErrPs can be utilized to correct errors [2, 3] or improve adaptation of the BCI [4, 5]. For the analysis of ErrPs it is necessary to have stimulus-locked data and therefore the previous studies have only investigated time-discrete feedback, in which feedback is given once at the end of a trial.

Kreilinger et al. [6] studied ErrPs during continuous arm movement and tried to classify ErrPs by mapping the continuous feedback to time-discrete feedback and additionally displaying the discrete feedback. That a discretisation of the feedback is not needed was shown by Milekovic et al. [7] in a study using Electrocorticography (ECoG) instead of EEG. They could show that an error-related response during continuous feedback can be observed in the ECoG signal and also classified [8]. For this paper we evaluated if ErrPs can also be measured in the EEG during only continuous feedback and if two different types of errors can be discriminated.

## 2 Methods

### 2.1 Task description

The experimental task was similar to the one described by Milekovic et al. [7] in which the subject had to play a simple video game. The subject used the thumbstick of a gamepad to control the angle in which the cursor on the screen moved. The task was to avoid collisions of the cursor with blocks dropping from the top of the screen with a constant speed. The speed of the falling blocks was set to a level that the game was challenging and the player collided with a

block from time to time. In case of a collision, the game resumed for 1 second and then stopped. The delay of 1 second was introduced to make sure that the reaction measured in the EEG originates from the subject recognizing the collision (outcome error) and not from the game stopping or restarting. To study the execution error, which is happening when the interface delivers erroneous feedback, the angle of the cursor movement was modified for the duration of 2 seconds. The degree of modification was randomized (45 °, 90 °, 180 ° to either the left or the right side). The time between two execution errors was randomized to be between 5 and 8 seconds.

## 2.2 Experimental setup

10 healthy subjects (mean age:  $24.1 \pm 1.1$  years) were recruited for this study. EEG was measured with two g.tec g.USBamp amplifiers and a Brainproducts Acticap System. 29 electrodes were placed on the scalp of the subject to measure EEG, while 3 electrodes were placed below the outer canthi of the eye and above the nasion for electrooculogram (EOG) recordings. The data was recorded with a sampling rate of 512 Hz and a 50 Hz notch filter was applied to filter out power line noise, as well as an additional bandpass filter between 0.5 Hz and 60 Hz. The position of the thumbstick as well as information about outcome or execution errors was transmitted to the recording software using the parallel port of the computer.

## 2.3 Signal processing and classification

The data was segmented into different trials with a length of 1 second: execution errors, time-locked to the start of an angle modification; outcome errors, time-locked to the collision event; and noError trials, where neither a collision nor an angle modification has happened during the trial or in the 1 second before or after the trial. For each subject about 1 hour of EEG was recorded resulting on average in  $597 \pm 22$  execution errors,  $86 \pm 30$  outcome errors and  $475 \pm 39$  noError trials.

An EOG-based regression method was used to reduce the effect of eye artifacts. To estimate classification accuracies we used a 10-fold cross-validation. For classification we used a Support Vector Machine (SVM) with a linear kernel. For all channels the samples in the time range 0.2 s to 0.9 s were used for classification and the signal was rereferenced to the common average. To investigate how well the error can be classified, outcome error and execution error, respectively, were classified against noError trials. To see if the two types of errors can be discriminated, we also classified execution errors against outcome errors. Since the number of trials were different for each class, the dataset was balanced to obtain an even amount of trials for each class.

To test if the subject's movements (due to gamepad control) or eye movements influence classification, classification was also done on the EOG data and on the recorded position of the thumbstick.

## 3 Results

Averaged over all subjects, execution errors and noError trials could be classified with an average accuracy of 65.0 % based on EEG, 50.9 % based on EOG and 52.5 % based on the thumbstick position. For outcome error against noError trials, average accuracies of 73.9 % (EEG), 54.9 % (EOG) and 56.0 % (thumbstick) could be reached. For the classification of the two error types, execution error and outcome error, we obtained average accuracies of 75.3 % (EEG), 56.4 % (EOG) and 55.3 % (thumbstick). Detailed results of the classification on the EEG data is shown

in table 1. For the classification results of the EEG data we performed a permutation test (1200 permutations) to test for significance and all EEG results were found to be significantly above chance level.

Table 1: Classification accuracies based on EEG data obtained by 10-fold cross-validation. Classes were balanced and therefore the chance level is at 50 %. All results are significantly above chance level.

subject	Execution vs. Outcome	Outcome vs. noError	Execution vs. noError
S01	77.9 %	74.4 %	69.7 %
S02	76.3 %	78.5 %	65.4 %
S03	69.6 %	68.2 %	59.9 %
S04	72.1 %	75.0 %	60.1 %
S05	70.2 %	60.3 %	64.3 %
S06	67.7 %	76.5 %	63.4 %
S07	73.6 %	76.3 %	62.8 %
S08	85.0 %	80.4 %	68.6 %
S09	78.1 %	71.3 %	64.5 %
S10	82.1 %	78.0 %	71.2 %
mean	75.3 %	73.9 %	65.0 %

Figure 1 shows the average waveform of the execution error at electrode Cz for all subjects as well as the topographic distribution of the potential. Although the waveform of the two error potentials differed strongly, the topographic distribution was very similar for both errors and all subjects with the maximum around electrode FCz and Cz.

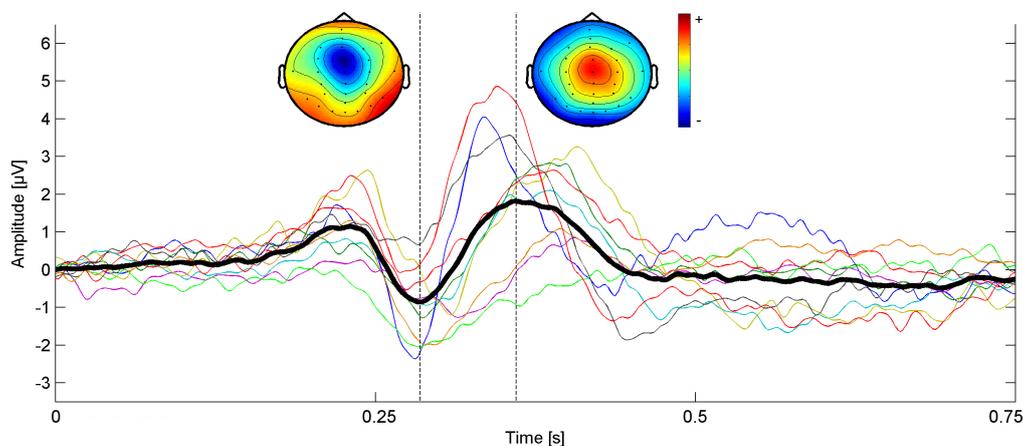


Figure 1: Execution error-related potential at electrode Cz. The colored lines depict the ErrP for the different subjects, while the bold black line is the average over all subjects. The topographic distribution of the potential averaged over all subjects at 285 ms and 360 ms is shown at the top. For display of the ErrPs, the difference between the error trials and noError trials was calculated.

Since the angle of the movement was randomly modified with different degrees, we also tested if execution errors with a different degree can be classified, e.g. 45 ° against 180 °, but did not achieve significant results.

## 4 Discussion and Conclusion

In this study we could show that ErrPs can be measured in the EEG, due to an erroneous response during continuous feedback. While we have further shown that two different types of errors can be discriminated in EEG (outcome error vs. execution error), we could not detect the severity of an execution error (e.g. 45 ° or 180 °). When looking at the shape and topographic distribution of the execution error, it is similar to the results of earlier studies [2, 6], thereby showing that classification is based on an electrophysiological response and not on artifacts. This is supported by the lower classification accuracies obtained on EOG and gamepad data, which are not significantly above chance level for most subjects.

Although we were able to classify both types of error, the classification accuracy needs to be improved to be useable for adaptation or error correction in a BCI application. Therefore it needs to be tested if the power spectrum of the EEG yields additional information to classify those ErrPs. While we could show that ErrPs are elicited during continuous feedback and we are able to classify them based on the EEG, the classification itself is still event-locked and it needs to be tested in a further study how well classification works if it is done continuous [8].

## Acknowledgements

The authors would like to thank Johannes Anufrienko for programming the experimental task. This study was funded by the German Federal Ministry of Education and Research (BMBF, BFNT F\*T, Grant UTü 01 GQ 0831) and the *Deutsche Forschungsgemeinschaft* (DFG, Grant RO 1030/15-1, KOMEG).

## References

- [1] Pierre W Ferrez and José del R Millan. Error-related eeg potentials generated during simulated brain-computer interaction. *Biomedical Engineering, IEEE Transactions on*, 55(3):923–929, 2008.
- [2] Martin Spüler, Michael Bensch, Sonja Kleih, Wolfgang Rosenstiel, Martin Bogdan, and Andrea Kübler. Online use of error-related potentials in healthy users and people with severe motor impairment increases performance of a P300-BCI. *Clinical Neurophysiology*, 123(7):1328–1337, 2012.
- [3] Nico M Schmidt, Benjamin Blankertz, and Matthias S Treder. Online detection of error-related potentials boosts the performance of mental typewriters. *BMC neuroscience*, 13(1):19, 2012.
- [4] Martin Spüler, Wolfgang Rosenstiel, and Martin Bogdan. Online adaptation of a c-VEP Brain-Computer Interface (BCI) based on Error-related potentials and unsupervised learning. *PloS one*, 7(12):e51077, 2012.
- [5] Alberto Llera, Marcel AJ van Gerven, Vicenç Gómez, Ole Jensen, and Hilbert J Kappen. On the use of interaction error potentials for adaptive brain computer interfaces. *Neural Networks*, 24(10):1120–1127, 2011.
- [6] Alex Kreiling, Christa Neuper, and Gernot R Müller-Putz. Error potential detection during continuous movement of an artificial arm controlled by brain-computer interface. *Medical & biological engineering & computing*, 50(3):223–230, 2012.
- [7] Tomislav Milekovic, Tonio Ball, Andreas Schulze-Bonhage, Ad Aertsen, and Carsten Mehring. Error-related electrocorticographic activity in humans during continuous movements. *Journal of neural engineering*, 9(2):026007, 2012.
- [8] Tomislav Milekovic, Tonio Ball, Andreas Schulze-Bonhage, Ad Aertsen, and Carsten Mehring. Detection of Error Related Neuronal Responses Recorded by Electrocorticography in Humans during Continuous Movements. *PLoS ONE*, 8(2):e55235, 02 2013.

# Optimization design for the SSVEP brain-computer interface (BCI) system

Erwei Yin<sup>1</sup>, Jun Jiang<sup>1</sup>, Dewen Hu<sup>1</sup> and Zongtan Zhou<sup>1</sup>

<sup>1</sup>College of Mechatronic Engineering and Automation, National University of Defense Technology, Changsha, Hunan, People's Republic of China  
narcz@163.com

## Abstract

In this paper, we present a optimized steady-state visually evoked potential (SSVEP) brain-computer interface (BCI) system to enhance the information transfer rate (ITR). First, to increase the number of available items, a two-step paradigm was employed, called row/column (RC) paradigm. Second, to improve accuracy, a new signal processing method (called CCA-RV) and a real-time feedback mechanism were designed. Finally, a fixed optimal approaches for setting the optimal stimulus duration was implemented to reach a reasonable online spelling speed. The experimental results with six subjects suggest that the CCA-RV method and the real-time feedback effectively increased accuracy comparing with CCA and without real-time feedback, respectively. Additionally, the fixed optimal approach achieved reasonable online spelling performance.

## 1 Introduction

Recently, an increasing number of researchers have focused on steady-state visually evoked potential (SSVEP) brain-computer interface (BCI) systems [1]. Here SSVEP is a periodic neural response located in the subject's central visual field, which is induced by the visual repetitive stimulus [2]. Due to the high information transfer rate (ITR), simple system configuration, and minimal required user training time, the SSVEP BCI has become one of the most promising paradigms for both disabled and healthy participants in practical BCI application. ITR is an important metric to measure the performance of the BCI system [3]. To increase ITR, we first adapted stimulus mechanism with a row/column (RC) paradigm to increase the number of items. Furthermore, we proposed a new signal processing method for reducing the inter-frequency variation in SSVEP responses (called CCA-RV method), and a real-time feedback mechanism to increase the attention on the visual stimuli, thus enhance accuracy. Finally, an optimal approach for setting the stimulus time was implemented for online spelling.

## 2 Method

### 2.1 RC paradigm

Because the available frequencies in a SSVEP BCI are often restricted, as the RC paradigm in P300 BCI, the targets in SSVEP BCI system are also detected by their row and column coordinates (see Figure 1). 36 characters arranged in a  $6 \times 6$  matrix are employed in our paradigm because it includes all the basic items necessary for spelling. Each cell in the proposed SSVEP BCI flickers between white and black at a constant frequency. Furthermore, we employed six frequencies, set at 8.18, 8.97, 9.98, 11.23, 12.85, and 14.99 Hz, in the design of a periodic stimulus mechanism. In Figure 1, the numbers on the top and left sides of the graphical user interface (GUI) denoting the stimulus frequencies. Specifically, to determine the row coordinate

	1	2	3	4	5	6
1	A	B	C	D	E	F
2	G	H	I	J	K	L
3	M	N	O	P	Q	R
4	S	T	U	V	W	X
5	Y	Z	1	2	3	4
6	5	6	7	8	9	del

Figure 1: Illustrations for the stimulus configuration of the proposed SSVEP BCI.

of the target, all the items in same row flickered at same frequencies. Subsequently, all items in the same column flickered at same frequencies to detect the target’s column coordinate.

### 2.2 Real-time feedback mechanism

During the spelling process, the subjects might reduce their level of attention on the centre of the SSVEP flicker, thus lead to a decrease in accuracy [4]. To this end, we designed a real-time feedback mechanism to improve the efficiency of the SSVEP speller. Feedback is given in real-time to the subject by changing the color of the current target character to green. Specifically, the system first changes the color of the characters within the selected row, and then changes the color of the character associated with selected row and column coordinates. During the spelling process, if subjects find that the character they are looking at has not changed color, they must increase their focus on the target stimuli.

### 2.3 CCA-RV method

In our approach, canonical correlation analysis (CCA) was used to calculate the correlation between the stimulus frequency and the multi-channel electroencephalogram (EEG) data. The correlation coefficient of each stimulus frequency is calculated in real-time ( $score_i(t)$ ). Then we obtained the average non-target scores ( $score_i^{NT}(t)$ ) of each frequency associated with different time points following the initial stimulus presentation. Finally, the SSVEP response scores were evaluated in the following manner:

$$Score_i(t) = \frac{score_i(t) - score_i^{NT}(t)}{score_i(t) + score_i^{NT}(t)} \tag{1}$$

### 2.4 Fixed optimal approach.

We employ the parctical ITR (PITR) proposed by Townsend *et al.* for the optimal spelling time selection, which provides a more realistic estimation of ITR [5]. The time corresponding to maximum PITR is selected as optimal stimulus time. The PITR can be expressed as follows:

$$PITR = \begin{cases} (2P - 1) \log_2 N/T, & P > 0.5 \\ 0, & P \leq 0.5 \end{cases} \tag{2}$$

where  $N$  is the number of items and  $P$  is the spelling accuracy. Here,  $T$  is the time interval per selection, which is computed from the following expression:

$$T = [(t^{row} + t^{column}) + I]/60 \quad (3)$$

where  $t^{row}$  and  $t^{column}$  are the row and column stimulus times, respectively, and  $I$  is the time between successive selections.

The optimal stimulus time is estimated using calibration data and is fixed prior to online BCI use. During online spelling, the SSVEP speller provides the spelling results once the optimal stimulus time is met. To avoid visual fatigue, we set the maximum stimulus time to 10 s. The peak PITR was taken as a measure of the possible performance of the SSVEP speller.

## 2.5 Experimental Setup

Six healthy subjects participated in the study. All subjects had normal or corrected-to-normal vision. All subjects signed an informed consent form in accordance with the Declaration of Helsinki. EEG signals were recorded using a BrainAmp DC Amplifier. Nine-channel active electrodes were selected for the SSVEP detection and were placed at Pz, P3, P4, Oz, O1, O2, POz, PO7 and PO8, referenced to P8 and grounded to Fpz. The EEG signals were sampled at 250 Hz and filtered using a 50-Hz notch filter. The experiments were performed in a normal office room. The subjects were seated in a comfortable chair located approximately 70 cm from a 27" LED monitor. The offline session consisted of two different stimuli conditions - one with real-time feedback and one without. Six runs of each condition were collected. In each run, the subjects were required to input 12 symbols in a random order to avoid adaptation. For each letter selection, the SSVEP stimuli appeared on the screen and remained for 10 s. During online spelling, the subjects were required to spell their own names in Latin letters three times. When an incorrect symbol was detected, the subject had to correct the misspelling by selecting the 'del' option located at the bottom right of the matrix, followed by the correct letter.

## 3 EXPERIMENTAL RESULTS

Figure. 2 presents a comparison of the average classification results achieved using the proposed approaches (CCA-RV and real-time feedback) and traditional SSVEP approaches (CCA and without real-time feedback). The results indicate that the proposed approaches achieved higher average classification accuracy than the traditional SSVEP approaches. The online performance of the proposed SSVEP speller is summarized in Table 1. The average PITR obtained across all subjects using the fixed optimization approach reached  $37.59 \pm 5.55$  bits/min.

Subject	Task Characters	Stimulus Time (s)	Practical Selections	PITR (bits/min)
S1	30	8.24	30	30.29
S2	24	3.92	28	44.91
S3	24	3.92	32	39.30
S4	15	4.48	19	35.59
S5	36	4.72	50	33.24
S6	27	4.40	31	42.21
AVG	26	4.95	31.67	37.59
STD	7.01	1.64	10.13	5.55

Table 1: Online performance of the proposed SSVEP BCI

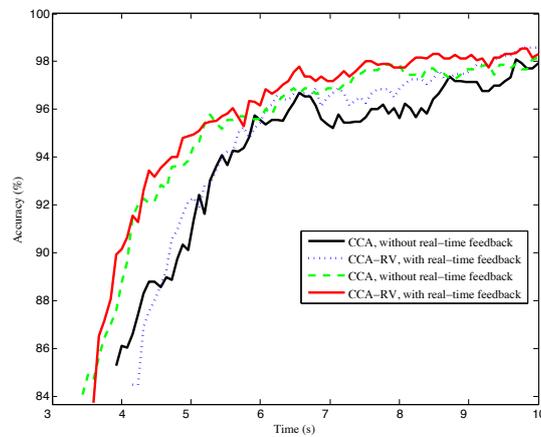


Figure 2: Comparisons of offline results associated with different approaches.

## 4 Conclusions

In this paper, a optimized SSVEP BCI system was proposed in order to increase ITR. Only six frequencies was used to establish the SSVEP BCI with 36 items. In addition, we proposed the CCA-RV method and the real-time feedback mechanism to enhance the accuracy. To achieve reasonable online performance, we designed a fixed optimal approach for selecting stimulus time. Experimental results suggest that the proposed SSVEP speller can provide improved performance compared to traditional BCI approaches. More specifically, online spelling PITS using the fixed optimization approach achieved  $37.59 \pm 5.55$  bits/min.

## 5 Acknowledgments

This work was supported in part by the National High-Tech R&D Program of China (863 Program) under Grant 2012AA011601, the National Natural Science Foundation of China under Grant 61375117.

## References

- [1] H. Cecotti. Spelling with non-invasive brain-computer interfaces-current and future trends. *J. Physiol. - Paris*, vol. 105, pp. 106–114, 2011.
- [2] F. Vialatte, M. Maurice, J. Dauwels, and A. Cichocki. Steady-state visually evoked potentials: Focus on essential paradigms and future perspectives. *Prog. Neurobiol.*, vol. 90, pp. 418–438, 2010.
- [3] J. R. Wolpaw, D. J. M. N. Birbaumer, P. G, and T. M. Vaughan. Brain-computer interfaces for communication and control. *Clin. Neurophysiol.*, vol. 113, no. 6, pp. 767–791, 2002.
- [4] S. Kelly, E. Lalor, C. Finucane, R. Reilly, and J. Foxe. Visual spatial attention tracking using high-density SSVEP data for independent brain-computer communication. *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 13, no. 2, pp. 172–178, 2005.
- [5] G. Townsend, B. K. LaPallo, C. B. Boulay, D. J. Krusienski, G. E. Frye, C. K. Hauser, N. E. Schwartz, T. M. Vaughan, J. R. Wolpaw, and E. W. Sellers. A novel P300-based brain-computer interface stimulus presentation paradigm: moving beyond rows and columns. *Clin. Neurophysiol.*, vol. 121, no. 7, pp. 1109–1120, 2010.

# Brain Painting Version 2 - evaluation with healthy end-users

Loic Botrel<sup>1</sup>, Paul Reuter<sup>1</sup>, Elisa Holz<sup>1</sup> and Andrea Kübler<sup>1</sup>

<sup>1</sup>University of Würzburg, Würzburg, Germany  
loic.botrel@uni-wuerzburg.de

## Abstract

Brain Painting is an application for a brain-computer-interface (BCI) that allows for painting using the brain activity generated by event related potentials (ERPs) in response to maintained attention on visual stimulations. Considering feedback from the extended use of Brain Painting V1 at home by two locked-in patients, we developed a new version of Brain Painting including auto-calibration, new painting features, and a tutorial video explaining its functions. This new version was tested with 10 healthy participants. Results ascertained that the tutorial video is sufficient for BCI Brain Painting novices to perform the subsequent Brain Painting session. Copy Painting was achieved with 94% accuracy and subjects used the application for on average 51 minutes in the free painting mode after termination of copy painting. Feedback from healthy subjects will help us to refine the application before bringing it to the two end-users in the locked-in state.

## 1 Introduction

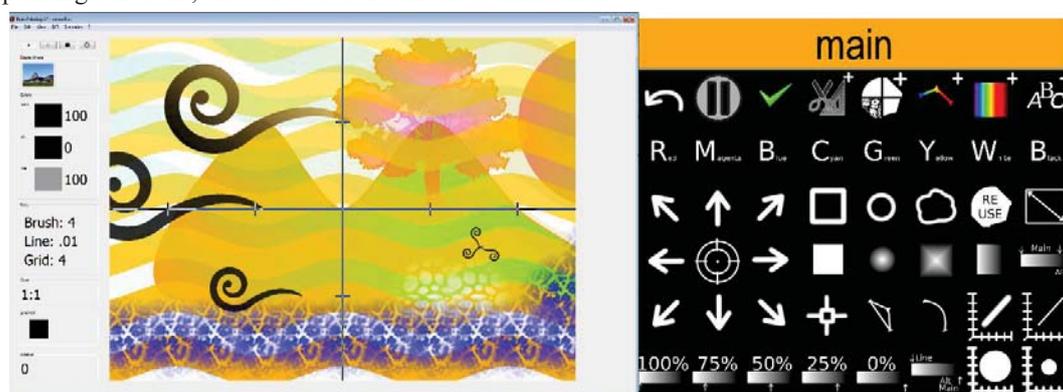
The BCI application Brain Painting (Münßinger, et al., 2010) allows an end-user to paint on a virtual canvas using brain activity without the requirement of motor pathways. Brain Painting is controlled by a BCI with event-related potentials (ERP) as input signal.

Brain Painting allows the end-user to select basic shapes of different color, opacity and size on a virtual canvas displayed on a monitor, such that the user can combine them and compose pieces of art. This system was designed to match the requirements of paralyzed or “locked-in” end-users (Zickler, Halder, Kleih, Herbert, & Kübler, 2013). Formerly restricted to lab or demonstration purposes, the system has been adapted for home use by Holz and colleagues (Holz, Botrel, Kaufmann, & Kübler, under review) and was given to two locked-in patients so they could use Brain Painting on a regular basis for several months. Our first end-user reported to be highly satisfied with the system, and we could measure an improvement in the quality of life. The end-user expressed the wish to have more painting features and a signal validation tool that would ensure a correct setup prior to every use.

Following the user-centered design (Kübler, et al., Manuscript submitted for publication), Brain Painting Version (V) 2 has been developed to face the new challenges encountered in home use environment. Thus, Brain Painting V2 is not only an extension of Version 1, but has been newly programmed to incorporate additional painting features such as lines and text. An autocalibration was also implemented that can easily be performed prior to every session. With the following study, we aimed at evaluating the usability of the system with a sample of healthy users in the laboratory.

## 2 Software

The Brain Painting V2 application uses, Python 2.7, Qt, Pygame and BCI2000. It is composed of 3 independent but connected modules: (I) The graphical user interface (GUI), coded in Python and Qt, contains the painting canvas in the middle and a status bar on the left (see Figure 1) to display important information such as selected colors and brush size. The menu bar on the top allows to load or save paintings. Below the menu bar, there are play, pause and stop buttons that respectively allow to start, pause and stop a Brain Painting session. (II) Running in background, BCI2000 processes the EEG signal, classifies and sends stimulations to the speller. The GUI uses the Telnet protocol interface of BCI2000 to load parameters, start and stop signal processing. (III) The custom P300 speller matrix feedback uses Pygame. It displays the Brain Painting matrix such that one symbol represents one painting function on the canvas. Famous faces flashes are implemented for stimulation to increase the signal-to-noise ratio (Kaufmann, et al., 2013). Several matrices were created to include all requested painting functions, and can be reached from the main menu.



**Figure 1:** (left) Graphical User interface with the Canvas and the status bar, positioned on the side. The session control buttons are above the status bar. (right) 6x8 Speller matrix with symbols associated with Brain Painting functions positioned in front of the participant. The painting on the left was composed during a live demonstration in the ‘Psychologie und Gehirn’ conference in Würzburg (PuG 2013; <http://tagung.dgpa.de/>).



**Figure 2:** Timeline of the study

## 3 Methods

Ten healthy participants (mean age 25 SD=2,8, all students) were included. Each participant was seated in a comfortable chair for the duration of the study (see Figure 2) with a speller stimulation monitor 0.5 to 1 meter distance in front. The second monitor, placed aside, displayed the Brain Painting canvas. A custom 14 electrodes EEG cap (Debener, Minow, Emkes, Gandras, & Vos, 2012) was used for EEG recording. The cap is a low-budget combination of a customer grade EEG amplifier with passive EEG electrodes commonly used in lab experiments.

Teaching someone how to use the Brain Painting is the first step for its use. Bearing in mind independent home use of Brain Painting, we created a tutorial video that would ensure a complete understanding of Brain PaintingV2 without the experimenter being present. The tutorial video lasted 29 min and presented all Brain PaintingV2 functions available in the matrix, by displaying their effect on the painting Canvas (see Figure 1). A female speaker uses a mouse cursor to navigate in the matrices and select available functions while commenting every action. Brain PaintingV2 features were summarized in 12 different categories such as “colors”, “lines” or “text” and presented point per point.

Calibration trials comprised 15 stimulations of each rows and columns. Stimulated symbols were overlaid during 62.5ms with a picture of Einstein. The inter stimulus interval was set to 125ms. The user was instructed to focus on one symbol and count how many times the overlaying Einstein face was visible. After 10 trials, classifier weights were adjusted using a stepwise-linear discriminant analysis for 800 ms post-stimulus. The calibration process returned the minimum number of sequences required to reach 100% spelling accuracy. To ensure high accuracy the number of sequences returned was increased by 2; and maintained throughout copy and free painting.

During the “Copy Painting” task participants were asked to select target indicated on the matrix. Each and every selection had an effect on the canvas. If a wrong selection was made, the feedback instructed the user to revert the error using the “undo” function within the matrix. Accuracy was retrieved during this Copy Painting task, and reported the number of correct selections divided by the number of total selections.

In the Free Painting mode, participants were given 15 to 60 minutes to compose two paintings of their own with the new Brain PaintingV2. The experimenter did not provide any hint or help concerning the painting interface during the Free Painting task nor in the Copy Painting task. Time spent on Free Painting was recorded

Participants were given visual analogue scales (VAS) at three different time points during the session: after the tutorial (t1), the Copy Painting task (t2) and the Free Painting task (t3). They were asked to rate, on a scale from 0 to 10, how well they understood each of the 13 Brain Painting features we distinguished.

After the Free Painting task, users were asked to answer on a VAS from 0 to 10, how satisfied and how frustrated they were during the Free Painting session. Participants also filled in the NASA-TLX to indicate the workload of the Brain Painting V2 (0 for no workload to 100 for maximum workload).

Subject	Nb. Seq.	Accuracy	Work load	Satisfac- tion	Frustra- tion	Free Painting duration	Selection / minute	Evalu- ation (t3)
A	4	.97	19	7.7	.9	32	3.3	9.9
B	4	.95	66	5.1	6.5	23	3.6	8.6
C	5	.95	75	3.7	4	68	2.9	6.8
D	5	1	54	9.6	0	60	2.8	9.2
E	5	1	n/a	9.4	1.2	60	2.9	1
F	5	.91	59	6.2	4.5	51	2.9	6.7
G	6	.90	40	7.1	2.5	58	2.5	8.4
H	4	.87	74	4.7	8.1	53	3.3	8.5
I	5	.97	20	5.8	4.8	52	2.9	8.7
J	5	.84	60	7	.8	55	2.8	7.9
Mean	4.8	.94	52	6.6	3.3	51	3	8.5
SD	.63	.05	21	1.9	2.7	12	.31	1.1

**Table 1:** Individual results indicating parameters, performance and evaluation results. Mean and standard deviation are mentioned at the bottom of the table. “Nb. Seq.” states for number of sequences.

## 4 Results

Participants achieved mean accuracy of 94% (see Table 1 for results of all participants) during the Copy Painting task. Participants spelled with an average of 4.8 sequences which allowed an average of

3 selections per minute. According to the VAS participants understood well the Brain Painting features with an average rating of 9 after the tutorial video (t1), 9 after the Copy Painting task (t2) and 8.5 after the Free Painting task (t3). During the Free Painting task, participants painted during on average 51 minutes. In the VAS applied after Free Painting, participants reported being satisfied with their Free Painting session (M=6.6) and reported low to average frustration (M=3.3). The total workload according to the NASA TLX was M= 52.

## 5 Discussion

Firstly, our results indicate that the auto-calibration worked fine for all subjects. Further, the tutorial video was very effective because participants reported good understanding of the Brain Painting features. Accuracy was high during Copy Painting corroborating the functionality of Brain Painting V2. After more than 1 hour of sustained use, participants reported average to high satisfaction and low frustration for a medium workload. That indicated good efficiency and high overall satisfaction (Kübler, et al., Manuscript submitted for publication). They spent on average 51 minutes with Free Painting although the experimenter only asked for at least 15 minutes, meaning that participants truly enjoyed Brain Painting. The low-budget EEG cap renders the system even more applicable. It is now safe to provide our locked-in patients with the Brain Painting V2, and investigate whether they will benefit from the practical outcomes of the enhancements of Brain PaintingV2.

### Acknowledgments

The study was funded by the European Community for research, technological development and demonstration activities under the Seventh Framework Programme (FP7, 2007-13), project grant agreement number 288566 (BackHome). This paper reflects only the authors views and funding agencies are not liable for any use that may be made of the information contained herein.

## References

- Debener, S., Minow, F., Emkes, R., Gandras, K., & Vos, M. (2012). How about taking a low-cost, small, and wireless EEG for a walk? *Psychophysiology*, S. 1617-1621.
- Holz, E. M., Botrel, L., Kaufmann, T., & Kübler, A. (under review). A. Long-term independent brain-computer interface home use improves quality of life of a patient in the locked-in state: A case study. *Archives PMR*.
- Kaufmann, T., Schulz, S. M., Köblitz, A., Renner, G., Wessig, C., & Kübler, A. (2013). Face stimuli effectively prevent brain-computer interface inefficiency in patients with neurodegenerative disease. *Clinical Neurophysiology*, pp. 893-900.
- Kübler, A., Holz, E. M., Kaufmann, T., Zickler, C., Kleih, S. C., Staiger-Sälzer, P., . . . Mattia, D. (Manuscript submitted for publication). The User-Centered Design in Brain-Computer Interface (BCI) research: A novel perspective for evaluating the usability of BCI- controlled applications. *Journal of Neurorehabilitation and Neural Repair*.
- MünBinger, J. I., Halder, S., Kleih, S. C., Furdea, A., Raco, V., Höhle, A., & Kübler, A. (2010). Brain Painting: first evaluation of a new brain-computer interface application with ALS-patients and healthy volunteers. *Frontiers in neuroscience*, S. 4.
- Zickler, C., Halder, S., Kleih, S. C., Herbert, C., & Kübler, A. (2013). Brain Painting: Usability testing according to the user-centered design in end users with severe motor paralysis. *Artif Intell Med.*, 99-110.

# Cognitive Rehabilitation through BNCI: Serious Games in BackHome

Eloisa Vargiu<sup>1</sup>, Jean Daly<sup>2</sup>, Stefan Dauwalder<sup>1</sup>, Elaine Armstrong<sup>2</sup>,  
Suzanne Martin<sup>3</sup> and Felip Miralles<sup>1</sup>

<sup>1</sup> Barcelona Digital Technology Center, Barcelona, Spain

{[evargiu](mailto:evargiu@bdigital.org), [sdauwalder](mailto:sdauwalder@bdigital.org), [fmiralles](mailto:fmiralles@bdigital.org)}@bdigital.org

<sup>2</sup> Cedar Foundation, Belfast Northern Ireland, U.K.

{[j.daly](mailto:j.daly@cedar-foundation.org), [e.Armstrong](mailto:e.Armstrong@cedar-foundation.org)}@cedar-foundation.org

<sup>3</sup> Faculty of life and Health Sciences, University of Ulster, Northern Ireland, U.K.

[s.martin@ulster.ac.uk](mailto:s.martin@ulster.ac.uk)

## Abstract

This paper presents a framework for the development of cognitive skills. In particular, we focus on the definition and provision of cognitive rehabilitation tasks to people with neurological disorders returning at home after a discharge. Those tasks will be executed through a Brain-Neural Computer Interface system and will consist of desktop-based serious games. The framework and the implemented games are part of the BackHome project and are currently under testing at the end-users facilities, in Belfast and Würzburg.

## 1 Introduction

Cognitive Rehabilitation (CR) is defined as “a systematic, functionally orientated service of therapeutic activities that is based on assessment and understanding of the patient’s brain-behavioral deficits” directed toward many areas of cognition, including (but not necessarily limited to) attention, concentration, perception, memory, comprehension, communication, reasoning, problem solving, judgment, initiation, planning, self-monitoring, and awareness [2].

CR is a core intervention in rehabilitation, aimed at restoring function or compensating for a cognitive defect. When cognition has been adversely affected, this is considered a predictor of other important aspects of psycho-social recovery [6]. When a cognitive disability is present, people are limited in their capacity to plan and sequence thoughts and actions, conceptualize ideas, and to interpret the meaning of social and emotional cues, and numbers and symbols [5] [1]. Therapeutic interventions may emerge from established approaches which assist the therapist in the planning and delivery of treatment. A restorative approach focuses on strengthening and improving functional performance that has been impaired by developing cognitive skills and retraining.

In this paper, we focus on the definition and provision of CR tasks to people with neurological disorders (traumatic brain injury) going back to home after a discharge. This research adopted a user centric design philosophy in the development of a rehabilitation tool that would focus on CR to complement and improve upon traditional methods. Tasks will be executed through a Brain-Neural Computer Interface (BNCI) system and will consist of desktop-based serious games [4]. Using desktop-based serious games within therapy is not yet established but emerging. This study provides the novel and unique opportunity to provide CR tasks by relying on a BNCI system.

This work is part of the BackHome project<sup>1</sup> [3]. The CR framework is one of the BNCI

---

<sup>1</sup><http://www.backhome-fp7.eu/>

services provided by the BackHome platform. Moreover, the implemented games are fully integrated in the telemonitoring and home support system developed within the project [7].

## 2 Method

Working directly with service providers for people with traumatic brain injury, a framework for the development of cognitive skills has been created against which cognitive tasks will be mapped. The adopted framework as well as the implemented games have been defined together with Occupational Therapists and Speech and Language Therapists. Initially, two focus groups were held (N= 8; N=58) to identify the user requirements for the overall system. A research collaboration was then formed with Specialist neurological Occupational Therapists and Speech and Language therapist who have worked with us over the past year to develop and review the various prototype of the CR application. The therapists have meet with the researchers on three occasions to date in the lab (N=10; N=9; N=3). The technical developers integrated the therapist’s recommendation into the system for the therapists to review the next iteration of the CR prototype. The system development has been iterative and incremental to enable reflections with the therapists prior to changes in the system becoming permanent.

As shown in Figure 1, the domains emerging with the framework reflect various levels of cognitive complexity: perception, attention and concentration, memory, and executive functions. The sequential approach is important, as problems in lower levels will impact on the ability of the person to advance in their cognitive competencies. The framework provides a backdrop to CR tasks performed using a BNCI system. It also provides a base for the future development of validated CR interventions by therapists with expertise in this area.

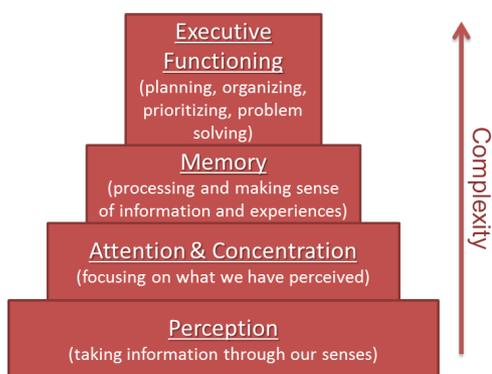


Figure 1: Cognitive skills.



Figure 2: The memory-card game starts with all the cards covered.

## 3 The Memory-card Game

Currently, two serious games have been defined and implemented: find-a-category and memory-card. In this paper, we presents the memory-card game that is aimed at enhancing memory skills. The game starts showing on the P300 interface all the cards covered (see Figure 2), the user chooses which card turning out and then the second one looking for its pair. If the pair is

not found, the cards are covered again. On the contrary, the game follows with the remaining cards. The game ends when all the pairs have been found.

Different levels of difficulty have been defined, depending on the ability of the user. In particular, difficulty increases by augmenting the number of available cards -i.e., the dimension of the P300 matrix (level 1, 8 cards; level 2, 12 cards; and level 3, 20 cards)- and the complexity of the adopted images (level 1, shapes; level 2, fruits; and level 3, animals). Figure 3-a shows an example for each of the given levels (for the sake of clarity, the P300 matrix has been omitted).

Results are then shown to the user in numerical and graphical view (see Figure 3-b). The numbers of corrected moves with respect to the total moves represents the score. The score for the last 8 played games (a line for each level) is presented. Moreover, the number of moves together with the elapsed time are shown and some statistics given. All the results are also sent to the therapist.

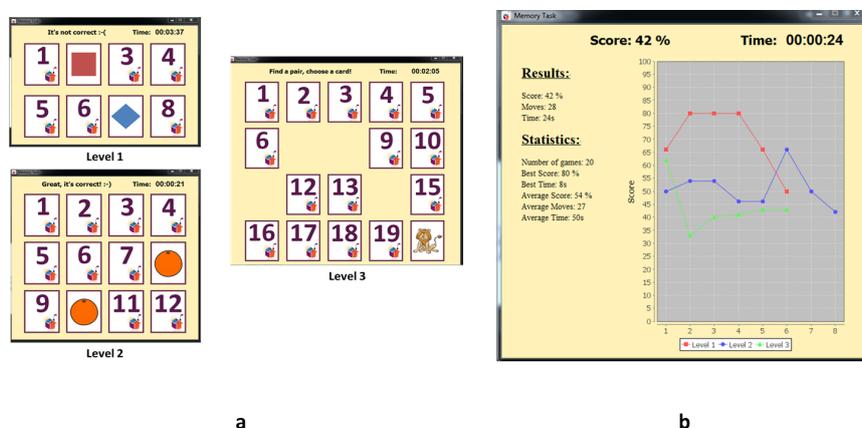


Figure 3: A screenshot of memory-card game for each given level (a) and the results display (b).

## 4 Discussion

In this qualitative research a user centered approach has been adopted. User centered design is a process of engagement with end users that adopts a range of methods to place the end of user of a technology at the center of the design process in terms of development and evaluation.

It is beyond the scope of our work to develop, validate, and illustrate clinical outcomes. Nonetheless, therapists are able to interact with users in real time, assigning game sessions depending on the level of disability, thanks to a therapist station<sup>2</sup>. Then, they monitor the use and outcomes of the CR tasks to attain therapeutic results. In fact, the ability for the therapist to plan, schedule, telemonitor and personalize the prescription of CR tasks using a therapist station will facilitate that the user performs those tasks inside her therapeutic range (i.e. motivating and supporting her progress), in order to help to attain beneficial therapeutic results.

The second-year-prototype of BackHome integrates all the implemented games as well as the therapist station. That prototype is currently under testing by the end-users of the project,

<sup>2</sup><https://station.backhome-fp7.eu:8443/BackHome>

located at Cedar Foundation in Belfast and at the University of Würzburg.

Preliminary experiments have been performed in Belfast at Cedar Foundations premises, a total of N=5 end users with a varying degree of cognitive impairments were involved in the evaluation. We calculated the accuracy (as percentage of correctly chosen symbols) and therefore, asked users during every trial, which cards they wanted to uncover. Results showed an average selection accuracy of 78%. Those users, who had no or low control over the games, had no or low control over the BNCI in general. This was either caused by the severity of their physical impairment or bad EEG signals due to problems with data acquisition.

## 5 Conclusions and Future Work

In this paper, we presented a framework that provides a backdrop to cognitive rehabilitation tasks performed using a BNCI system. The framework is part of the BackHome project and, according to the cognitive skills that it defines, two serious games have been defined and implemented. The corresponding system is currently under testing by people with traumatic brain injury.

As for the future work, we are currently implementing a daily-life-activity game aimed at improving perception. Moreover, apart serious games, we are also considering defining and implementing cognitive rehabilitation tasks relying on an e-puck robot, which might become especially engaging to users.

## Acknowledgments

The research leading to these results has received funding from the European Community's, Seventh Framework Programme FP7/2007-2013, BackHome project grant agreement n. 288566.

## References

- [1] David Braddock, Mary C Rizzolo, Micah Thompson, and Rodney Bell. Emerging technologies and cognitive disability. *Journal of Special Education Technology*, 19(4), 2004.
- [2] Keith D Cicerone, Cynthia Dahlberg, Kathleen Kalmar, Donna M Langenbahn, James F Malec, Thomas F Bergquist, Thomas Felicetti, Joseph T Giacino, J Preston Harley, Douglas E Harrington, et al. Evidence-based cognitive rehabilitation: recommendations for clinical practice. *Archives of physical medicine and rehabilitation*, 81(12):1596–1615, 2000.
- [3] J. Daly, E. Armstrong, F. Miralles, E. Vargiu, G.R. Müller-Putz, C. Hintermüller, C. Guger, A. Kübler, and S. Martin. Backhome: Brain-neural-computer interfaces on track to home. In *RAatE 2012 - Recent Advances in Assistive Technology & Engineering*. 2012.
- [4] P. Rego, P.M. Moreira, and L.P. Reis. Serious games for rehabilitation: A survey and a classification towards a taxonomy. In *Information Systems and Technologies (CISTI), 2010 5th Iberian Conference on*, pages 1–6, June 2010.
- [5] Marcia J Scherer. *Assistive technologies and other supports for people with brain impairment*. Springer Publishing Company, 2011.
- [6] David J Schretlen and Anne M Shapiro. A quantitative review of the effects of traumatic brain injury on cognitive functioning. *International Review of Psychiatry*, 15(4):341–349, 2003.
- [7] Eloisa Vargiu, Juan Manuel Fernández, Sergi Torrellas, Stefan Dauwalder, Marc Solà, and Felip Miralles. A sensor-based telemonitoring and home support system to improve quality of life through bnci. In *Assistive Technology: From Research to Practice, AAATE 2013. Encarnao, P., Azevedo, L., Gelderblom, G.J., Newell, A., Mathiassen, N.-E. (Eds.)*, 2013.

# Multi-state driven decoding for brain-computer interfacing

Minkyu Ahn<sup>1</sup>, Hohyun Cho<sup>1</sup>, Sangtae Ahn<sup>1</sup> and Sung Chan Jun<sup>1\*</sup>

<sup>1</sup>Gwangju Institute of Science and Technology, Gwangju, South Korea  
minkyuahn@gmail.com, augustcho@gist.ac.kr, stahn@gist.ac.kr, and  
scjun@gist.ac.kr

## Abstract

Performance variation is one of the critical issues to be resolved in brain-computer interface field. Subjects exhibit different performances from session to session and even across trials. To overcome this issue, three strategies have been commonly proposed, including extraction of robust features, adaptation methods, and monitoring and rejection of bad trials. In this work, we suggest a new multi-classifier strategy using trial data shuffling. This strategy generates different classifiers according to various noise states. Our proposed strategy showed an improvement in classification performance of approximately 3 percent, and a trial-wise quality measure facilitated to monitor bad trials. This seems a promising method to improve the reliability of the BCI system.

## 1 Introduction

Brain-Computer Interface (BCI) suffers from performance variation. Within subjects, the non-stationary properties of brain signals are considered to be a major cause of this problem. Existing strategies can be categorized as follows: 1) extraction of robust features (Cho et al., 2012), 2) adaptation of feature or classifier (Krusienski et al., 2011; Shenoy et al., 2006; Satti et al., 2010), and 3) monitoring and rejecting bad trials (Ferrez and del R Millan, 2008). Brain signals yield both meaningful (which is used to decode a user's intention or thought) and non-relevant information (background noise). We focused primarily on the first type of information under the assumption that meaningful information is the same over trials. Therefore, most studies have assumed that features are extracted well and reflect task-related information. In practice, however, noise varies over time; this variation influences the overall signal property, which degrades the performance of the feature extractor. Thus, noise information should be considered carefully. Moreover, a user may change the way s/he conducts a mental task or may not be in a proper mental state. Therefore, these trials may violate the assumption above, thus calling into question the use of the methods in Strategies 1 and 2. Grosse-Wentrup and Schölkopf (2012) demonstrated that the prediction of poor mental state, in which a user is not likely to generate a motor imagery-related signal, is possible before the user begins the imagination task. It is remarkable that BCI can predict a user's condition in advance; therefore, the system enables us to deal more readily with the potentially unreliable state of a user. In this study, we propose a new multi-state driven (MSD) method that deals with the problem of fluctuations in BCI. The concept of this MSD strategy is described in Section 2. Materials, evaluation and results are presented in the subsequent sections. Finally, we discuss related ideas and our future work.

---

\*Corresponding author, [scjun@gist.ac.kr](mailto:scjun@gist.ac.kr): Research supported by National Research Foundation of Korea (Grant No. NRF-2013R1A1A2009029).

## 2 Methods and materials

### 2.1 Multi-State Driven method

Typically, every trial obtained in the calibration phase is used to construct one classifier. However, this is likely to overlook two points. First, why is the entire trial dataset used simultaneously in order to construct a classifier? In general, a classifier depends heavily on the given dataset. Thus, performance will vary over the datasets used. Second, why is only one classifier used? As reported in (Lotte et al., 2007), a multi-classifier method gives better and more stable results. The use of one classifier is risky, as its performance may vary greatly over trials. To reduce this risk, we may separate one dataset into multiple sub-datasets containing a small number of trials and use those to generate multiple classifiers (one per sub-dataset). Finally, a decoder is constructed with a reasonable combination of these classifiers. Given trial dataset  $T$ , it is assumed that each trial of dataset  $T$  consists of the sum of  $s$  (task-involved information) and noise  $n$ . With just several trials, a classifier can be constructed through a function associated with preprocessing, and a variety of filtering and feature selections. Therefore, Strategy 1 focuses on extracting the signal  $s$  only, while Strategy 2 considers  $s + n$  simultaneously. If noise variability over trials is high and sub-dataset  $D \subset T$  has its own unique noise state  $N (= \sum_i n_i)$ , these will influence classifier construction and classification is likely to work poorly for trials not contained in  $D$ . We defined this unique noise property as state  $N$ . This noise state  $N$  definitely varies over sub-datasets and some sub-datasets may contain noise, as well as information unrelated to the task that is caused by a user. Thus, we expect that such a variable noise state may become involved in the construction of various classifiers; we call this idea a MSD strategy. Any type of pre-processing and feature selection can be applicable to this MSD strategy. Many ideas may exist with respect to generating a decoder that considers outputs from these multiple classifiers. In this study, a “voting” method (Lotte et al., 2007) was introduced and a single value was used as a quality score for each trial. By introducing MSD, multiple outputs (one output per classifier) coming from many classifiers are generated when a trial is given as input. These multiple outputs may be used to evaluate the measure of quality for a given input trial. The number of class labels classified is counted from multiple outputs and the probability for each trial is estimated. This probability shows how well each trial is discriminated and thus we used this trial discriminant score (DS) to monitor bad trials in order to confirm the improvement in accuracy.

### 2.2 Motor imagery experiment

The following experiment was approved by the Institutional Review Board at Gwangju Institute of Science and Technology. All subjects were informed of the experimental process and purpose before the experiment and signed letters of consent were collected from every subject. With four subjects, we recorded electroencephalographs (EEG) through a Biosemi active2 (64 channels, sampling rate: 512 Hz). Three sessions were conducted with each subject and each session consisted of offline and online experiments. In the offline experiments, a trial began with a blank screen; the instruction bar appeared on the left, right or bottom side of the screen after 2 sec. Subjects were instructed to imagine movement of a body part for 2 sec when a ball was presented. After another 2 sec, the ball moved in the instructed direction. 1<sup>st</sup> and 2<sup>nd</sup> runs consisted of 30 trials, as these two runs were used to identify the best pair of motor imagery (i.e., left and right hand, and foot movement imagination). At the end of the 2<sup>nd</sup> run, we estimated cross-validated accuracy from 60 trials for three conditions through Common Spatial Pattern (CSP) and Fisher Linear Discriminant Analysis (FLDA). Here we applied bandpass filtering (8-30 Hz) and the temporal interval was selected manually after examining event-related (de)synchronization patterns. For the 3<sup>rd</sup> and 4<sup>th</sup> runs, subjects conducted twenty trials per condition. Therefore, this yielded forty more trials per condition. Finally, we collected sixty trials for each condition in a two class motor imagery experiment. From these trials, we constructed a classifier

using invariant Common Spatio-Spectral Pattern (iCSP) and FLDA, as in (Cho et al., 2012). This classifier was applied in the online experiment. Subjects conducted 150 trials for two conditions (75 trials for each) over 3 runs. A feedback trial similar to the trials in the offline experiment was designed. However, in the feedback trials, we gave a classified result and moved the ball in that direction.

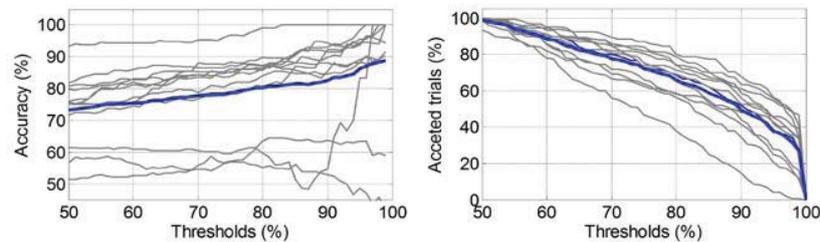
## 2.3 Evaluation

The purpose of this study was to evaluate our proposed strategy and compare it to the conventional strategy that uses whole trials for one classifier. For this evaluation, we fixed informative intervals as 8-30 Hz and 0.4-2.4 sec. for spectral and temporal filtering (Ahn et al., 2012). The evaluation was conducted through ‘in-phase’ and ‘phase to phase’ performances. To estimate performance in the calibration/feedback phase, we first applied cross-validation to check in-phase performance. Whole trials were divided into 10 groups and 7 groups were chosen as a training set, while the remaining 3 groups were used as a testing set. CSP and FLDA were applied, such that 120 iterations yielded 120 accuracy estimates; we assigned the mean accuracy to conventional (CV) performance. We also estimated MSD accuracy. In 10 groups of trials, we chose nine groups as training data and the remaining group was used as testing. From 9 groups in the training set, 7 groups were selected to generate a classifier; therefore, we were able to construct an MSD decoder consisting of 36 classifiers. This decoder evaluated each trial in the test data, and the process was repeated through 10-fold validation; thereafter, a trial-wise discriminant value was estimated and we calculated MSD accuracy. Finally, we obtained both CV and MSD accuracy for each calibration and feedback phase. Conventional phase-to-phase performance is designed so that a classifier constructed by calibration data evaluates the trials in the feedback phase. We applied this conventional method for our calibration and feedback data through CSP and FLDA. In addition, MSD was implemented and applied. The MSD decoder constructed from calibration data was applied to discriminate the trials in the feedback phase.

MSD produces trial-wise discriminability from classifiers, which facilitates the evaluation of trial quality. It is possible to reject a bad trial that falls beneath a certain threshold. This was evaluated for phase to phase performance. The MSD decoder constructed using calibration data evaluated trials in the feedback phase and bad trials were rejected. For this evaluation, we used a 50% threshold, which means that the MSD decoder identified a certain class that received at least more than one evaluation. However, we applied different criteria to examine how the accuracy changed. Based on this criterion, we accepted trials showing higher probability than the threshold and the hit rate was calculated.

**Table 1:** Performance (%) comparison between conventional approach (CV) and MSD

Datasets	In-phase (Calibration)		In-phase (Feedback)		Phase to phase	
	CV	MSD	CV	MSD	CV	MSD
A1	85.0	87.5	72.9	80.7	75.3	78.7
A2	92.8	95.8	58.9	62.0	75.3	75.3
A3	86.6	90.8	65.8	71.4	70.7	75.7
B1	80.6	84.2	94.1	95.3	70.0	71.3
B2	88.5	89.2	80.5	83.3	80.7	80.7
B3	80.9	85.0	84.8	87.3	69.3	75.3
C1	89.4	91.7	76.5	79.3	77.3	81.3
C2	81.8	86.7	62.9	66.0	58.0	61.3
C3	60.2	60.0	51.6	51.3	51.3	56.0
D1	60.8	61.7	61.7	62.0	50.0	51.3
D2	75.3	78.3	87.1	87.3	72.0	76.0
D3	85.7	87.5	94.7	94.0	92.0	93.3
Mean	80.6	83.2	74.3	76.7	70.2	73.0
Standard Deviation	10.5	11.3	14.3	14.0	12.1	11.7
<i>p</i> value (*:<.01, **:<.005, ***:<.001)	0.00097656 (***)		0.0068359 (*)		0.0019531 (**)	



**Figure 1:** Accuracy (left) and accepted trials (right) in % over varying threshold levels (50% to 100%). Each thin, gray line denotes the results for each dataset and the thick, blue line indicates the mean.

### 3 Results and conclusions

As shown in Table 1, MSD resulted in higher performance than CV accuracy showing the improvement 2.6% (calibration), and 2.4% (feedback). Wilcoxon signed rank tests revealed the significance at  $p < 0.001$  and  $p < 0.01$ , respectively. In phase to phase performance, the MSD was also superior. The mean accuracy increased from 70.17% to 73.03% ( $p < 0.005$ ). In the test of the applicability of the trial rejection method, we observed that, for most datasets, the accuracy increased as the criterion became stronger (Figure 1). Meanwhile, the number of trials accepted decreased steadily. This demonstrated that most datasets, except for C2, C3 and D1, generated reasonably good task-related information. Thus, trial rejection might be a more efficient technique to apply. In this study, we proposed a new strategy to make the BCI system more reliable. Our results showed notable improvement in performance. To overcome performance variation, further investigations of trial-wise and session-wise data are required. Our future studies will investigate this issue.

### References

- Ahn, M., Hong, J., & Jun, S. (2012). Feasibility of approaches combining sensor and source features in brain-computer interface. *Journal of Neuroscience Methods*, 204(1), 168 - 178.
- Cho, H., Ahn, M., Sangtae, A., & Jun, S. (2012). Invariant Common Spatio-Spectral Patterns. *3rd TOBI Workshop*, (pp. 31-32).
- Ferrez, P., & del R Millan, J. (2008). Error-related EEG potentials generated during simulated brain-computer interaction. *IEEE transactions on bio-medical engineering*, 55(3), 923-929.
- Grosse-Wentrup, M., & Schölkopf, B. (2012). High gamma-power predicts performance in sensorimotor-rhythm brain-computer interfaces. *Journal of Neural engineering*, 9(4), 046001.
- Krusienski D. J., Grosse-Wentrup M., Galan F., Coyle D., Miller K. J., Forney E., Anderson C. W. (2011) Critical issues in state-of-the-art brain-computer interface signal processing. *Journal of Neural engineering*, 8(2), 025002.
- Lotte, F., Congedo, M., Lécuyer, A., Lamarche, F., & Arnaldi, B. (2007). A review of classification algorithms for EEG-based brain-computer interfaces. *Journal of Neural Engineering*, R1.
- Satti, A., Guan, C., Coyle, D., Prasad, G. (2010) A Covariate Shift Minimisation Method to Alleviate Non-stationarity Effects for an Adaptive Brain-Computer Interface. *20th International Conference on Pattern Recognition*, (pp. 105–108).
- Shenoy, P., Krauledat, M., Blankertz, B., Rao, R., & Müller, K.-R. (2006). Towards adaptive classification for BCI. *Journal of neural engineering*, R13-23.

# Fixation-related potentials: Foveal versus parafoveal target identification

Anne-Marie Brouwer<sup>1</sup>, Manje Brinkhuis<sup>1,2</sup>, Boris Reuderink<sup>3</sup>, Maarten Hogervorst<sup>1</sup>, Jan van Erp<sup>1</sup>

<sup>1</sup>TNO, Soesterberg, The Netherlands

<sup>2</sup>University of Iceland, Reykjavik, Iceland; Utrecht University, Utrecht, the Netherlands

<sup>3</sup>Cortex, Amersfoort, The Netherlands

{anne-marie.brouwer, maarten.hogervorst, jan.vanerp}@tno.nl,  
manjebrinkhuis@gmail.com, boris@cortex.nl,

## Abstract

The P300 event-related potential (ERP) can be used to infer whether an observer is looking at a target or not. Common practice in P300 BCIs and experiments is that observers are asked to fixate their eyes while stimuli are presented. First studies have shown that it is also possible to distinguish between target and non-target fixations on the basis of single ERPs in the case that eye movements are made, and ERPs are synchronized to fixations (fixation-related potentials or FRPs) rather than to stimulus onset. However, in these studies small object sizes ensured that participants could only identify whether the object was a target or non-target after fixating on it. We here compare (non-)target FRPs when objects are identified before versus after fixation. We also examine ERPs from static eyes conditions. FRP shapes are in accordance with the notion that the late component of the P300 is associated with identifying a target, and eye movements do not substantially affect the P300. Even when the time of object identification is unknown, it is possible to distinguish between target and non-target FRPs on a single FRP basis. These results are important for fundamental science and for developing applications to covertly monitor observers' interests.

## 1 Introduction

The P300 event related potential (ERP) occurs 300-500 ms after a stimulus has been presented that draws attention, either through bottom-up mechanism as in odd-ball paradigms or through conscious guidance, as made use of in P300 BCIs. In P300 experiments and BCI paradigms, participants are usually asked not to move their eyes around the time that the P300 is expected to occur, in order to avoid eye movement artefacts in the EEG. It is however expected that the same kinds of processes and ERPs occur when observers actively sample their visual environment themselves by making eye movements. In this case, one would need to examine the EEG traces relative to self-paced fixation onset (fixation-related potential, or FRP). There are only few studies that demonstrate that fixation-related P300s are elicited by top-down determined target identification (Kamienkowski et al., 2012; Brouwer et al., 2013). These studies control for potentially confounding factors that (may) differ systematically between target and non-target fixations, such as fixation duration, preceding saccade length and low-level visual properties of the objects of interest. Brouwer et al. (2013) showed that target and non-target FRPs could be distinguished above chance on a single FRP level. In the

experiments by Kamienkowski et al. and Brouwer et al., small object size ensured that participants could only identify the object (target or non-target) after fixating it. However, in real life, targets can also be identified parafoveally ('in the corner of the eyes'), prior to fixation on the target. ERP experiments in which participants fixated a fixation cross demonstrated that target stimuli presented in the periphery also elicit P300s. We expect P300 FRPs to occur earlier when targets are identified before than after fixation. In the current study, we examine FRPs when a target has been identified before fixating it (large target) and after (small target). We also include a condition in which the eyes do not move in order to examine whether planning (or inhibiting) a saccade affects the P300 latency and amplitude. Our main interest is in whether we can still distinguish between targets and non-targets at a single FRP level in the case that time of target identification relative to fixation is unknown. This would be required if one is interested in monitoring the interest of naturally behaving observers.

## 2 Methods

EEG-electrodes were placed at Fz, Pz, Cz, POz, Oz, P3, P4, PO7 and PO8 with a ground electrode at FPz. The EEG electrodes were referenced to linked mastoid electrodes. Four electrodes were used for EOG recording: two at the outer canthi left and right; one above and one below the left eye. Both horizontal and vertical EOG-electrodes were referenced to each other. Data were recorded at 256 Hz and filtered online using a 0.1 Hz high-pass, a 100 Hz low-pass and a 50 Hz notch filter.

Stimuli were shown on a 20" tft-screen with a refresh rate of 60 Hz and at a viewing distance of 60 cm. They were circular patches with a black center and white surround, or the other way around. One type was designated as target, the other as non-target. Center and surround had an equally sized surface area. They were shown against a gray background in one of four locations (top-left, top-right, bottom-left, or bottom-right) at 12° of visual angle from a fixation cross, located in the center. The stimuli were sized either 4° or 0.25° in diameter, so that only the large stimuli could be identified when gaze was at the fixation cross, whereas an eye movement was required for the small stimuli. All patches were surrounded by a black square frame which had thickness of two pixels, allowing the participants to easily detect the location of the (small) stimulus.

Twelve observers participated. During the task, the fixation cross was present continuously. One after the other, the stimuli were presented for 1000 ms in one of the four corners (random order) with an inter-stimulus interval of 750 ms. The participants were asked to count targets. After every 21 trials, the participant indicated the number of counted targets through the keyboard.

There were three conditions: one in which participants were asked to keep fixating the cross and that contained only large stimuli (static-large); one in which they were asked to make an eye movement to the stimulus as soon as it appeared that contained only large stimuli (saccade-large); and the same one but with small stimuli (saccade-small). Each condition contained 336 trials, 112 of which were target presentations. Targets and non-targets were presented in random order. Trials from the two saccade conditions were presented in random order within one saccade block. Half of the participants started with the saccade block, the other with the static fixation block. Type of target (black center or white center) was also counterbalanced within these two participant groups.

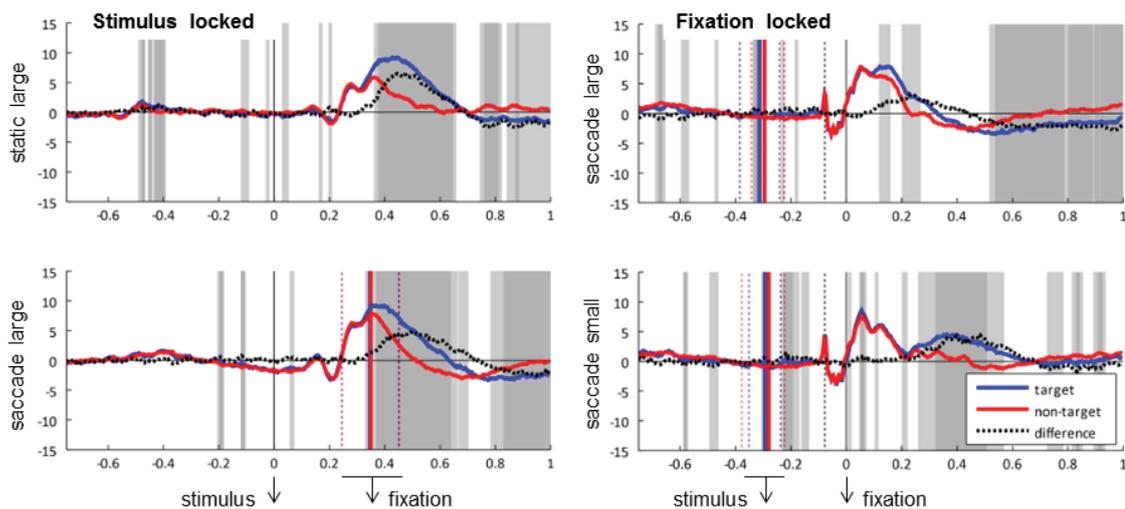
To determine fixation onset in the saccade condition, we detected the peak EOG eye movement speed after stimulus onset and set eye fixation onset at the end of the saccade related peak. For some analyses, we identified and removed eye movement artefacts from the EEG using ICA (Jung et al., 2000). Independent components that reflected EM activity were selected manually by comparing components with the original EOG data. For eleven participants the first two components and for one the first three components were identified to reflect EM and removed from the original data.

For each participant, we estimated whether ERPs (running from stimulus onset until 500 ms later) and FRPs (from fixation onset until 500 ms later) could be correctly classified as associated with a

target or non-target. These estimates were produced using a five-fold cross-validation procedure. We used the same classification pipeline as described in Brouwer et al. (2013). In short, an L2-regularized logistic regression classifier was trained on (9 electrodes \* 129 samples) 1161 dimensional feature vectors. All adjustable parameters were optimized independently of the test sets. The classification procedure was performed on 8 different sets of data as indicated in Table 1. For ERPs, we determined classification performance for data from the static large condition. For FRPs, we determined classification performance for data from the two saccade conditions separately as well as together. All sets of data were examined with and without eye movement artifacts removed. We used Equal Error Rate as a measure of classification performance since this measure is independent of the percentage of misses (targets classified as non-targets) and false alarms (non-targets classified as targets). It reflects the proportion of errors for the case that the percentage of false alarms and misses are equal.

### 3 Results

Figure 1 shows the average target- and non-target ERPs and FRPs in the most relevant comparison conditions for electrode POz (where effects tended to be clearest, consistent with Brouwer et al., 2013), after removing the eye movement artefacts. The shaded areas represent p-values lower than 0.01 (dark) or 0.05 (light) as indicated by paired sample t-tests performed on every time sample of target and non-target values. As expected, ERPs (i.e. stimulus locked traces) and especially the target minus non-target difference trace are similar, regardless of eye movements being made. Also as expected, the difference peak of FRPs (i.e., fixation locked traces) occurs earlier when objects could be identified in the periphery (large object) compared to when not (small objects).



**Figure 1:** ERPs (left) and FRPs (right) for target (blue) and non-target (red) traces in different conditions, averaged over participants for electrode POz. Dotted curves indicate the difference wave.

Table 1 shows the results of the ERP and FRP classification analysis, separately for each condition, for the two saccade conditions grouped, and both with and without eye movement artefacts removed. Single trial classification is possible above chance in all cases as indicated by Wilcoxon signed rank tests on 0.5 (chance level) minus Equal Error Rates per participant (all eight p-values < 0.01). On average, the lowest error rates are achieved when eyes are static and EEG traces are locked to stimulus onset, while the highest error rates are observed for the most difficult case, i.e. traces

locked to fixation onset and unknown time of object identification (large and small stimuli mixed). Removing eye movement artefacts does not increase error rate, confirming the idea that distinction between targets and non-targets is based on brain rather than eye signals.

		Eye artefacts removed	Raw traces
Stim locked	Static large	0.29±0.06 [0.20-0.39]	0.29±0.06 [0.21-0.38]
Fix locked	Sacc large	0.35±0.08 [0.15-0.42]	0.35±0.06 [0.23-0.43]
	Sacc small	0.31±0.06 [0.17-0.38]	0.40±0.05 [0.30-0.48]
	Sacc small & large	0.37±0.06 [0.19-0.44]	0.41±0.04 [0.32-0.46]

**Table 1:** Equal Error Rates averaged across participants  $\pm$  the standard deviation. Between brackets are the minimum and maximum Equal Error Rates as observed across participants. Chance level is at 0.5.

## 4 Discussion

We found that we can still distinguish between target and non-target FRPs in the case that time of target identification relative to fixation is unknown. This is important when one is interested to use FRPs to monitor an observer's interest since for most visual exploration tasks, time of target identification is unknown. Equal Error Rates may seem high at first sight in comparison to other P300 studies. However, one has to note that these results reflect single trial classification and are based on data with a relatively large proportion of targets which is expected to result in relatively low P300 amplitudes.

Since attention and eye movements are intimately connected (Kowler, 2011) one might have expected that planning or inhibiting saccades would interfere with P300 latency and amplitude. However, we found that the static-large and saccade-large FRPs are very similar.

As expected, when objects are identified before fixation, the FRP P300 difference peak between targets and non-targets occurs earlier compared to identification after fixation. The FRPs in the saccade-small condition nicely reflect parafoveal and foveal visual processing. The first peak (after a peak that corresponds with a presaccadic spike potential) is similar for target and non-target traces, as well as for the traces in the saccadic-large condition. This peak probably reflects the detection of a relevant visual event in the periphery, i.e. stimulus onset. Then, there is a peak only for targets which is delayed in the case of small target size since targets can only be identified after an eye movement has been executed. The target trace of saccade-small may show us the P3a and P3b (Polich, 2007) more separated in time than is usually the case in experimental settings.

## References

- Brouwer, A-M., Reuderink, B., Vincent, J., van Gerven, M.A.J., van Erp, J.B.F. (2013). Distinguishing between target and nontarget fixations in a visual search task using fixation-related potentials. *Journal of Vision*, 13(3):17, 1–10.
- Jung, T.P., Makeig, S., Humphries, C., Lee, T.W., McKeown, M.J., Iragui, V., Sejnowski, T.J. (2000). Removing electroencephalographic artifacts by blind source separation. *Psychophysiology*, 37, 163-78.
- Kamienkowski, J.E., Ison, M.J., Quiroga, R.Q., Sigman, M. (2012). Fixation-related potentials in visual search: A combined EEG and eye tracking study. *Journal of Vision*, 12(7):4, 1-20.
- Kowler, E. (2011). Eye movements: the past 25 years. *Vision research*, 51(13), 1457-83.
- Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical neurophysiology*, 118(10), 2128-48.

# Acquiring control of auditory assistive communication devices

Sebastian Halder<sup>1</sup>, Ivo Käthner<sup>1</sup> and Andrea Kübler<sup>1</sup>

University of Würzburg, Würzburg, Germany  
sebastian.halder@uni-wuerzburg.de

## Abstract

Brain-computer interfaces (BCIs) often require intact gaze control. Users with brain-injuries or diseases that lead to a loss of gaze control need alternative BCI paradigms. Auditory P300 BCIs can provide such an alternative.

Previously we were able to show that training with an auditory P300 BCI based on natural stimuli and spatial cues leads to a substantial increase in information transfer rate (ITR) in a sample of healthy participants. Recently, we performed this training (five sessions) with a sample of N=5 motor impaired end-users.

The five end-users started with an average ITR of 0.17 bits/min. After the five training sessions were completed the average online ITR was 3.09 bits/min.

Three out of five end-users acquired control of the BCI system used in this study. This is the first time that severely motor impaired end-users achieved a level of control this high with a multi choice auditory P300 BCI.

## 1 Introduction

Brain-computer interfaces (BCIs) can provide a means of communication for severely paralyzed persons. Many potential users of BCIs lose gaze control due to progressive neurological disorders or acquired brain injuries. Thus, these users can benefit from BCIs that are controlled without visual stimulation.

The P300 event-related potential (ERP) is a component of the electroencephalogram (EEG) that can be used for controlling BCIs in the visual and also in the auditory domain. Controlling a P300 BCI requires a focus of attention on one of several stimuli that are presented in a random order. When selecting symbols with an auditory P300 BCI one can use this method to first select the row in the letter matrix and then the column. How easy this is for the user depends on the auditory stimuli. If they differ too much some stimuli may involuntarily attract the attention of the user. If they are very similar the user may not be able to differentiate the stimuli, especially if presentation times are short. Our first implementation of an auditory P300 BCI used spoken numbers as stimuli [1, 11]. These can be differentiated but require a long presentation time. Ideally, for a P300 BCI the sound can be differentiated with the onset of the stimulus. Other early implementations also needed long letter selection times and did not achieve high accuracies [9]. Our second implementation, inspired by [16], used artificial tones with additional spatial information [7]. This enabled faster presentation of the stimuli but they were difficult to discriminate for some users. In a refined version animal sounds were shown to be much better stimuli for auditory P300 BCIs [14]. They require only short presentation times but can still be easily discriminated. Additionally, using groups of alike sounding animals, stimuli sets that are similar but still differentiable can be created. Most recently, training was shown to lead to substantial improvements in communication speed with healthy participants [2]. So far multi-choice auditory BCIs have not been successfully used by a person with severe motor impairments [11, 17]. In this study we proceed to show for the first time that training enables severely motor impaired users to control an auditory P300 BCI.

## 2 Methods

### 2.1 Participants

Five users with severe motor impairments took part in the study. User 1 (male, 70 years old) was diagnosed with muscle dystrophy (ALS-functional rating scale revised (ALS-FRSR) score 40; range 0-48), User 2 (male, 56 years) was diagnosed with diffuse brain damage due to hypoxia (ALS-FRSR 18), User 3 (female, 45 years) was diagnosed with multiple sclerosis (ALS-FRSR 24), User 4 (female, 53 years) was diagnosed with ALS (ALS-FRSR 25) and User 5 (male, 73 years) was also diagnosed with ALS (ALS-FRSR 35).

### 2.2 Procedure

Each of the participants performed five separate sessions on different days with the auditory P300 BCI system. Each session consisted of spelling two words with five symbols (two times AGMSY) for calibration of the classifier (stepwise linear discriminant analysis; which was re-trained for every session) and five words with five symbols (chosen to make the stimuli needed for selection equally distributed) for feedback (VARIO, GRUEN, RUBIO, TUMBI, PHLEX). There was a pause of twelve seconds (to give the user enough time to between letter selections). Between rows and columns there was a pause of two seconds. Each stimulus was presented for 187.5 ms with 250 ms between stimuli. During training each stimulus was presented ten times. For the five words spelled with feedback stimulus repetitions was adjusted individually to the number of sequences needed to reach 70% plus three (to prevent ceiling effects).

The stimuli used in this study were a duck, bird, frog, gull and pigeon sound appearing to originate from the left, middle left, front, middle right and right side of the participant (via simulation of direction using stereo headphones). Each sound codes for a row and a column in a matrix with 5x5 symbols (the letters A-Y). The user attends to one of the sounds to select the row, then after a short pause to one of the sounds to select the column.

### 2.3 Data acquisition

EEG was recorded with a g.tec g.USBamp EEG amplifier with a bandpass from 0.1 to 30 Hz, notchfilter at 50 Hz and 256 Hz sampling rate. Sixteen g.gamma electrodes were positioned at FC3, FCz, FC4, C3, Cz, C4, CP5, CPz, CP6, P3, Pz, P4, Po5, Poz, Po6 and Oz.

BCI2000 was used for controlling all aspects of stimulus presentation, signal processing and data recording [15]. Recordings were made on a Hewlett-Packard ProBook 6460b with a dual-core CPU, 4 GB of RAM and a 64-bit Windows 7.

### 2.4 Signal processing

Stepwise linear discriminant analysis was used for online classification of the data. Information transfer rates (ITRs) were calculated with the formula suggested by Wolpaw [19].

## 3 Results

Average ITR increased from 0.17 bits/min to 3.09 bits/min for all users. Note that these values include the two users that did not learn to control the BCI (User 2 and User 3). The three users that learned to control the BCI increased their ITRs from 0.15 bits/min to 5.12 bits/min.

Symbol selection accuracies increased from 11.2% to 52.8% for all users and from 9.3% to 84% for the three users that learned to control the BCI.

## 4 Discussion

Three out of five users learned to control the BCI through training. None of the users were able to control the BCI in the first session. This shows that the training is not only a reliable method to increase ITR but mandatory for users with severe motor impairments.

Compared to the data from previous implementations of auditory P300 BCI spellers the extent of the performance increase becomes apparent. In the initial implementation by Furdea et al. [1] the ITR with healthy participants was on average 1.54 bits/min. The same implementation was evaluated with patients and none of the four participants achieved accuracies above 50%, which would be needed to transfer information [11]. A first implementation with spatial cues and artificial tones lead to an increase in ITR to 2.76 bits/min when using artificial tones [7] and to 4.23 bits/min when using animal sounds [14]. Finally, we were able to show that training leads to an ITR of 5.59 bits/min in healthy participants [2], an increase to 360% compared to 2009. In this paper we were able to show that motor impaired users can reach 3.09 bits/min (this includes two users that did not learn to control the BCI) which is an ITR unprecedented by motor impaired users with auditory P300 BCIs.

Unrelated to BCIs it has been shown before that training can increase the amplitude of ERPs [10, 18]. We believe that the increase in BCI performance can be attributed to similar effects. Users 2 and 3 did not learn to control the BCI. This may be related to the diagnosis: user 2 was diagnosed with diffuse brain damage due to hypoxia and user 3 with multiple sclerosis. Both may have an impact on the generation of the ERPs needed for BCI control. This was particularly evident for user 3 who did exhibit not any stimulus locked responses, also to auditory oddballs.

The data shown here underlines again that no BCI works for every user. This is especially the case with BCIs that have high attentional demands, such as auditory spellers, which also a large percentage of healthy controls fails to use [3]. Using binary choice auditory BCIs may be a viable alternative [5, 4, 6, 13, 12]. Nonetheless, it cannot be concluded from this data that the presented auditory BCI will enable users who have previously been unable to use an auditory BCI will now be able to do so [11, 8, 17]. The diversity of the epidemiologies that lead to such severe motor impairments that only communication with non-visual BCIs is possible will always require an individualized solution.

## 5 Acknowledgements

The study was funded by the European Community for research, technological development and demonstration activities under the Seventh Framework Programme (FP7, 2007-13), project grant agreement number 288566 (BackHome). This paper reflects only the authors views and funding agencies are not liable for any use that may be made of the information contained herein.

## References

- [1] A. Furdea, S. Halder, D. J. Krusienski, D. Bross, F. Nijboer, N. Birbaumer, and A. Kübler. An auditory oddball (P300) spelling system for brain-computer interfaces. *Psychophysiology*, 46(3):617–25, May 2009.

- [2] S. Halder, E. Baykara, C. Fioravanti, N. Simon, I. Käthner, E. Pasqualotto, S. Kleih, C. Ruf, N. Birbaumer, and A. Kübler. Training effects of multiple auditory BCI sessions. In *International BCI Meeting* <http://dx.doi.org/10.3217/978-3-85125-260-6-80>, 2013.
- [3] S. Halder, E. M. Hammer, S. C. Kleih, M. Bogdan, W. Rosenstiel, N. Birbaumer, and A. Kübler. Prediction of auditory and visual P300 brain-computer interface aptitude. *PLoS ONE*, 8(2):e53513, 02 2013.
- [4] S. Halder, M. Rea, R. Andreoni, F. Nijboer, E. M. Hammer, S. C. Kleih, N. Birbaumer, and A. Kübler. An auditory oddball brain-computer interface for binary choices. *Clin Neurophysiol*, 121(4):516–23, Apr 2010.
- [5] N. Hill, T. Lal, K. Bierig, N. Birbaumer, and B. Schölkopf. An auditory paradigm for brain-computer interfaces. In *Advances in Neural Information Processing Systems 17*, 2004.
- [6] N. J. Hill and B. Schölkopf. An online brain-computer interface based on shifting attention to concurrent streams of auditory stimuli. *J Neural Eng*, 9(2):026011, Apr 2012.
- [7] I. Käthner, C. Ruf, E. Pasqualotto, C. Braun, N. Birbaumer, and S. Halder. A portable auditory P300 brain-computer interface with directional cues. *Clin Neurophysiol*, 124(2):327–38, Feb 2013.
- [8] T. Kaufmann, E. M. Holz, and A. Kübler. Comparison of tactile, auditory and visual modality for brain-computer interface use: A case study with a patient in the locked-in state. *Frontiers in Neuroscience*, 7(129), 2013.
- [9] D. S. Klobassa, T. M. Vaughan, P. Brunner, N. E. Schwartz, J. R. Wolpaw, C. Neuper, and E. W. Sellers. Toward a high-throughput auditory P300-based brain-computer interface. *Clin Neurophysiol*, 120(7):1252–61, Jul 2009.
- [10] N. Kraus, T. McGee, T. D. Carrell, C. King, K. Tremblay, and T. Nicol. Central auditory system plasticity associated with speech discrimination training. *J Cogn Neurosci*, 7(1):25–32, 1995.
- [11] A. Kübler, A. Furdea, S. Halder, E. Hammer, F. Nijboer, and B. Kotchoubey. A brain-computer interface controlled auditory event-related potential (P300) spelling system for locked-in patients. *Ann N Y Acad Sci*, 1157:90–100, Mar 2009.
- [12] C. Pokorny, D. S. Klobassa, G. Pichler, H. Erlbeck, R. G. L. Real, A. Kübler, D. Lesenfants, D. Habbal, Q. Noirhomme, M. Riseti, D. Mattia, and G. R. Müller-Putz. The auditory p300-based single-switch brain-computer interface: paradigm transition from healthy subjects to minimally conscious patients. *Artif Intell Med*, 59(2):81–90, Oct 2013.
- [13] E. Ricci, S. Haider, T. Vaughan, and N. Hill. An improved auditory streaming BCI with voice stimuli. In *International BCI Meeting*, 2013 <http://dx.doi.org/10.3217/978-3-85125-260-6-1>.
- [14] C. Ruf, N. Simon, I. Käthner, E. Pasqualotto, N. Birbaumer, and S. Halder. Listen to the frog! An auditory P300 brain-computer interface with directional cues and natural sounds. In *International BCI Meeting* <http://dx.doi.org/10.3217/978-3-85125-260-6-157>, 2013.
- [15] G. Schalk, D. McFarland, T. Hinterberger, N. Birbaumer, and J. Wolpaw. BCI2000: a general-purpose brain-computer interface (BCI) system. *IEEE Trans Biomed Eng*, 51(6):1034–1043, 2004.
- [16] M. Schreuder, B. Blankertz, and M. Tangermann. A new auditory multi-class brain-computer interface paradigm: spatial hearing as an informative cue. *PLoS One*, 5(4):e9813, 2010.
- [17] M. Schreuder, A. Riccio, M. Riseti, S. Dähne, A. Ramsay, J. Williamson, D. Mattia, and M. Tangermann. User-centered design in brain-computer interfaces—a case study. *Artif Intell Med*, 59(2):71–80, Oct 2013.
- [18] K. Tremblay, N. Kraus, T. McGee, C. Ponton, and B. Otis. Central auditory plasticity: changes in the n1-p2 complex after speech-sound training. *Ear Hear*, 22(2):79–90, Apr 2001.
- [19] J. Wolpaw, N. Birbaumer, D. McFarland, G. Pfurtscheller, and T. Vaughan. Brain-computer interfaces for communication and control. *Clin Neurophysiol*, 113(6):767–791, 2002.

# Including temporal dynamics in single trial Motor Imagery classification using Empirical Mode Decomposition

Simon R.H. Davies<sup>1</sup> and Christopher J. James<sup>2</sup>

<sup>1</sup> Institute of Digital Healthcare, University of Warwick, Coventry, Warwickshire, U.K.  
davies\_s@wmg.warwick.ac.uk

<sup>2</sup> School of Engineering, University of Warwick, Coventry, Warwickshire, U.K.  
c.james@warwick.ac.uk

## Abstract

This paper describes the inclusion of temporal signal information in the process of Empirical Mode Decomposition (EMD) applied to the electroencephalogram (EEG). Key information in the classification of motor imagery EEG is very frequency centric but also contains strong temporal cues due to the nature of the experimental paradigm. Taking the temporal dynamics of the EEG into account, multiple snapshots of the signal in time are input into a multi-variate EMD process. Features were created from the processed signal using Common Spatial Patterns and these were input to a Support Vector Machine for classification. The results showed that the added temporal dynamics gave no major improvement to the sensitivity or specificity compared to regular EMD.

## 1 Introduction

Motor Imagery (MI) works by using the changes in brain activity over the sensorimotor cortex during imagined localised limb movements [5]. These changes include the suppression of the  $\mu$  rhythm on the contralateral hemisphere. The  $\mu$  rhythm is limited to a frequency band of 8-13 Hz, predominantly around 10 Hz. There is also weaker resonance activity in the beta bands around 20 Hz. During the commencement of MI there will be a large decrease in potential followed immediately by a large increase and then a return to pre-MI levels in the contralateral hemisphere as the limb being imagined moved. The lateral hemisphere will see a similar effect but to a much lesser degree, so called Event Related Desynchronisation/Event Related Synchronisation (ERD/ERS).

Empirical Mode Decomposition (EMD) is a process that can decompose non-stationary and non-linear signals into a group of frequency harmonics called Intrinsic Mode Functions (IMFs) and residual signal noise [2]. This makes it readily applicable to electroencephalogram (EEG) recordings. In this paper we work with the specific Brain-Computer Interface (BCI) [5] paradigm of MI. EMD has already been shown to be a useful signal processing technique for MI [1]. Here we introduce a new version of EMD that uses Taken's theorem [?] to incorporate temporal data into the EMD sifting process. This new algorithm will be tested in classifying MI trials, compared to standard EMD.

Any continuous single-spatial channel signal can be converted into a multi-temporal channel signal using Taken's theorem as described in reference [3]

$$x(t) = (x(t - \tau), x(t - 2\tau), \dots, x(t - (m - 1)\tau)) \in R,$$

where  $\tau$  is the lag and  $m$  is the embedding dimension. A similar method was used to apply Independent Component Analysis (ICA) to a single EEG channel [3]. The embedding dimension

needs to be greater than  $2D + 1$ , where  $D$  is the number of signal sources. From [3] we set the dimension size of  $m = 30$  and  $\tau = 1$ .

## 2 Methodology

### 2.1 Dataset

The EEG dataset contains 90 MI trials from 11 different subjects. Each trial is 8.2 seconds long, consisting of 4.1 seconds neutral activity, a stimulus cue that is randomly either left or right, and 4.1 seconds of sustained MI. The recordings were made with a 64-channel EEG running BCI2000 software with a sampling frequency of 160 Hz [6]. The electrodes selected for analysis were FC3, FC4, C5, C3, Cz, C4, C6, CP3, CP4, T7 and T8 according to the 10-10 system and were determined to cover the motor cortex. This is the same electrode selection as previous EMD studies [4].

### 2.2 Feature Extraction

EMD works by applying an iterative sifting process to the data, where the mean envelope of each channel is subtracted repeatedly until the data has been decomposed into a group of oscillations of varying frequencies that are symmetrical with respect to the  $x$  axis. However, Park et al [4] came up with an enhanced version that uses spatial information to help form the envelopes, causing each channel's IMFs to be equal in number to the rest and occupy the same frequency bands, making them consistent across all channels and far easier to analyse, this is called Multi-variate EMD (MEMD) and was shown to perform well compared to existing standard methods such as wavelet transforms and the application of a Butterworth filter.

MEMD works by calculating the mean envelope of  $n$ -directional signals, instead of just one signal in isolation, by projecting the signal along different directions in  $n$ -variate spaces and then averaging the projections to get the local mean. Low discrepancy Hammersley sequences are used to obtain quasi-uniform points on high dimensional spheres to form the projection vectors. In this case instead of applying MEMD to EEG data of multiple channels, we use dynamical embedding to create a multi-temporal signal and apply MEMD to that. The aim of this is to use temporal data to form more accurate and consistent IMFs from a signal channel using the temporal data, this is called Temporal MEMD (T-MEMD). Channels FC3, FC4, C5, C3, Cz, C4, C6, CP3, CP4, T7 and T8 according to the 10-10 system were selected for analysis.

The resulting IMFs for each trial, for either EMD or T-MEMD, are un-embedded and then selected or discarded based on two different methods, one knowledge-based and one performance-based. The knowledge-based method attempts to select relevant IMFs based on previous knowledge of their features. A Fast-Fourier Transform (FFT) is calculated for the duration of the signal that was recorded after the cue as it contains the MI. IMFs with at least 5% of their total power between 8-13 Hz (the  $\mu$  rhythm) are considered to have possible MI relevant data and are kept. The performance-based method uses a brute force method to identify the combination of IMFs that gives the best result per subject in terms of the number of trials correctly classified. As with this dataset the number of IMFs per channel never exceeds 13 and the IMFs containing useful information never exceeds 5 in number, it means all possible combinations of IMFs (maximum number of combinations: 2379) can be tested in a practical amount of time.

The selected IMFs are summed to form the processed signal and features are created for each trial using Common Spatial Patterns (CSP). CSPs calculate a set of spatial filters that

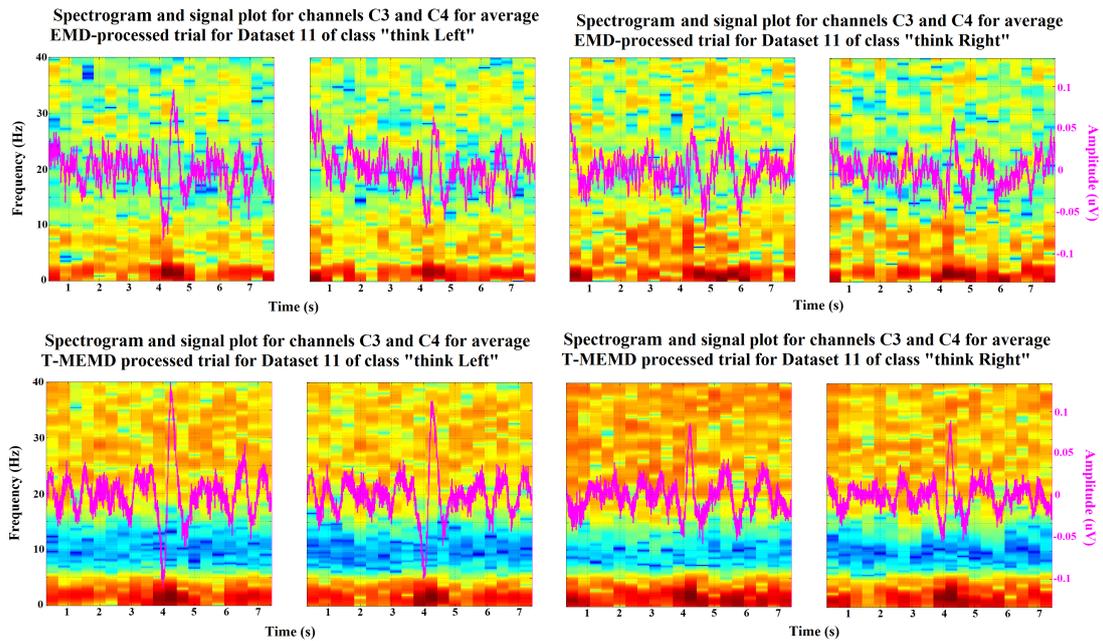


Figure 1: This figure shows the spectrograms for the average processed signal with the waveform itself superimposed in purple. The plots are for electrodes C3 and C4, for the classes “think Left” and “think Right” respectively for the EEG data of User 11. The first row is regular EMD and the bottom row is T-MEMD with the stimulus occurring exactly half way across the x-axis. In this case, T-MEMD has created clearer ERD/ERS peaks.

Method	Dataset 11		Avg., all channels			Avg., C3			Avg., C4		
	Sens.	Spec.	Sens.	Spec.	S.D.	Sens.	Spec.	S.D.	Sens.	Spec.	S.D. (%)
EMD knowledge-based	62.2	53.3	50.3	49.2	10.6	46.5	38.7	24.9	52.1	28.9	30.6
T-MEMD knowledge-based	35.6	37.8	53.2	49.0	11.6	37.5	44.2	28.2	45.6	30.1	29.3
EMD brute force	73.3	75.6	68.9	68.9	5.3	69.0	53.2	16.2	68.4	50.6	15.2
T-MEMD brute force	62.2	71.1	70.7	69.6	6.8	65.0	46.5	20.3	66.8	51.9	14.4

Table 1: The sensitivity, specificity and standard deviation of each EMD and T-MEMD method for a single dataset applied to all channels, the average sensitivity, specificity and standard deviation of each EMD and T-MEMD method for all datasets applied to all channels and applied to a single channel, C3 and C4 respectively.

maximise the variances of one class and minimises them in the other [7]. The features are then input to a Support Vector Machine (SVM). As there were only 90 trials per user it was decided to use the Leave One Out method (LOO). None of the CSP features or brute force classifier made use of the test trials.

### 3 Results & Analysis

EMD and T-MEMD both achieved very similar performance with the difference between the two being well within their respective standard deviations in every case. The brute force method

returned significantly higher performance than the knowledge based method for both decomposition methods. Whilst the IMFs chosen with the brute force method were not consistent between users or decomposition methods, they did focus on the frequency bands known to contain MI information that were identified in the introduction (Figure 1). A single channel analysis of the high-performing brute force method was carried out to further see if embedding temporal data added anything to the EMD process. As CSPs need multiple channels to function the variance of each processed trial was used as the input of the SVM.

As Table 1 shows, there is negligible difference in performance indicating that the added temporal dynamics do not contain any new information. In part, this may be due to the fact that the underlying processes for MI affect all channels, and whilst ERD/ERS and the  $\mu$  rhythm are expected to be stronger on one side of the brain versus the other, the changes still occur simultaneously for both hemispheres. This implies that the information content for MI is inherent in the lateralisation of the changes - i.e. spatially. In [3] adding temporal dynamics to ICA had a greater impact than with EMD because ICA contains zero temporal information, whilst EMD still uses the data laid out in chronological order. EMD can also only discard background noise if it is unstructured due to its criteria for identifying IMFs.

## 4 Conclusion

Ultimately the added temporal dynamics did not significantly improve the classification performance. However it might still have some effect on performance if applied to an EEG signal with significant temporally independent features, which is not the case in MI. Another way to enhance the method could be to incorporate both spatial and temporal information by decomposing all channels in their multi-temporal form simultaneously using MEMD.

## References

- [1] S.R.H. Davies and C.J. James. Novel use of empirical mode decomposition in single-trial classification of motor imagery for use in brain-computer interfaces. In *Engineering in Medicine and Biology Soc. (EMBC), 2013 35th Annual Intl. Conf. of the IEEE*, pages 5610–5613, July 2013.
- [2] Norden E. Huang, Zheng Shen, Steven R. Long, Manli C. Wu, Hsing H. Shih, Quanan Zheng, Nai-Chyuan Yen, Chi Chao Tung, and Henry H. Liu. The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 454(1971):903–995, 1998.
- [3] C. James and D. Lowe. Using dynamical embedding to isolate seizure components in the ictal EEG. *Science, Measurement and Technology, IEE Proceedings -*, 147(6):315–320, Nov 2000.
- [4] Cheolsoo Park, D. Looney, N. ur Rehman, A. Ahrabian, and D.P. Mandic. Classification of motor imagery bci using multivariate empirical mode decomposition. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 21(1):10–22, Jan 2013.
- [5] G. Pfurtscheller, C. Brunner, A. Schlgl, and F.H. Lopes da Silva. Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks. *NeuroImage*, 31(1):153 – 159, 2006.
- [6] Gerwin Schalk, Dennis J. Mcfarl, Thilo Hinterberger, Niels Birbaumer, and Jonathan R. Wolpaw. Bci2000: A general-purpose brain-computer interface (bci) system. *IEEE Transactions on Biomedical Engineering*, 51, 2004.
- [7] Zhisong Wang, Alexander Maier, Nikos K. Logothetis, and Hualou Liang. Single-trial classification of bistable perception by integrating empirical mode decomposition, clustering, and support vector machine. *EURASIP J. Adv. Sig. Proc.*, 2008.

# Strategies for adaptive motor imagery classification using error-related potential derived labels have unique risk profiles

Tim Zeyl<sup>1,2</sup> and Tom Chau<sup>1,2</sup>

<sup>1</sup> University of Toronto, Toronto, Ontario, Canada

[tim.zeyl@utoronto.ca](mailto:tim.zeyl@utoronto.ca), [tom.chau@utoronto.ca](mailto:tom.chau@utoronto.ca)

<sup>2</sup> Bloorview Research Institute, Toronto, Ontario, Canada

## Abstract

Signals measured during brain-computer interface (BCI) tasks are nonstationary, which can lead to classification errors. The error-related potential (ErrP) has been proposed for BCI error detection as well as partially-supervised classifier adaptation. We discuss how the ErrP can be incorporated into several adaptive classification methods, and the unique sensitivity that these methods have to misidentification of the ErrP. We find that the risk associated with these methods varies as a function of false positive rate for a realistic ErrP detector receiver operating characteristic and we recommend individualized biasing of the ErrP detector to account for these effects.

## 1 Introduction

Adaptive classification of brain-computer interfaces (BCIs) can be used to address the inherent non-stationarity of EEG data during mental tasks such as motor imagery. Class labels typically used for classifier adaptation are not available in a true BCI session, thus unsupervised adaptation has been employed as an alternative to supervised adaptation [9]. However, unsupervised methods may not be suitable when the nonstationarity affects relative class positions.

A potential compromise is to use error-related potentials (ErrPs) to generate labels for partially-supervised adaptation. Adaptation using such uncertain labels has been proposed by [5, 10], and [7] adapted an SVM classifier in the context of a code-modulated visual evoked potential speller, with benefits for participants. However, the validity of the labels depend on the accuracy of the ErrP detector, with some correct trials inevitably being interpreted as incorrect, and some incorrect trials being interpreted as correct. There is limited discussion on the risk associated with adaptation using incorrect labels and what methods are most suitable in this situation. To this end, we evaluate two adaptation methods across several ErrP detection accuracies. We assume stationary performance of the ErrP detector after the results of [2], but note that risks would increase with a nonstationary assumption.

The performance of motor imagery classifiers is dependent on choice of frequency band, and the authors of [8] showed that the most discriminative frequency changes from session to session. Therefore we build a classifier based on the filter bank common spatial pattern (FBCSP) [1] framework that uses a majority weighted vote from linear discriminant analysis (LDA)-based classifiers in each FBCSP band. This framework allows us to adapt both the ensemble weights and the base LDA classifiers separately or concurrently to either re-weight individual frequency components or change the decision boundaries in each band. In this study we evaluate the consequences of incrementally adapting these two components of the classifier at several accuracies of the ErrP detector.

## 2 Methods

Data are taken from dataset IIB of the IVth BCI competition [4], which is comprised of 9 participants performing 5 sessions of left and right hand motor imagery, with 120-160 trials in each session. EEG is recorded with three bipolar electrodes above C3, Cz, C4 at 250 Hz with a 50 Hz notch filter. We epoch the data and apply a zero-phase filter bank of 4 Hz pass-band non-overlapping filters from 4 – 40 Hz. In each band we extract CSP features from a 2s window starting 1.5s post cue. The first and last CSP features are used, which results in two features from each filter band.

We train LDA classifiers on the features from each band and combine their decisions using a weighted majority vote. Data from session 1 are used for training and the remaining sessions are used for testing. Initial weights are determined using a 10-fold cross-validation evaluating Cohen’s kappa from each base classifier and normalizing the results. Then, the base LDA classifiers are retrained using all the data from session 1.

To simulate ErrPs with realistic detection accuracies, we estimated the values of the receiver operating characteristic (ROC) curves from the online ErrP detectors found in [6] and evaluated several points along the median curve, which appear in Table 1. Using these false positive rates ( $\alpha$ ) and true positive rates, we simulated ErrPs for each trial in sessions 2-4 as in [10].

We adapt both the base LDA classifiers (denoted ‘BaseAdapt’) and the weights of the ensemble (denoted ‘Reweight’) incrementally after every trial. Our estimate of the true class,  $\hat{y} \in \{0, 1\}$ , is derived from the classifier’s output on the current trial,  $\tilde{y} \in \{0, 1\}$ , and our belief that an error occurred,  $\tilde{E} \in \{0, 1\}$  (i.e. we detect an ErrP). Thus,  $\hat{y}$  is incorrect whenever  $\tilde{E}$  is incorrect. BaseAdapt updates the class means and global covariance according to the supervised LDA classifier in [9]; the ‘Pmean-Gcov’ unsupervised adaptive classifier from this group is included for comparison. The learning parameter,  $\eta$ , is set to 0.05. The Reweight strategy uses  $\tilde{E}$  to implement a variant of the dynamic weighted majority of [3] that decrements the weight of incorrect experts by a factor of 0.9 and increments correct experts by a factor of 1.1. No experts are removed or created. We simulate the performance of these two adaptation strategies, as well as their combination (denoted ‘Hybrid’), 50 times with independent randomly generated  $\tilde{E}$  on each repetition.

False Positive Rate ( $\alpha$ )	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40
True Positive Rate	0.70	0.79	0.87	0.93	0.93	0.95	0.95	0.98

Table 1: False positive and corresponding true positive rates of the ErrP detector.

## 3 Results

The average classification accuracies across all 4 evaluation sessions and all participants are shown in Fig. 1a for each adaptation strategy. We see that across most values of  $\alpha$ , on average the semi-supervised adaptation improved the accuracies over the case of no adaptation (horizontal line). Errorbar length indicates 2 standard deviations ( $2\sigma$ ) of all 50 simulation repetitions averaged across participant and session.  $\sigma$ , shown as a function of  $\alpha$  in Fig. 1b, gives a quantitative measure of the risk associated with each adaptive method. With increasing  $\sigma$ , there is increasing risk that the adaptation could be detrimental instead of beneficial. In general, Fig. 1 indicates that the accuracy of the BaseAdapt method decreases with increasing  $\alpha$ , while  $\sigma$  increases. Even at low  $\alpha$ , it performs no better than its unsupervised counterpart.

The Reweight method has optimal accuracy and minimal  $\sigma$  at  $\alpha \approx 0.2$ . The Hybrid method obtains the best optimal average classification accuracies, but it also has the highest  $\sigma$  across most values of  $\alpha$ ; this is likely because it combines variability from both methods.

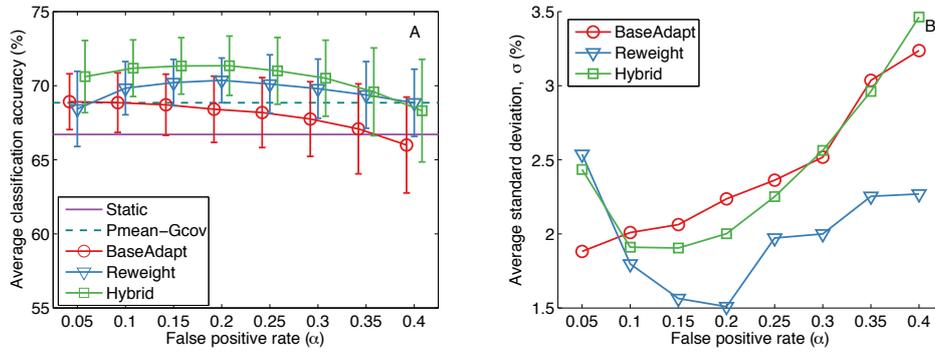


Figure 1: a. Average classification accuracies (across session and participants) as a function of false positive rate (jittered from true value for visibility). b. Standard deviation of 50 simulations as a function of false positive rate, averaged across sessions and participants.

Choosing an appropriate  $\alpha$  for the ErrP detector involves maximizing average performance while minimizing risk, or  $\sigma$  equivalently. The lower quartile of the 50 simulation repetitions is a convenient measure for quantifying these two goals. The best  $\alpha$  for each participant and each adaptation type was chosen as the one that gave the optimum lower quartile of simulation repetitions averaged across sessions. Fig. 2a compares the average classification accuracy and  $\sigma$  at the best  $\alpha$  for each adaptation type and for each participant. This figure indicates highly variable performance of the adaptation strategies across participants. For some participants adaptation is not beneficial, while for others one adaptation type clearly outperforms the other. This participant dependency is also seen in the best  $\alpha$  for each method (shown in Fig. 2b). The BaseAdapt method tends to prescribe lower  $\alpha$  than do the Reweight and Hybrid methods where best  $\alpha$  varies more with participant.

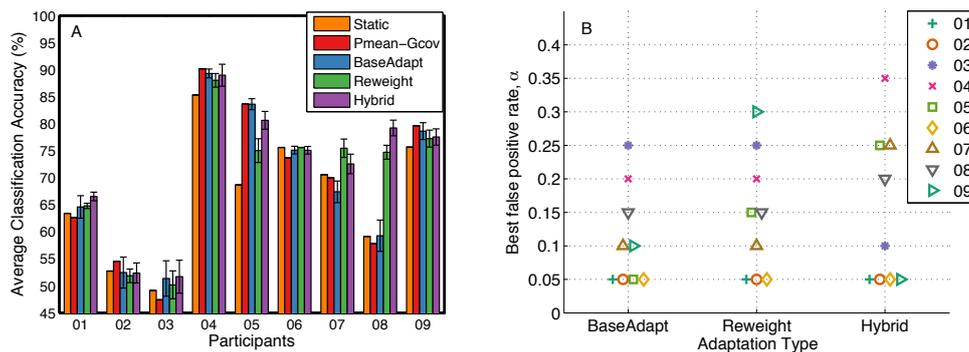


Figure 2: a. Barplot of average classification accuracies (across session) at the best false positive rate for each participant. Error bars indicate  $\sigma$  of 50 simulations averaged across session. b. Best false positive rate chosen for each participant (refer to legend) for each adaptation method.

## 4 Discussion

We find that the variance, and thus the risk associated with employing these adaptation methods changes across  $\alpha$  in a manner unique to each method. The adaptation of the LDA base classifiers has lower risk at low  $\alpha$ , while the re-weighting method has lower risk when  $\alpha$  and true positive rate are balanced. The hybrid method combines the risk from both methods such that, although the average performance is the highest, it also has a high variance for most  $\alpha$ .

LDA may be more sensitive to high  $\alpha$  because when a false positive occurs, the error is due to a sample more likely to be further from the adapted class mean than in the case of a false negative. This drives classes closer together, which can lead to erratic movement of the class boundary as discussed in [10]. The Reweight method may not favor a single error type as such.

We found that, averaged across all participants, re-weighting the ensemble had the lowest risk. However, we find that the adaptation method with the best performance is unique to individuals, so that participant specific adaptation methods may be required. A few participants appear to benefit much more from the Reweight method compared to the BaseAdapt method. This may be due to particularly strong shifts in the most discriminative frequency for these participants. However, re-weighting may not be helpful for individuals with stationary discriminative frequency. These findings also suggest that ErrP detectors should be biased to particular  $\alpha$  depending on the adaptation method chosen and the individual.

Future work should attempt to reduce the risk associated with adaptation using potentially uncertain labels. This may be achieved by combining the detection of an ErrP with the confidence of the BCI task classifier, or using a fuzzy classifier where weights of training samples are determined by the ErrP strength on individual trials.

## References

- [1] K. K. Ang, Z. Y. Chin, H. Zhang, and C. Guan. Filter bank common spatial pattern (FBCSP) in brain-computer interface. In *IEEE IJCNN*, pages 2390–2397. IEEE, 2008.
- [2] P. W. Ferrez and J. R. Millán. Error-related EEG potentials generated during simulated brain-computer interaction. *IEEE Transactions on Biomedical Engineering*, 55(3):923 – 929, 2008.
- [3] J. Z. Kolter and M. A. Maloof. Dynamic weighted majority: An ensemble method for drifting concepts. *J Mach Learn Res*, 8(12):2755 – 2790, 2007.
- [4] R. Leeb, F. Lee, C. Keinrath, R. Scherer, H. Bischof, and G. Pfurtscheller. Brain-computer communication: motivation, aim, and impact of exploring a virtual apartment. *IEEE T Neur Sys Reh*, 15(4):473–482, 2007.
- [5] A. Llera, M. van Gerven, V. Gómez, O. Jensen, and H. Kappen. On the use of interaction error potentials for adaptive brain computer interfaces. *Neural Networks*, 24(10):1120–1127, 2011.
- [6] N. M. Schmidt, B. Blankertz, and M. S. Treder. Online detection of error-related potentials boosts the performance of mental typewriters. *BMC Neurosci*, 13:19, Feb. 2012.
- [7] M. Spüler, W. Rosenstiel, and M. Bogdan. Online adaptation of a c-VEP brain-computer interface(BCI) based on error-related potentials and unsupervised learning. *PLOS ONE*, 7(12):e51077, Dec. 2012.
- [8] K. P. Thomas, C. Guan, C. T. Lau, A. P. Vinod, and K. K. Ang. Adaptive tracking of discriminative frequency components in electroencephalograms for a robust brain-computer interface. *J Neural Eng*, 8(3):036007, 2011.
- [9] C. Vidaurre, M. Kawanabe, P. von Büna, B. Blankertz, and K. R. Müller. Toward unsupervised adaptation of LDA for brain-computer interfaces. *IEEE T Biomed Eng*, 58(3):587–597, 2011.
- [10] T. J. Zeyl and T. Chau. A case study of linear classifiers adapted using imperfect labels derived from human event-related potentials. *Pattern Recogn Lett*, 37(0):54–62, Feb. 2014.

# A P300 BCI for e-inclusion, cognitive rehabilitation and smart home control

I. Käthner<sup>1\*</sup>, J. Daly<sup>2\*</sup>, S. Halder<sup>1</sup>, J. Räderscheidt<sup>1</sup>, E. Armstrong<sup>2</sup>,  
S. Dauwalder<sup>3</sup>, C. Hintermüller<sup>4</sup>, A. Espinosa<sup>4</sup>, E. Vargiu<sup>3</sup>, A. Pinegger<sup>5</sup>,  
J. Faller<sup>5</sup>, S. C. Wriessnegger<sup>5</sup>, F. Miralles<sup>3</sup>, H. Lowish<sup>6</sup>, D. Markey<sup>6</sup>,  
G. R. Müller-Putz<sup>5</sup>, S. Martin<sup>2</sup>, A. Kübler<sup>1</sup>

<sup>1</sup>Institute of Psychology, University of Würzburg, Würzburg, Germany, <sup>2</sup>The Cedar Foundation, Belfast, United Kingdom, <sup>3</sup>Barcelona Digital Technology Center, Barcelona, Spain, <sup>4</sup>G.tec Medical Engineering GmbH, Schiedlberg, Austria, <sup>5</sup>Institute for Knowledge Discovery, Graz University of Technology, Graz, Austria, <sup>6</sup>Telehealth Solutions, Medvivo Group, Wiltshire, United Kingdom  
ivo.kaethner@uni-wuerzburg.de, j.daly@cedar-foundation.org

## Abstract

We implemented an easy-to-use P300 BCI system that allows users to control a variety of applications for communication, creative expression, training of cognitive abilities and environmental control. In this paper we present an evaluation of the following four applications: a speller, two games that can be used for cognitive rehabilitation or entertainment, twitter (via web browser) and a webcam. All fourteen healthy participants had control over the BCI and reached high accuracies (>85%). The results of the evaluation informed the development of the next prototype. With a user-centered approach we aim to further improve the prototype and ultimately provide end users with a multifunctional system that can be used as assistive technology in a home environment.

## 1 Introduction

One of the main goals of the project “BackHome” is to advance existing BCI systems such that they can be used as assistive technology in a home environment and operated without the need of technical skills. Based on previous work (Faller, et al., 2013), we implemented a BCI system that provides end users with a simple graphical interface and several applications for communication, creativity, cognitive rehabilitation, entertainment and smart home control. Thereby we aim at facilitating activities of daily living for persons with severe paralysis and improving the users’ social integration. From the beginning of the project, we engaged with potential end-users to gather users’ needs and feedback on the functionalities of a first prototype. This user-centered approach (Kübler, Holz, & Kaufmann, 2013) informed the development of the current BCI prototype that allows users to control the following applications: a multimedia player and web browser (see Halder et al., in preparation), a speller, Brain Painting for creative expression (Münßinger et al., 2010), a TV, a webcam and two games (see Vargiu et al., this conference). The first game (find-a-category) aims at improving the semantic and reasoning skills of the participant. A certain category is presented and users have to identify the picture that matches the category among four alternatives (see Figure 1A). The level of difficulty is varied by the level of abstraction needed to identify the correct picture. The second game (pairs) is an implementation of the memory-cards game, in which users have to find a pair of matching cards. In the beginning, all cards are depicted face down. Each card is assigned a

---

\* These authors contributed equally

number and users can uncover the cards by choosing the corresponding number in the control matrix. The game is finished once all pairs are found. The level of difficulty is varied by the number of cards and the level of complexity of the depicted pictures/symbols shown. At the end of each game, the results are shown in a numerical and graphical display. The cognitive rehabilitation games can be assigned to specific users via a web application (therapist station). The sessions can then be loaded by the users by choosing the “session” symbol from the main games menu. All applications are controlled with a P300 system based on control matrices that were first proposed for a spelling system (Farwell & Donchin, 1988).

## 2 Methods

### 2.1 Framework of Evaluation

The evaluation of the BCI Prototype followed the framework for evaluation suggested by Zickler et al. (2011). It is based on ISO recommendations for the evaluation of usability (ISO-9241-210) and requires the estimation of effectiveness and efficiency of the system and of user satisfaction. Effectiveness is operationalized as the accuracy of symbol/command selections in percent correct. Efficiency is estimated through the calculation of the information transfer rate using the formula suggested by Wolpaw et al. (2002) and the NASA-Task Load Index (NASA-TLX) as an indicator of the subjective workload. Satisfaction was assessed with an extended version of the Quebec User Evaluation of Satisfaction with assistive Technology (extended QUEST 2.0, Demers, Weiss-Lambrou & Ska, 2002), a visual analogue scale (VAS, ranging from 0=not at all satisfied to 10=completely satisfied) and a custom usability questionnaire concerning the design of the system.

### 2.2 Participants, Data Acquisition and Procedure

Fourteen volunteers participated in the study (9 females,  $M=28.1$  years  $\pm$  8.6, range: 21-46,  $N=7$  have never used a BCI before). They signed informed consent prior to participation in the study that was conducted in accordance with the guidelines of the Declaration of Helsinki. The EEG was recorded with 8 active electrodes positioned at Fz, Cz, P3, P4, PO7, POz, PO8 and Oz. It was amplified using a g.USBamp amplifier (g.tec GmbH, Austria). Participants were seated about one meter from the monitor that displayed the symbol matrices used to control the applications. The laptop was placed next to the monitor during usage of the twitter and gaming applications to display the browser or the windows displaying the games (see Figure 1).

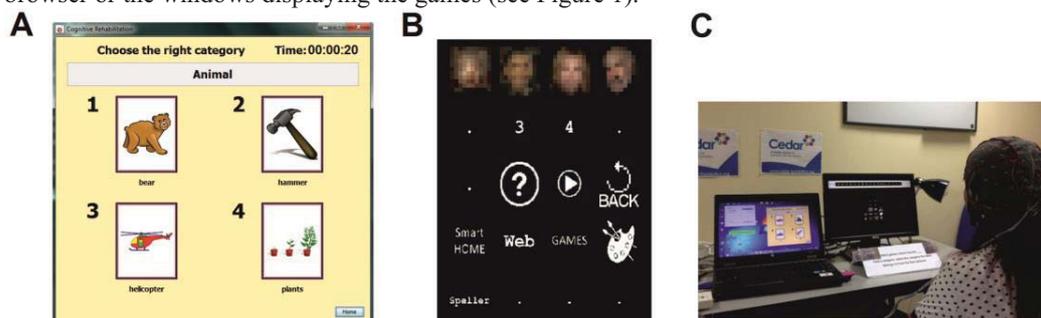


Figure 1. Exemplary question of the category game (A) and matrix used to control the application (B). Users choose a number to indicate the right answer. By focusing on the particular icon, they can also switch to a different application. In Figure B the first row is overlaid with pictures of faces (in this publication they are pixelated due to copyright). Figure C shows a user in front of the two screens during BCI operation.

Users were asked to complete an experimental protocol that included writing ten letters with the spelling application, playing the two games (minimum of 5 and 8 selections required), writing and posting a short twitter message (18 selections) and controlling a webcam (3 selections). To navigate the menus and switch between applications 7 selections were required. To make a certain selection users were asked to focus on the symbol they should select and silently count the number of times it was highlighted. During the screening run, used to acquire data for the classifier, users had to select 5 letters from a 3x3 matrix. Each row and column flashed for a total 15 times during the screening run. After feature weights had been determined, the number of flashes was adaptive and varied for every selection (dynamic stopping). Rows and columns were flashed with pictures of famous faces to improve classification accuracies (Kaufmann et al., 2013). During all tasks (except spelling), users could correct mistakes. If, with three attempts, users were unable to select a certain symbol, it was selected by the investigator, who then indicated the next selection. After completing the BCI tasks users were asked to fill in the questionnaires.

### 3 Results

All users had control over the BCI and reached high accuracies (>85%) with four of the five applications. The average accuracies for the different tasks are depicted in Figure 2. Accuracy for menu navigations was 72%. Due to the varying number of stimulus repetitions needed for one selection and different matrix sizes, bitrates can only be estimated for individual tasks and participants. Highest bitrates were achieved with the web browser (Twitter), with bitrates ranging from 11-30 bits/minute.

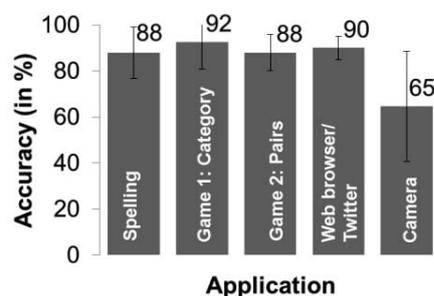


Figure 2. Average accuracies for the five applications.

It is more important to note that 98% of symbols could be selected, with a maximum of three attempts for one selection. Subjective workload was indicated with the NASA-TLX as moderate (score of  $51.84 \pm 14.34$ ; highest possible score is 100). The overall satisfaction score assessed with a VAS was  $7.48 \pm 1.37$  (range 6-10). The global score of the QUEST 2.0 was also high ( $4.33 \pm 0.59$ ), indicating that users were “quite satisfied” with the system. The score included items concerning the dimensions, weight, adjustment, safety, comfort, ease of use, effectiveness of the system and professional services (instructions). The average score for the BCI specific items reliability, speed, learnability and aesthetic design was lower, but still high ( $4.05 \pm 0.68$ ). While users criticized the aesthetic design of the electrode cap (“not suited for everyday life”, “clinical/scientific look”, “necessity of gel”, “restricting cables”), they did not name it as one of the most important aspects. Instead, they stressed the importance of effectiveness, ease of use and learnability. Generally, users were satisfied with the high selection accuracies, but stated that the reliability of the software should be improved (it had to be restarted occasionally). Further, some users

suggested to improve the design of the software that they criticized as “pixelated”. Several users remarked positively on the many different applications that the system offers.

## 4 Discussion

With the proposed P300 BCI system study participants could switch between and control multiple applications ranging from entertainment to smart home control. The web browser allows to gain access to internet based communication (e.g. via social media platforms). The evaluation revealed some aspects of the system that could be improved to increase its usability. This includes improvements of the hardware, such as less conspicuous EEG equipment as well as improvements of the software (e.g. design of the user interface and stability). Some symbols were particularly difficult to select (e.g. for the camera task), probably because they differed in size. We will continue to evaluate the system with end users and, in an iterative process, implement new features to increase user friendliness of the system to ultimately reach our goal of implementing a system that can be employed in daily life to reduce social exclusion and increase independence of the end users.

## Acknowledgments

This work was funded by the FP7 EU research project BackHome (No. 288566). This paper reflects only the authors’ views and funding agencies are not liable for any use that may be made of the information contained herein.

## References

- Demers, L., Weiss-Lambrou, R., & Ska, B. (2002). Quebec user evaluation of satisfaction with assistive technology (QUEST 2.0): An overview and recent progress. *Technology and Disability, 14*, pp. 101-105.
- Faller, J., Torrellas, S., Costa, U., Opisso, E., Fernández, J. M., Kapeller, C., . . . Müller-Putz, G. (2013). Autonomy and Social Inclusion for the Severely Disabled: The BrainAble Prototype. *Proc. of the 5th Int. BCI Meeting 2013*, (pp. 372-3). Asilomar, CA, USA.
- Farwell, L., & Donchin, E. (1988). Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroen Clin Neuro, 70(6)*, pp. 510-23.
- Kaufmann, T., Schulz, S., Köblitz, A., Renner, G., Wessig, C., & Kübler, A. (2013). Face stimuli effectively prevent brain-computer interface inefficiency in patients with neurodegenerative disease, 124(5). *Clin Neurophysiol, 893-900*.
- Kübler, A., Holz, E., & Kaufmann, T. (2013). A user centred approach for bringing BCI controlled applications to end-users. In R. Fazel-Rezai, *Brain-Computer Interface Systems - Recent Progress and Future Prospects* (pp. 1-19). Rijeka, Croatia: InTech.
- Münßinger, J. I., Halder, S., Kleih, S. C., Furdea, A., Raco, V., Hösle, A., & Kübler, A. (2010). Brain Painting: First Evaluation of anew Brain-Computer Interface Application with ALS-Patients and Healthy Volunteers. *Frontiers in Neuroscience, 4:182*, doi:10.3389/fnins.2010.00182.
- Wolpaw, J., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain-computer interfaces for communication and control. *Clin Neurophysiol, 113(6)*, pp. 767-91.
- Zickler, C., Riccio, A., Leotta, F., Hillian-Tress, S., Halder, S., Holz, E., . . . Kübler, A. (2011). A Brain-Computer Interface as Input Channel for a Standard Assistive Technology Software. *Clinical EEG and Neuroscience, 42(2)*, pp. 236-244.

# Towards Neurofeedback Training of Associative Brain Areas for Stroke Rehabilitation

Ozan Özdenizci<sup>1</sup>, Timm Meyer<sup>2</sup>, Müjdat Çetin<sup>1</sup> and Moritz Grosse-Wentrup<sup>2</sup>

<sup>1</sup> Sabancı University, Faculty of Engineering and Natural Sciences, Istanbul, Turkey  
oozdenizci@sabanciuniv.edu, mcetin@sabanciuniv.edu

<sup>2</sup> Max Planck Institute for Intelligent Systems, Dept. Empirical Inference, Tübingen, Germany  
tmeyer@tuebingen.mpg.de, moritzgw@tuebingen.mpg.de

## Abstract

We propose to extend the current focus of BCI-based stroke rehabilitation beyond sensorimotor-rhythms to also include associative brain areas. In particular, we argue that neurofeedback training of brain rhythms that signal a state-of-mind beneficial for motor-learning is likely to enhance post-stroke motor rehabilitation. We propose an adaptive neurofeedback paradigm for this purpose and demonstrate its viability on EEG data recorded with five healthy subjects.

## 1 Introduction

While initially conceived as communication devices for severely paralyzed patients, brain-computer interfaces (BCIs) have recently been considered in the context of post-stroke motor rehabilitation [1]. Here, BCIs are employed to synchronize movement intent, as decoded by a BCI from sensori-motor rhythms, with congruent haptic feedback, e.g., as delivered by an exoskeleton [2]. This artificial re-establishment of the sensorimotor feedback loop has been shown to support modulation of sensorimotor-rhythms [3] and result in enhanced post-stroke recovery [4]. In this paper, we argue that, motivated by the impressive results achieved to date, the current focus should be extended beyond sensori-motor- to also include associative brain areas. This argument is based on empirical evidence that a variety of brain rhythms beyond those in primary sensorimotor areas are related to the extent of post-stroke impairment. For instance, the global ratio of  $\delta$ - to  $\alpha$ -power of the brain's electromagnetic field has been found to correlate with the extent of post-stroke disability and predict subsequent recovery [5]. We interpret such abnormal activation patterns as disturbances of the balance of large scale cortical networks [6], and argue that re-establishing their natural balance by means of BCI-based neurofeedback is likely to support the brain in post-stroke recovery.

In order to turn this hypothesis into a viable stroke rehabilitation protocol, several interrelated problems need to be addressed. Firstly, we need to identify which large scale networks are involved in sensori-motor learning in healthy control subjects. Next, we have to investigate how activation patterns of and between these networks are disturbed in stroke patients, and elucidate how these disturbances relate to post-stroke recovery. And thirdly, we need to train patients via neurofeedback to establish activation patterns of large scale networks that are associated with good post-stroke recovery.

In the present work, we address the viability of the last issue. In a recent publication, we have identified EEG correlates of motor-learning performance [7]. In particular, we have provided evidence that parieto-occipital  $\alpha$ -power during rest as well as during movement preparation predicts performance in a subsequent reaching movement. This led us to hypothesize that parieto-occipital  $\alpha$ -power reflects activity in a cortical network that is tuned by the brain to optimize motor-learning performance. Here, we propose an adaptive neurofeedback training

scheme to modulate parieto-occipital  $\alpha$ -power in a way that we predict will enhance motor-learning performance. Based on experimental data from five healthy subjects we argue that, firstly, subjects are able to modulate parieto-occipital  $\alpha$ -power by means of neurofeedback, and, secondly, that this modulation is not task-specific but extends into the inter-trial resting-periods. Taken together, these results indicate the feasibility of using neurofeedback to support subjects in generating brain activation patterns that are associated with good motor-learning performance. This constitutes an important building block towards a stroke therapy based on neurofeedback training of associative brain areas.

## 2 Methods

### 2.1 Experimental Data

Five healthy subjects were recruited for this study, each of which completed two training sessions. Each session lasted one hour with one week interval between sessions. Prior to the training sessions all participants gave their informed consent after the training procedure was explained to them in accordance with guidelines set by the Max Planck Society. During the training sessions, a 120-channel EEG was recorded at 1 kHz sampling rate, using active EEG electrodes and a QuickAmp amplifier (BrainProducts, Gilching, Germany). Electrodes were placed according to the 10-20 system, with Cz as the initial reference electrode. All data was re-referenced to common average reference.

### 2.2 Adaptive Online Feedback

Each training session included one resting-state baseline and eight training blocks with one minute breaks in between. For the resting-state baseline, subjects were instructed to relax for five minutes with eyes open, looking at a fixation cross displayed centrally on a screen approximately 1.5 m in front of the subjects. This resting-state data was used to calibrate the feedback for the subsequent training session. Specifically, the 120-dimensional raw EEG data  $\mathbf{x}[t]$  was first spatially filtered by a filter  $\mathbf{w}$  to obtain a one-dimensional signal  $y[t] = \mathbf{w}^T \mathbf{x}[t]$ . The spatial filter  $\mathbf{w}$  was taken from a group-wise independent component analysis (ICA) of a previous study [7], in which the corresponding independent component's  $\alpha$ -power was found to predict performance in a motor-learning task. Figure 1a displays the source localization results obtained for this independent component in parieto-occipital areas (adapted from [7]). Next, log-bandpower of the spatially filtered signal was computed in the  $\alpha$ -band (8-14 Hz), using a FFT in conjunction with a Hanning window, in a sliding window of 2 s with a window step size of 100 ms. This signal processing pipeline was used to compute the mean and standard

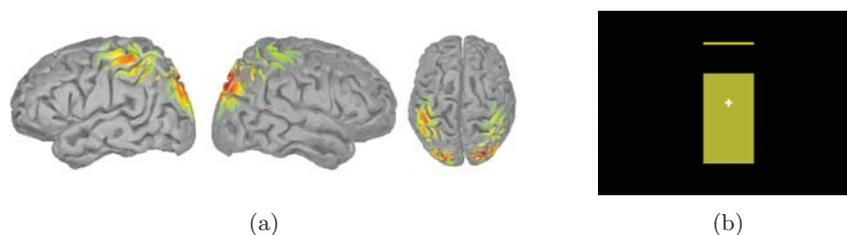


Figure 1: a) Parieto-occipital region used for neurofeedback [7]. b) Visual stimulus.

deviation of resting-state  $\alpha$ -power.

In each training block, the same signal processing pipeline was used to provide subjects with visual feedback on parieto-occipital  $\alpha$ -power. The current estimate of parieto-occipital  $\alpha$ -power was displayed as the vertical height of a rectangular visual stimulus on the screen (Figure 1b), which was updated at 25 Hz. The bottom of the rectangle corresponded to the mean log-bandpower of the resting-state baseline. One training block consisted of 15 trials. In each trial, subjects were given the objective to increase the vertical height of the rectangle to reach an adaptively determined target height level marked by an upper bar on the screen. Subjects had to learn to up-regulate the presented brain activity with eyes open and keep the activity at that level or above for a cumulative time of two seconds. No instructions were given to the subjects on how to achieve this goal. If subjects succeeded in this task in a 15 second trial, the rectangle turned green. Otherwise, the next trial began after a short resting-period, with a randomly determined length between 4.5 to 5.5 seconds. For the first training block, the target distance was set to one standard deviation of the log-bandpowers in the resting-state baseline. Depending on the performance of the subject, the target distance changed. It increased by 0.2 standard deviations of the resting-state baseline phase in the next block, if the success rate of the previous block over 15 trials was higher than 70%. The target distance decreased by the same amount if the success rate was lower than 60%. This adaptive approach was implemented to balance any potential frustration or negligence of the subject on the task [8]. During the training sessions, presentation of the stimulus and real-time data processing was performed with the BCI2000 software [9] and its extension BCPy2000 [10].

### 2.3 Offline Data Analysis

Following the last training session, the data of all subjects and sessions was visually inspected for contamination by ocular artifacts. The data of one session of one subject was discarded, as it was heavily contaminated by eye blinks. To assess the overall effect of neurofeedback training, we then pooled the data of all subjects and sessions. Specifically, we computed the correlation of trial-number, ranging from one to 120, with parieto-occipital  $\alpha$ -power, averaged across subjects and sessions, in each trial. In this way, we estimated the linear trend in parieto-occipital  $\alpha$ -power across a training session. We also computed this correlation for  $\alpha$ -power at each individual recording channel in order to investigate the topography of  $\alpha$ -power modulation. Finally, we computed the same correlation only using the EEG data of the resting-periods prior to each trial.

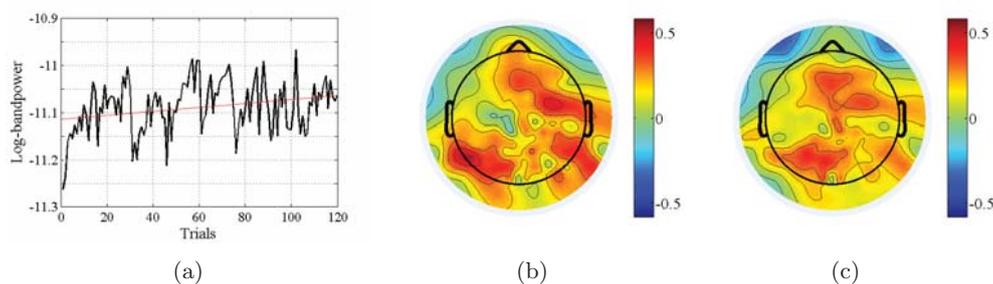


Figure 2: a) Grand average of trial  $\alpha$ -log-bandpower values (red is the linear fit). b) Topography of grand average log-bandpower correlations with trial numbers within a session of 120 trials in  $\alpha$ -band (for trial phases). c) Same topography in (b) obtained for rest phases.

### 3 Results

Parieto-occipital  $\alpha$ -power displays a positively-sloped linear trend within a training session of 120 trials (Figure 2a) with a correlation coefficient of  $\rho = 0.26$ . A permutation test on the temporal order of trials with  $10^4$  permutations rejected the null-hypothesis of zero correlation with  $p = 0.002$  ( $N = 120$ ). Correlation values between  $\alpha$ -power and trial number at each individual electrode are shown in Figure 2b and 2c for trial- and pre-trial data, respectively. These results indicate that subjects learned to modulate parietal  $\alpha$ -power and that this self-regulation extended beyond the actual feedback phase into the inter-trial resting phases.

### 4 Discussion

In this pilot study, we presented an adaptive neurofeedback training paradigm that enables subjects to up-regulate parieto-occipital  $\alpha$ -power. Importantly, we could provide evidence that up-regulation of  $\alpha$ -power was sustained in the inter-trial periods, indicating that subjects learned to induce a stable state-of-mind that we predict will be beneficial for motor-learning [7]. In future work, this training paradigm will be carried out by stroke patients prior to a rehabilitation session. We hypothesize that this will be beneficial for motor-learning and hence support post-stroke motor recovery.

### References

- [1] M. Grosse-Wentrup, D. Mattia, and K. Oweiss. Using brain-computer interfaces to induce neural plasticity and restore function. *Journal of Neural Engineering*, 8(2):025004, 2011.
- [2] M. Gomez-Rodriguez, M. Grosse-Wentrup, J. Hill, A. Gharabaghi, B. Schölkopf, and J. Peters. Towards brain-robot interfaces in stroke rehabilitation. In *IEEE International Conference on Rehabilitation Robotics (ICORR)*, pages 1–6. IEEE, 2011.
- [3] M. Gomez-Rodriguez, J. Peters, J. Hill, B. Schölkopf, A. Gharabaghi, and M. Grosse-Wentrup. Closing the sensorimotor loop: haptic feedback facilitates decoding of motor imagery. *Journal of Neural Engineering*, 8(3):036005, 2011.
- [4] A. Ramos-Murguialday et al. Brain-machine interface in chronic stroke rehabilitation: A controlled study. *Annals of Neurology*, 74(1):100–108, 2013.
- [5] S. P. Finnigan, M. Walsh, S. E. Rose, and J. B. Chalk. Quantitative EEG indices of sub-acute ischaemic stroke correlate with clinical outcomes. *Clinical Neurophysiology*, 118(11):2525–2532, 2007.
- [6] S. L. Bressler and V. Menon. Large-scale brain networks in cognition: emerging methods and principles. *Trends in Cognitive Sciences*, 14(6):277–290, 2010.
- [7] T. Meyer, J. Peters, T. O. Zander, B. Schölkopf, and M. Grosse-Wentrup. Predicting motor learning performance from electroencephalographic data. *Journal of NeuroEngineering and Rehabilitation*, 11:24, 2014.
- [8] S. Othmer, S. F. Othmer, and D. A. Kaiser. EEG biofeedback: An emerging model for its global efficacy. In J. Evans and A. Abarbanel, editors, *Introduction to Quantitative EEG and Neurofeedback*, pages 243–310. Academic Press, New York, 1999.
- [9] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw. BCI2000: A general-purpose brain-computer interface (BCI) system. *IEEE Transactions on Biomedical Engineering*, 51(6):1034–1043, 2004.
- [10] BCPy2000, <http://bci2000.org/downloads/BCPy2000/>.

# Towards a passive brain computer interface for improving memory

Eunho Noh<sup>1</sup>, Matthew V. Mollison<sup>2</sup>, Grit Herzmann<sup>3</sup>, Tim Curran<sup>2</sup> and Virginia R. de Sa<sup>4</sup>

<sup>1</sup> Department of Electrical and Computer Engineering, University of California, San Diego,

<sup>2</sup> Department of Psychology and Neuroscience, University of Colorado Boulder

<sup>3</sup> Department of Psychology, The College of Wooster

<sup>4</sup> Department of Cognitive Science, University of California, San Diego

## Abstract

We propose a passive BCI system based on our previous results on deciphering neural correlates of memory from single-trial EEG. Our system will measure the brain activity of a user, and infer the user's preparedness for learning to present study items at estimated optimal times. The system will also monitor the brain activity during learning/encoding to assess whether the encoding process was successful or not. These studied item will be presented in the future with an appropriate lag depending on the predicted level of encoding to strengthen retention. The system will also extract information related to the user's confidence during re-presentation of the item to assess the level of reinstatement. Items with low reinstatement will be presented again to ensure encoding. Spacing models can be incorporated with the system to determine these lags for optimal retention. Other systems that monitor

## 1 Introduction

Faced with rising classroom sizes and the higher incidence of learning difficulties, such as ADD, we need a tool that can improve a user's ability to remember study material. Such a tool may also be useful for addressing age-related memory decline. We propose a system that tailors instruction to each individual's brain dynamics. The system will use neural activity reflecting memory encoding and retrieval to choose optimal presentation times and intervals to improve memory. This system can be considered a passive BCI where the system monitors the brain in real-time to extract information related to memory encoding/retrieval.

## 2 Types of memory related neural processes

We review findings on the neural correlates of long-term memory which can be used for an EEG-based passive BCI system and give a summary of the classification results applied to single-trial EEG data to identify neural signatures reflecting memory encoding and retrieval.

### 2.1 Encoding Success

There are significant differences in the spectral and temporal patterns between remembered and forgotten trials during study item presentation which are known as subsequent memory effects (SMEs). The difference in event-related potential (ERP) between subsequently remembered and forgotten items is also known as the Dm (or difference due to memory) effect [13, 12]. It is observed around 400 to 800ms after study item presentation. Brain oscillations during encoding

also distinguish between subsequently remembered and forgotten items (see [6] for a review). Power decreases for the remembered items typically occur in the alpha (7-12 Hz) and low beta (12-19 Hz) bands [7, 5] of the EEG signal.

We used single-trial classification to successfully distinguish between remembered and forgotten pictures using the temporal and spectral information in the EEG during stimulus presentation [10]. By combining the information from multiple time windows, the *encoding success classifiers* achieved an overall accuracy of 58 % across all 18 subjects. The classifiers gave accuracies significantly over chance ( $p < 0.05$ ) for 9 subjects [9].

## 2.2 Encoding preparedness

EEG, MEG, and ECoG studies have shown that spectral differences in brain activity immediately preceding study item presentation also show significantly different patterns for the subsequently remembered and forgotten trials. This difference is observed in multiple frequency bands ranging from theta (4-7 Hz) [4] to the high beta (19-30 Hz) [3] bands.

These good/bad (for subsequent memory retention) brain states were successfully identified from the spectral information in the EEG signal between -300 to 0 ms before stimulus presentation of pictures in [10]. The classifier combined information from 9 overlapping subbands of the EEG signal (4-7, 6-10, 7-12, 10-15, 12-19, 15-25, 19-30, 25-35, and 30-40 Hz). The overall classification accuracy of the *encoding preparedness classifiers* across 18 subjects was 57.2 % where 9 subjects showed significantly over chance results.

Because study items were presented at a fixed interval without awareness of the subjects' brain state, we used the 10 % of presentations with the highest and lowest classifier confidence as analogs for the best and worst sets in a real-time system. The rate of remembered items gave a 50 % improvement when the study items were presented at the best times compared to the worst (59.6 % items remembered during *good* brain states vs. 40 % during *bad* brain states). For the 10 subjects with the highest overall classification accuracy, the rate of remembered items was a 74 % improvement when the study items were presented at the best times compared to the worst (63.4 % items remembered during *good* brain states vs. 36.5 % during *bad*).

## 2.3 Confidence at retrieval

The parietal ERP old/new effect is the difference in ERP between correctly rejected new and correctly detected old trials during memory retrieval. The old trials show a positive-going deflection compared to the new trials in the left parietal channels between 500-800 ms after stimulus presentation [14]. The frontal ERP old/new effect also distinguishes between correctly rejected new and correctly detected old trials. The ERPs show a more negative peak for less familiar trials in the frontal channels between 300-500 ms after stimulus presentation [2].

We used the EEG data during the recognition phase of a memory experiment and found that it is possible to distinguish unsuccessfully from successfully retrieved studied items with 58.4% accuracy [11] where 20 out of 34 subjects gave significantly over chance results. The likelihood of remembering a study item for trials with the 10 % highest and lowest classifier outputs were 0.8 and 0.45 respectively suggesting that the *confidence at retrieval classifier* outputs reflect the level of retrieval strength during the test phase.

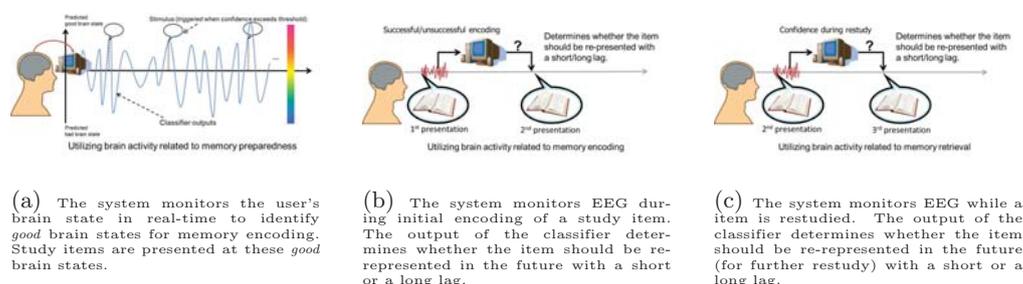


Figure 1: The three components of the proposed system

### 3 A system for improving memory

The three classifiers above (*encoding preparedness*, *encoding success*, *confidence at retrieval*) can be incorporated into a passive BCI system to assist users in improving memory. A common real-world learning task is paired-associate learning (e.g. word-meaning lists for students or vocabulary learning for a second language). In this section, we propose a system that incorporates the above classifiers to improve paired-associate learning. The system will present study items to the users following a continuous recognition paradigm where there is no distinction between study and test phases. Items may occur multiple times in the list for restudy.

The *encoding preparedness classifier* monitors the user's brain state in real-time to identify optimal brain states for study item presentation. When the system identifies a near-optimal time for presentation, it gives a scheduled study item (e.g. a target word in a foreign language) to the user. The user responds with either *New* for items they have not seen before, *Don't know* for items for which they do not remember the associations, or with their guess for the previously given association (e.g. the same word in English). After the response, the user receives feedback from the system with the correct association pair. When the user is finished studying/restudying the given association, the next target item appears on the screen at the next detected near-optimal time. With a first-time user, the classifiers can be initialized using training data from other users or the study items can be given at fixed time points. These initial classifiers can be *adapted* to the user in an online manner at the second (*test*) presentation of each item (when the behavioral results are obtained) with the stored EEG data from memory encoding/retrieval.

For a new study item, the lag for re-presentation is determined by the *encoding success classifier*. If the likelihood of encoding success is high (low) for a given item, the item may be re-presented with a long (short) lag. For a restudy item, the lag for re-presentation (for further restudy) is determined by the subject's response and the *confidence at retrieval classifier* output. Incorrect user responses will be re-presented with a short lag; for correct responses, the *confidence at retrieval classifier* output will determine the lag duration. Spacing models can also be incorporated to specifically determine these lags for optimal retention [1].

One disadvantage of the proposed system is that the throughput (rate of learned material presentation) may be lower than a conventional tutoring system due to the waiting period for good brain states. We can overcome this issue by using the good brain states to present the most critical or difficult<sup>1</sup> items. At intervening times, the remaining study material can be

<sup>1</sup>Difficult items may be identified from average hit rates from multiple users or based on previous behavioral/classification results for the specific user.

given to the user with a typical presentation rate [8]. As a long-term goal, we plan to explore the effects of long term use of our system. We hypothesize that the implicit neurofeedback users would get from being presented study items only when they are ready could help them remain in a receptive brain state more often.

## 4 Conclusion

In this paper, we proposed a passive BCI system for assisting memory formation and retention. Our system has three components: 1) it will infer the user's preparedness for learning to present study items at estimated optimal times; 2) it will monitor the brain activity during learning/encoding to assess whether the encoding process was successful or not; 3) it will also extract information related to the user's confidence during re-presentation of the item to assess the level of reinstatement. The incorporation of spacing models was also discussed to specifically determine these lags for optimal retention. Other systems that monitor study performance or user state could also be integrated with the proposed system.

This research was funded by NSF grants CBET-0756828, SBE-0542013 and SMA-1041755, NIH Grant MH64812, and a KIBM Innovative Research Grant.

## References

- [1] N. J. Cepeda, N. Coburn, D. Rohrer, J. Wixted, M. Mozer, and H. Pashler. Optimizing distributed practice: theoretical analysis and practical implications. *Exp Psychol*, 56:236–246, 2009.
- [2] T. Curran. Brain potentials of recollection and familiarity. *Memory & Cognition*, 28:923–38, 2000.
- [3] J. Fell, E. Ludowig, B. P. Staresina, T. Wagner, T. Kranz, C. E. Elger, and N. Axmacher. Medial temporal theta/alpha power enhancement precedes successful memory encoding: evidence based on intracranial EEG. *Journal of Neuroscience*, 31(14):5392–5397, 2011.
- [4] S. Guderian, B. H. Schott, A. Richardson-Klavehn, and E. Duzel. Medial temporal theta state before an event predicts episodic encoding success in humans. *PNAS*, 106:5365–70, 2009.
- [5] S. Hanslmayr, B. Spitzer, and K.-H. Bäuml. Brain oscillations dissociate between semantic and nonsemantic encoding of episodic memories. *Cerebral Cortex*, 19(7):1631–1640, 2009.
- [6] S. Hanslmayr and T. Staudigl. How brain oscillations form memories - a processing based perspective on oscillatory subsequent memory effects. *NeuroImage*, 85(2):648–655, 2013.
- [7] W. Klimesch, H. Schimke, M. Doppelmayr, B. Ripper, J. Schwaiger, and G. Pfurtscheller. Event-related desynchronization (ERD) and the Dm effect: Does alpha desynchronization during encoding predict later recall performance? *Intl. J. Psychophysiology*, 24:47–60, 1996.
- [8] R. Mayer. *Applying the Science of Learning*. Pearson Education, Incorporated, 2010.
- [9] G. Müller-Putz, R. Scherer, C. Brunner, R. Leeb, and G. Pfurtscheller. Better than random? A closer look on BCI results. *International Journal of Bioelectromagnetism*, 10(1):52–55, 2008.
- [10] E. Noh, G. Herzmann, T. Curran, and V. R. de Sa. Using single-trial EEG to predict and analyze subsequent memory. *NeuroImage*, 84:712–723, 2014.
- [11] E. Noh, M. V. Mollison, T. Curran, and V. R. de Sa. Single-trial identification of failed memory retrieval. submitted to the 48th Asilomar Conference on Signals, Systems, and Computers, 2014.
- [12] K. A. Paller, M. Kutas, and A. R. Mayes. Neural correlates of encoding in an incidental learning paradigm. *Electroencephalography and Clinical Neurophysiology*, 67(4):360–371, 1987.
- [13] T. F. Sanquist, J. W. Rohrbaugh, K. Sydulko, and D. B. Lindsley. Electrocortical signs of levels of processing: perceptual analysis and recognition memory. *Psychophysiology*, 17(6):568–576, 1980.
- [14] E. L. Wilding and M. D. Rugg. An event-related potential study of recognition memory with and without retrieval of source. *International Journal of Psychophysiology*, 119:889–905, 1996.

# Airborne Ultrasonic Tactile Display Brain-computer Interface Paradigm

Katsuhiko Hamada<sup>1,\*</sup>, Hiromu Mori<sup>2,†</sup>, Hiroyuki Shinoda<sup>1</sup>  
and Tomasz M. Rutkowski<sup>2,3,‡,§</sup>

<sup>1</sup> The University of Tokyo, Tokyo, Japan

<sup>2</sup> Life Science Center of TARA, University of Tsukuba, Tsukuba, Japan

<sup>3</sup> RIKEN Brain Science Institute, Wako-shi, Japan  
tomek@bci-lab.info & http://bci-lab.info/

## Abstract

We study the extent to which contact-less and airborne ultrasonic tactile display (AUTD) stimuli delivered to the palms of a user can serve as a platform for a brain computer interface (BCI) paradigm. Six palm positions are used to evoke combined somatosensory brain responses, in order to define a novel contact-less tactile BCI. A comparison is made with classical attached vibrotactile transducers. Experiment results of subjects performing online experiments validate the novel BCI paradigm.

## 1 Introduction

State-of-the-art brain computer interfaces (BCIs) are typically based on mental visual or auditory paradigms, as well as motor imagery paradigms, which require extensive user training and good eyesight or hearing. In recent years, alternative solutions have been proposed to make use of a tactile modality [1, 2, 3] to enhance brain-computer interfacing efficiency. The concept reported in this paper further extends the brain's somatosensory channel by the application of a contact-less stimulus generated with an airborne ultrasonic tactile display (AUTD) [4]. The rationale behind the use of the AUTD is that, due to its contact-less nature, it allows for a more hygienic application, avoiding the occurrence of skin ulcers (bedsores) in patients in a locked-in state (LIS). This paper reports very encouraging results with AUTD-based BCI (autdBCI) in comparison with the classical paradigm of vibrotactile transducer-based somatosensory stimulus (vtBCI) attached to the user's palms [3]. The rest of the paper is organized as follows. The next section introduces the materials and methods used in the study. The results obtained in online experiments with 13 healthy BCI users are then discussed. Finally, conclusions are formulated and directions for future research are outlined.

## 2 Materials and Methods

Thirteen male volunteer BCI users participated in the experiments. The users' mean age was 28.54, with a standard deviation of 7.96 years. The experiments were performed at the Life Science Center of TARA, University of Tsukuba, at the University of Tokyo and at RIKEN Brain Science Institute, Japan. The online (real-time) EEG autdBCI and vtBCI paradigm

---

\*K. Hamada is currently with DENSO Corporation, Japan.

†The author was supported in part by the Strategic Information and Communications R&D Promotion Program no. 121803027 of The Ministry of Internal Affairs and Communication in Japan.

‡The corresponding author.

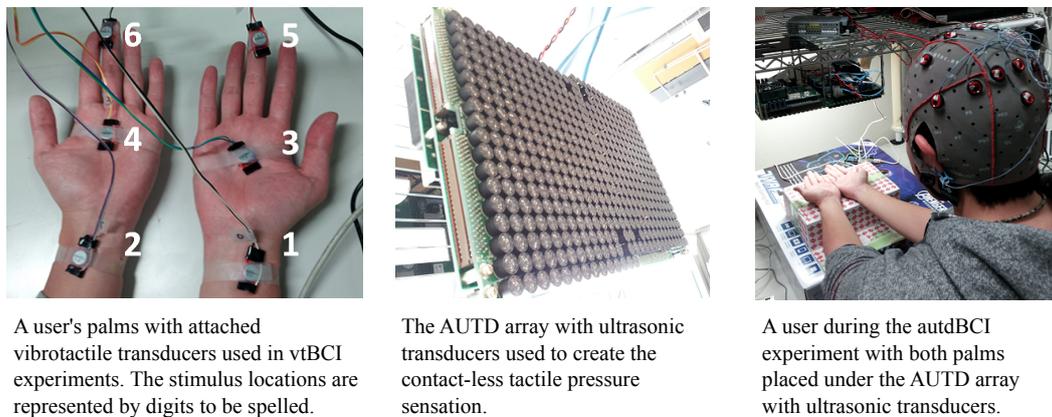


Figure 1: Tactile stimulus set-ups for autdBCI and vtBCI experiments with palms stimulated from the top in the both cases.

experiments were conducted in accordance with the *WMA Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects* and the procedures were approved and designed in agreement with the ethical committee guidelines of the Faculty of Engineering, Information and Systems at University of Tsukuba, Japan. The AUTD stimulus generator produced vibrotactile contact-less stimulation of the human skin via the air using focused ultrasound [4, 5]. The effect was achieved by generating an ultrasonic radiation static force produced by intense sound pressure amplitude (a nonlinear acoustic phenomenon). The radiation pressure deformed the surface of the skin on the palms, creating a tactile sensation. An array of ultrasonic transducers mounted on the AUTD created the focussed radiation pressure at an arbitrary focal point by choosing a phase shift of each transducer appropriately (the so-called phased array technique). Modulated radiation pressure created a sensation of tactile vibration similar to the one delivered by classical vibrotactile transducers attached to the user's palms. The AUTD device developed by the authors [4, 5] adhered to ultrasonic medical standards and did not exceed the permitted skin absorption levels (approximately 40 times below the limits). The effective vibrotactile sensation was set to 50 Hz [5] to match with tactile skin receptors and notch filters for power line interference rejection. As a reference, in the second vtBCI experiment, contact vibrotactile stimuli were also applied to locations on the users' palms via the transducers HIHX09C005-8. Each transducer in the experiments was set to emit a square acoustic frequency wave at 50 Hz, which was delivered from the ARDUINO micro-controller board with a custom battery-driven and isolated power amplifier and software developed in-house and managed from a *MAX 6* visual programming environment. The two experiment set-ups above are presented in Figure 1. Two types of experiments were performed with the volunteer users. Psychophysical experiments with foot-button-press responses were conducted in order to test uniform stimulus difficulty levels from response accuracy and time measurements. The subsequent tactile oddball online BCI EEG experiments evaluated the autdBCI paradigm efficiency and allowed for a comparison with the classical skin contact-based vtBCI reference. In both the above experiment protocols, the users were instructed to spell sequences of six digits representing the stimulated positions on their palms. The training instructions were presented visually by means of the *BCI2000* [6] and *MAX 6* programs with the numbers 1 – 6 representing the palm locations as depicted in the left panel of Figure 1. The EEG signals were captured with an EEG amplifier system g.USBamp by g.tec Medical Engineering GmbH,

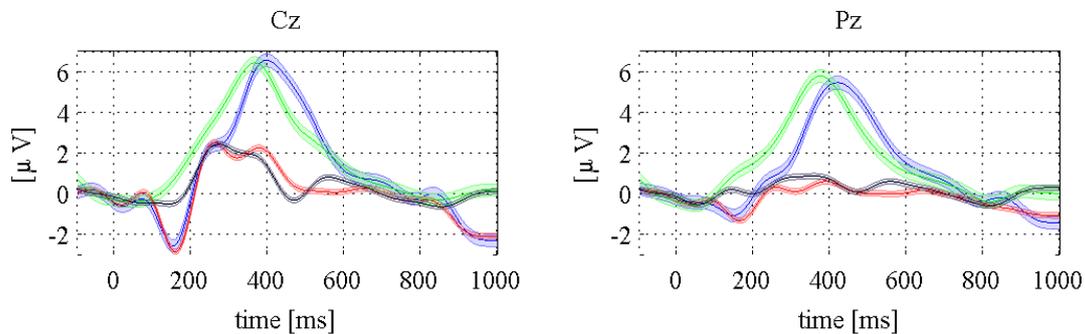


Figure 2: The autdBCI (blue - targets; red - non-targets) and vtBCI (green - targets; black - non-targets) grand mean averaged ERP responses, together with standard error bars. Due to limited space, of the 16 measured, only electrodes *Cz* and *Pz* are presented.

Austria, using 16 active electrodes. The electrodes were attached to the head locations: *Cz*, *Pz*, *P3*, *P4*, *C3*, *C4*, *CP5*, *CP6*, *P1*, *P2*, *POz*, *C1*, *C2*, *FC1*, *FC2*, and *FCz*, as in the 10/10 extended international system. The ground electrode was attached to the *FPz* position, and the reference was attached to the left earlobe. No electromagnetic interference was observed from the AUTD or vibrotactile transducers operating with frequencies notch-filtered together with power line interference from the EEG. The EEG signals captured were processed online with a BCI2000-based application [6], using a stepwise linear discriminant analysis (SWLDA) classifier [7] with features drawn from 0 – 800 ms ERP intervals decimated by a factor of 20. The stimulus length and inter-stimulus-interval were set to 400 ms, and the number of epochs to average was set to 15. The EEG recording sampling rate was set at 512 Hz, and the high and low pass filters were set at 0.1 Hz and 60 Hz, respectively. The notch filter to remove power line interference was set for a rejection band of 48 ~ 52 Hz. Each user performed three experiment runs (randomized 90 targets and 450 non-targets each). As a feedback the spelled numbers (palm position assigned digits as in Figure 1) were shown on a display to the user.

### 3 Results and Conclusions

The averaged evoked responses to targets and non-targets are depicted together with standard error bars in Figure 2. The BCI six digit sequences spelling accuracy analyses for both the experiments for the various averaging options are summarized in Figure 3. The chance level was of 16.6%. The mean accuracies for 15-trial averaged ERPs were 63.8% and 69.4% for autdBCI and vtBCI, respectively. The maximum accuracies were 78.3% and 84.6% respectively. The differences were not significant, supporting the concept of AUTD-based tactile stimulus usability for BCI. However, a single trial classification offline analysis of the collected responses resulted in mean accuracies of 83.0% for autdBCI and 53.8% for vtBCI, leading to a possible 19.2 bit/min and 7.9 bit/min, respectively. In the case of the autdBCI, only a single user's results were bordering on the level of chance, and four subjects attained 100% (10 trials averaging). On average, lower accuracies were obtained with the classical vtBCI, with which three users bordered on the level of chance, and only one user scored 100% accuracy level in SWLDA-classified averaged responses. This case study demonstrates results obtained with a novel six-command-based autdBCI paradigm. We compared the results with classical vibrotactile transducer stimuli already generated. The experiment results obtained in this study

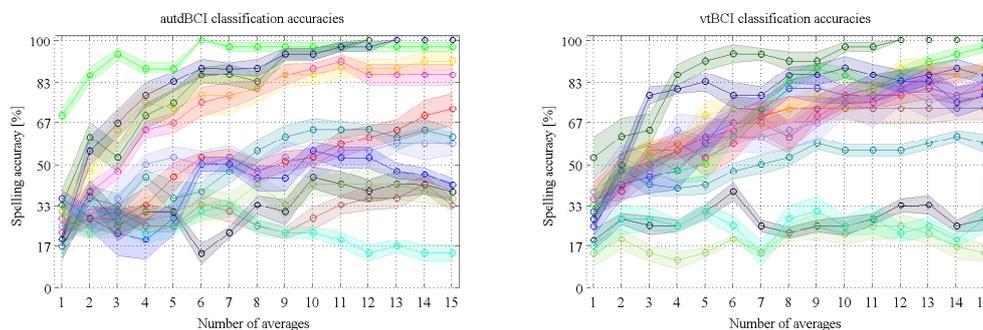


Figure 3: Averaged autdBCI and vtBCI six digits spelling accuracy results colour coded separately for each user, with standard error bars depicted.

confirm the validity of the contact-less autdBCI for interactive applications and the possibility to further improve the results with the utilization of single trial-based linear classification. The EEG experiment with the paradigm confirms that contact-less (airborne) tactile stimuli can be used to create six command-based interfaces. The results presented offer a step forward in the development of novel neurotechnology applications. Due to the still not very high interfacing rate achieved by users in the case of online BCI, the current paradigm obviously requires improvement and modification. These requirements determine the major lines of study for future research. However, even in its current form, the proposed autdBCI can be regarded as a practical solution for LIS patients (locked into their own bodies despite often intact cognitive functioning), who cannot use vision or auditory-based interfaces due to sensory or other disabilities.

## References

- [1] G.R. Muller-Putz, R. Scherer, C. Neuper, and G. Pfurtscheller. Steady-state somatosensory evoked potentials: suitable brain signals for brain-computer interfaces? *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 14(1):30–37, March 2006.
- [2] A-M. Brouwer and J.B F Van Erp. A tactile P300 brain-computer interface. *Frontiers in Neuroscience*, 4(19), 2010.
- [3] H. Mori, Y. Matsumoto, S. Makino, V. Kryssanov, and T.M. Rutkowski. Vibrotactile stimulus frequency optimization for the haptic BCI prototype. In *Proceedings of The 6th International Conference on Soft Computing and Intelligent Systems, and The 13th International Symposium on Advanced Intelligent Systems*, pages 2150–2153, Kobe, Japan, November 20-24, 2012.
- [4] T. Iwamoto, M. Tatzono, and H. Shinoda. Non-contact method for producing tactile sensation using airborne ultrasound. In Manuel Ferre, editor, *Haptics: Perception, Devices and Scenarios*, volume 5024 of *Lecture Notes in Computer Science*, pages 504–513. Springer Berlin Heidelberg, 2008.
- [5] K. Hamada. Brain-computer interface using airborne ultrasound tactile display. Master thesis, The University of Tokyo, Tokyo, Japan, March 2014. In Japanese.
- [6] G. Schalk and J. Mellinger. *A Practical Guide to Brain-Computer Interfacing with BCI2000*. Springer-Verlag London Limited, 2010.
- [7] D.J. Krusienski, E.W. Sellers, F. Cabestaing, S. Bayouth, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw. A comparison of classification techniques for the P300 speller. *Journal of Neural Engineering*, 3(4):299, 2006.

# The Evaluation of a Brain Computer Interface System with Acquired Brain Injury End Users

Jean Daly<sup>2</sup>, Elaine Armstrong<sup>2</sup>, Selina C. Wriessnegger<sup>3</sup>, Gernot R. Müller-Putz, Christoph Hintermüller<sup>4</sup>, Eileen Thomson<sup>2</sup> and Suzanne Martin<sup>1</sup>

<sup>1</sup>Faculty of life and Health Sciences, University of Ulster, Northern Ireland, UK.

<sup>2</sup>Cedar Foundation, Northern Ireland, UK

<sup>3</sup>Institute for Knowledge Discovery, Graz University of Technology, Graz, Austria

<sup>4</sup>G.tec Medical Engineering GmbH, Schiedlberg, Austria

s.martin@ulster.ac.uk

## Abstract

This paper focuses on the evaluation of a P300 BCI prototype to migrate from the lab to the home of people with acquired brain injury (ABI). Our evaluation involved both, end users and healthy controls. Overall, lower accuracy scores were recorded for the end-user group (55%) compared to 78% in the control group. The findings further indicate that participants were satisfied with the BCI but felt frustrated when it did not respond to their input. Overall, our evaluation indicated that BCI systems can work for people with ABI and the results will inform the development of subsequent prototypes.

## 1 Introduction

Brain-Computer Interfaces (BCI) are hardware and software systems that respond to brain signals, recorded by the electroencephalogram (Wolpaw et al. 2002). The goal of BCI technology is to increase independence, communication, rehabilitation outcomes, environmental control and social inclusion. Although BCIs can support a number of applications, little evidence of exploration of the systems beyond the laboratory has been undertaken. There are many reasons for this for example; complexity of the system and access to the target population makes this difficult. This research, carried out as part of the BackHome project (FP7/2007-2013), aims to build on laboratory-based results to develop a BCI system for home use.

Limited research has explored BCI technology with ABI end-users (Mulvenna et al. 2012). Post ABI a number of barriers can impact on a person's quality of life, including physical function, cognition and communication. This ambitious project will identify user requirements and system usability within this population by adopting a user-centered approach (Lightbody et al. 2011). End-user feedback will inform the technical developers through an iterative process in the design and development of the BCI system throughout the project. It is anticipated that the final prototype will be a system on which a number of services can be offered to support the transition from hospital to home, increase therapeutic outcomes, enable communication and home monitoring and control.

## 2 Methods

Eleven people were recruited to evaluate the prototype. Five participants (4 female, M= 35.6 years, range: 26-45; N= 3 had no prior experience of BCI) in a control group in advance of working directly with six target end users (1 female, M= 36 years, range: 25-48) who are living with ABI. The study design set out that each participant would aim to complete a predefined experimental protocol on three different occasions. The experimenter instructed all participants in detail prior to the measurement and checked whether they understood the paradigm before starting the measurement. Adherence to the protocol was monitored throughout the measurement.

The BCI prototype was implemented in Matlab Simulink (MathWorks, USA) and used a P300 based paradigm. One screen displayed the P300 speller matrix, while the other displayed the user interface used for controlling Facebook, Twitter and a desk light. The distance between user and screen was approximately one meter. EEG was acquired from eight active Ag/AgCl electrodes (g.Gamma, g.tec, Austria), at the positions Fz, Cz, P3, POz, P4, PO7, Oz and PO8. The channels were referenced to the right earlobe and a ground electrode was placed at FPz. Signals were amplified by a g.USBamp (g.tec, Austria), sampled at 256 Hz and band filtered between 0.5 and 30 Hz.

The set-up phase measured the time from sitting in front of the equipment until commencing the testing protocol. This included placing the cap/electrodes, adding gel, testing the signals, and creating the classifier. For classifier training, users were instructed to consecutively count the number of flashes of five specific letters in a 6 X 6 speller matrix. This data was then down-sampled to 64 Hz. Step-wise linear discriminant analysis then automatically determined the most discriminative features from the eight channels and the signal points in the 800 ms epochs after flash onset and setup the classifier model.

The testing phase required the participant to complete a 30-step protocol. The researcher guided the participant through the process, which included the selection of fifteen letters (Spelling task) and fifteen selections to navigate the system such as turn on/ off a light and read messages on Facebook or Twitter. Erroneous selections were not corrected. If users were unable to make the correct selection after three attempts the step was abandoned and they were directed to the next step in the protocol.

Each participant aimed to complete the protocol in three sessions on three different days, followed by the visual analogue scale (VAS) questionnaire to rate overall satisfaction. After the final evaluation session participants completed the extended Quebec User Evaluation of Satisfaction with Assistive Technology (QUEST) 2.0 (Demers, Weiss- Lambrou & Ska, 2002), a customized usability questionnaire and the (NASA-TLX) NASA-Task Load Index to assess workload Sharek (2009).

## 3 Results

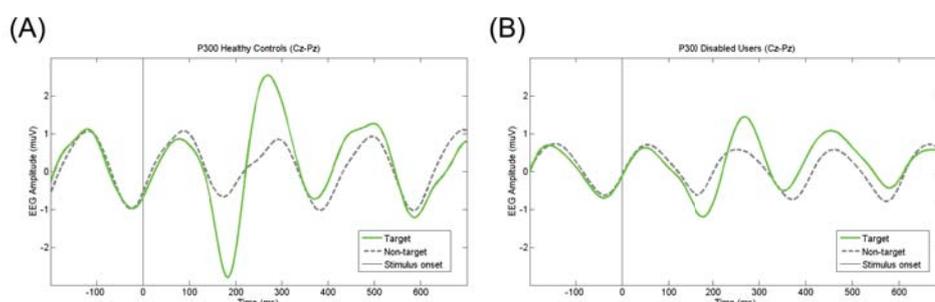
All of the set up was undertaken by non-BCI-experts and took 15.5 minutes on average with the control group as oppose to 27.6 minutes for the end users. The most challenging aspect of the set up was to assure good signal quality in all electrodes. Figure 1 below illustrates that the EEG responses are neurophysiologically sound over the multiple training sessions for both the control and end user's although some differences are noted in the strength of deflection of the P300 component. The Panels (A) and (B) show the averaged target and non-target EEG response in channel Cz over all healthy and disabled users. Both curves are averages over multiple training sessions. As baseline correction, we subtracted the mean of the 500 ms window before stimulus onset from every epoch. We rejected epochs with an overall average absolute amplitude higher than +8  $\mu$ V.

All five participants in the control group successfully completed the 30-step protocol over three sessions with an average accuracy of 78% (range: 65% to 91%). A total of 14 out of 18 attempted protocols were fully completed by the end users. The evaluation presents an average accuracy of 55%

(individual session range= 41% to 79%) for those completing a full protocol. Four protocols were partially completed up to step ten, three were stopped as the system was failing to respond and one because the participant had reached the cut off time of two hours using the system. The four protocols that were partially completed (accuracy ranged from 36% to 62%) brought the overall average accuracy down to 55%. One participant did not complete the evaluation as the system stopped responding to his commands after one complete protocol and two partly completed protocols. Specifically the final four steps of the protocol recorded considerably lower accuracy ranging from 37% to 45% especially in comparison to the control group that ranged from 71% to 83%.

**Table 1:** Outcomes from Control Group and End User Evaluation

Measurement Tool		Controls	Cedar Participants
Time	Set up	15.5 min ( $\pm 4.09$ )	27.6 min ( $\pm 14.48$ )
	Complete Protocol	15.8 min ( $\pm 8.17$ )	37.29 min ( $\pm 8.4$ )
Effectiveness: Accuracy	Protocol (navigation 30 selections)	78% ( $\pm 9.4$ )	55% ( $\pm 10.6$ )
	Spelling (15 letter selections)	83% ( $\pm 13.4$ )	62% ( $\pm 10.4$ )
Efficacy	NASA TLX	-----	58.56 ( $\pm 13.1$ : out of 100)
Satisfaction:	VAS	7.6 ( $\pm 1.67$ )	7.8 ( $\pm 1.9$ )
	QUEST 2.0	4.23 ( $\pm 1.9$ )	3.8 ( $\pm .86$ )
	QUEST added items	4 ( $\pm 0.88$ )	3.9 ( $\pm 1.19$ )



**Figure 1:** Panel (A) shows the brain response for healthy and Panel (B) for disabled users.

All participants indicated satisfaction following each session, with the control group reporting an average VAS score of 7.6 (range: 5-9) and target end-users average score of 7.8 (range: 5-10). The control group mean QUEST score was 4.23 (4= quite satisfied) and the average score of the added BCI related items was 4. The target end users reported a slightly lower average QUEST score of 3.8 (4= quite satisfied/3= more or less satisfied) and the average score of the added items was 3.9. The aspects rated as most important were effectiveness, comfort, ease of use, speed, and reliability. Additionally, target end users highlighted safety as an important fact to consider in the BCI design. The subjective workload using the NASA-TLX reported moderate to high workload scores (ranging from 44.66- 75.26 of 100 with a mean of 58.56).

## 4 Discussion

This evaluation demonstrates that it is feasible, valuable and worthwhile to engage directly with

people who have ABI in the development of a BCI system for home use. The work described here is part of an iterative process to improve the functionality and usability of BCI for users with ABI. Recruiting people without ABI, as a control, is helpful to test out the protocol and provide comparison data in advance of going to your target end user group. The results highlighted that although the end users were able to use the system the set up time was longer, tasks took longer to complete and the accuracy was lower. It is possible this is due to the participant's residual cognitive impairment as a result of ABI such as difficulty concentrating for periods of time as well as decreased stamina, memory and attention. Mental fatigue was indicated as an issue for the end-user group in the usability questionnaire and highlighted by the last four steps of the protocol recording considerably lower accuracy scores. The key findings from both groups included frustration when selections were incorrect and difficulty navigating through some aspects of the system. Equally, both groups recommended changes to the user interface, appearance of the cap/wires, and the control group suggested changes to the onscreen keyboard.

A number of limitations for every day use of the system by non-BCI-experts emerged from this evaluation. The most challenging aspect of the set up was achieving a stable signal from all the electrodes. It was also difficult to determine why the system was not responding to a particular user. There could be a number of reasons for the reduced response rate such as 'noisy' signals, a system failure, participant fatigue level or insufficient classifier accuracy however it is currently impossible for the non-expert user to resolve the issue to support the user.

The focus of the project is to move BNCI systems from the laboratory to the home of people with neurological conditions such as ABI. The findings indicated that BCI systems can work for people with ABI, which is promising, and will be used to inform the development and design of the subsequent two prototypes within the project. The evaluation provides important information to improve the prototype design and enhance the ability of the BCI to improve individual's functional ability, quality of life, and independence.

## 5 Acknowledgements

Funded by the FP7 EU research project BackHome (grant agreement n° 288566). We thank Josef Fallner and our partners from University of Würzburg and Barcelona Digital for technical help.

## 6 References

- Demers, L., Weiss-Lambrou, R., & Ska, B. (2002). Quebec user evaluation of satisfaction with assistive technology (QUEST 2.0): An overview and recent progress. *Technology and Disability, 14*, pp. 101-105.
- Lightbody, G., Ware, M., McCullagh, P., Mulvenna, M.D., Thomson, E., Martin, S., Todd, D., Medina, V.R., & Martinez, S.C. (2010). A user-centred approach for developing brain-computer interfaces. In: *Proceedings of Pervasive Computing Technologies for Healthcare, Munich 2010*, pp. 1-8.
- Mulvenna, M Lightbody, G., Thomson, E. McCullagh, P. Ware, M., and Martin, (2012). Realistic expectations with brain computer interfaces. *Journal of Assistive Technologies, 6 (4)*, pp. 233 – 244.
- Nicolas-Alonso, L. F. & Gonez-Gil, J. (2012). Brain computer interfaces, a review. *Sensors, 12*, pp. 1211 – 1279.
- Sharek, D. (2009). *NASA-TLX Online Tool (Version 0.06) [Internet Application]*. Research Triangle, NC. Retrieved from <http://www.nasatlx.com>
- Wolpaw, J., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain-computer interfaces for communication and control. *Clinical Neurophysiology, 113(6)*, pp. 767-791.

# Head-related Impulse Response-based Spatial Auditory Brain-computer Interface

Chisaki Nakaizumi<sup>1,\*</sup>, Toshie Matsui<sup>1</sup>, Koichi Mori<sup>2,\*</sup>, Shoji Makino<sup>1,\*</sup>  
and Tomasz M. Rutkowski<sup>1,3,\*</sup>,<sup>†</sup>

<sup>1</sup> Life Science Center of TARA, University of Tsukuba, Japan  
tomek@bci-lab.info & http://bci-lab.info/

<sup>2</sup> Research Institute of National Rehabilitation Center for Persons with Disabilities, Japan

<sup>3</sup> RIKEN Brain Science Institute, Japan

## Abstract

This study provides a comprehensive test of the head-related impulse response (HRIR) to an auditory spatial speller brain-computer interface (BCI) paradigm, including a comparison with a conventional virtual headphone-based spatial auditory modality. Five BCI-naive users participated in an experiment based on five Japanese vowels. The auditory evoked potentials obtained produced encouragingly good and stable P300-responses in on-line BCI experiments. Our case study indicates that the auditory HRIR spatial sound paradigm reproduced with headphones could be a viable alternative to established multi-loudspeaker surround sound BCI-speller applications.

## 1 Introduction

A brain-computer interface (BCI) is capable of providing a speller for disabled people with conditions such as amyotrophic lateral sclerosis (ALS). Although the currently successful visual modality may provide a fast BCI speller, patients at an advanced stage who are in a locked-in state cannot use the modality because they lose all intentional muscle control, including even blinking and movements of the eyes. An auditory BCI may be an alternative method because it does not require good eyesight. However, the modality is not as precise as the visual.

We propose an alternative method to extend the previously published spatial auditory BCI (saBCI) paradigm [1] by making use of a head-related impulse response (HRIR) for virtual sound image spatialization with headphone-based sound reproduction. Our research goal is a virtual spatial auditory BCI using HRIR-based spatialized cues in the part of the non-invasive, stimulus-driven, auditory modality which does not require long-term training. Experiments were conducted to reproduce and provide a comparison with previously reported vector-based amplitude panning (VBAP)-based spatial auditory experiments [1]. The more precise HRIR-based spatial auditory BCI stimulus reproduction was used to simplify previously reported real sound sources generated with surround sound loudspeakers [2].

HRIR appends interaural intensity differences (IID), interaural time differences (ITD), and spectral modifications to create the spatial stimuli, while VBAP appends only IID. HRIR allows for more precise and fully spatial virtual sound image positioning, even without utilizing the user's own HRIR measurements [3].

The next section of this paper describes the experiment set-up and the HRIR-based saBCI paradigm, together with EEG signal acquisition, pre-processing and classification steps. In the

---

\*The author was supported by the Strategic Information and Communications R&D Promotion Program (SCOPE) no. 121803027 of The Ministry of Internal Affairs and Communications in Japan.

<sup>†</sup>Corresponding author.

third section, the event related potentials (ERP), and especially the  $P300$  response latencies are described, with a classification and discussion of the HRIR-based saBCI paradigm information transfer rate (ITR) results, including a comparison with the conventional method. Finally, the conclusions and future research directions are indicated.

## 2 Methods

All of the experiments were performed at the Life Science Center of TARA, University of Tsukuba, Japan. Five paid BCI-naive users participated in the experiments. The average age of the users was 21.6 years (standard deviation 0.547 years; five females). The psychophysical and online EEG BCI experiments were conducted in accordance with *The World Medical Association Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects*. The experiment procedures were approved and designed in agreement with the ethical committee guidelines of the Faculty of Engineering, Information and Systems at the University of Tsukuba, Japan. Five Japanese vowels ( $a$ ,  $i$ ,  $u$ ,  $e$ ,  $o$ ) were used in this experiment. The vowels were taken from a sound dataset of female voices [4]. The monaural sounds were spatialized using the public domain CIPIC HRTF DATABASE provided by the University of California, Davis [5]. Each Japanese vowel was set on a horizontal plane at azimuth locations of  $-80^\circ$ ,  $-40^\circ$ ,  $0^\circ$ ,  $40^\circ$ ,  $80^\circ$  for the vowels  $a$ ,  $i$ ,  $u$ ,  $e$ ,  $o$ , respectively. The psychophysical experiments were conducted to investigate the response time and recognition accuracy. The users were instructed to respond by pressing the button as soon as possible after they perceived the *target* stimulus, as in a classical oddball paradigm [6]. In a single psychophysical experimental run, 20 *targets* and 80 *non-targets* were presented. An online EEG experiment was conducted to investigate the  $P300$  response with BCI-naive users. The brain signals were collected with a biosignal amplifier system g.USBamp by g.tec Medical Engineering GmbH, Austria. The EEG signals were captured by sixteen active gel-based electrodes attached to the following head locations  $Cz$ ,  $Pz$ ,  $P3$ ,  $P4$ ,  $Cp5$ ,  $Cp6$ ,  $P1$ ,  $P2$ ,  $Poz$ ,  $C1$ ,  $C2$ ,  $FC1$ ,  $FC2$ , and  $FCz$ , as in the extended 10/10 international system. The ground electrode was attached on the forehead at the  $FPz$  location, and the reference on the user's left earlobe. BCI2000 software was used for the saBCI experiments to present stimuli and display online classification results. A single experiment was comprised of five runs which contained 10 *target* and 40 *non-target* stimuli. Each run contained five selections. The stimulus duration was set to 250 ms, the interstimulus interval (ISI) to 150 ms, and brain signal ERPs were averaged 10 times for each vowel classification. In brief, the single experiment was comprised of 25 selections. The EEG sampling rate was set to 512 Hz, and a 50 Hz notch filter to remove electric power line interference was applied in a rejection band of 48 – 52 Hz. The band pass filter was set with 0.1 Hz and 60 Hz cut-off frequencies. The acquired EEG brain signals were classified online by the BCI2000 application using a stepwise linear discriminant analysis (SWLDA) classifier with features drawn from the 0 ~ 800 ms ERP interval.

## 3 Results

This section presents and discusses results obtained from the psychophysical and EEG experiments conducted with five users, as described in the previous section. In the psychophysical experiment, the accuracy rates for all stimuli were above 94%. The majority of responses were concentrated at the 350 ms latency. There were no significant differences in the response times between the target stimuli as tested by ANOVA ( $p < 0.05$ ). The results of the EEG experiment are depicted in Figure 1. The left panel shows the grand mean averaged ERP results at four

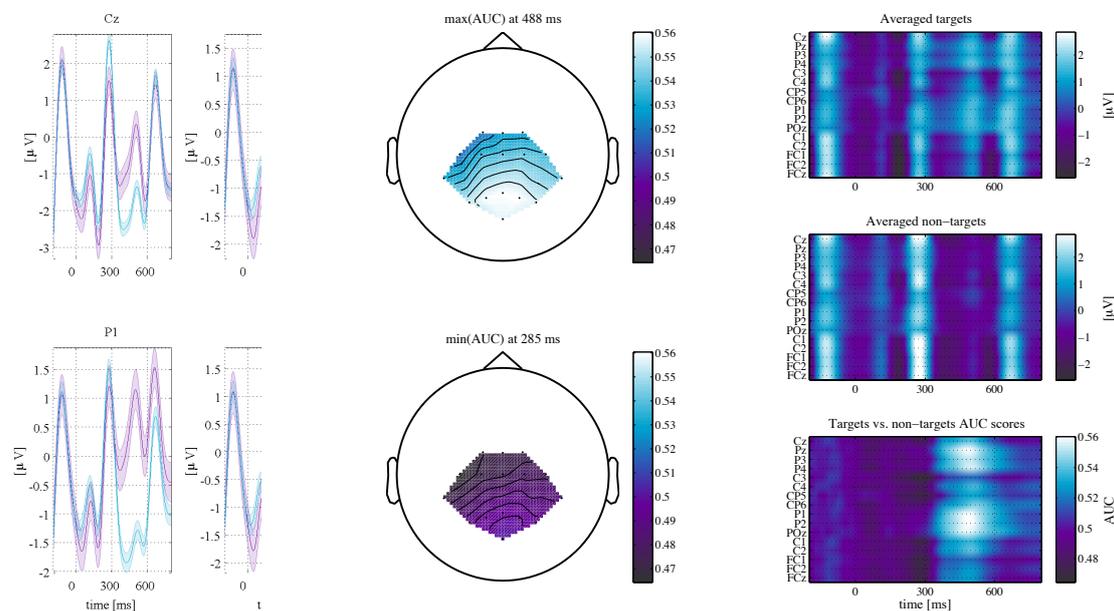


Figure 1: Grand mean averaged ERP and AUC scores leading to final classification results for the participants. The left panel shows the averaged ERP for all participants. The purple line shows the brain waves for *targets*, and blue line is for *non-targets*. The centre panel presents the head topographies at the maximum and minimum AUC scores as obtained from the right bottom panel. The top right panel presents averaged ERP responses to the *target* and the middle panel to the *non-target* stimuli. The right bottom panel visualizes the AUC analysis results of *target* versus *non-target* response distribution differences.

representative electrodes. The centre panel provides the results as scalp topographies at the maximum and minimum area under curve (AUC) of a receiver operating with characteristic values [7] for *target* vs. *non-target* latencies. It also demonstrates the EEG electrode positions used in the experiments. The top right panel indicates the averaged ERP responses of all electrodes to the *target*, and the second panel shows responses to the *non-target* stimuli. The bottom panel indicates the AUC of *target* versus *non-target* responses, clearly confirming the usability of 400 ~ 600 ms latencies for the subsequent classification. Table 1 presents the classification accuracies of the P300 responses as obtained with the SWLDA classifier and the ITR scores. The average score was obtained as a mean value calculated from 1 ~ 5 runs (the training run was not included in the calculation of accuracies). The ITR is a major comparison measure [7] among the BCI paradigms. All five users scored above the five vowel sequences spelling chance levels of 20%. There was one user who achieved 100% accuracy, which was the best in the experiments reported. We also compared the ITR scores with a VBAP-based spatial auditory BCI, which is regarded as a conventional method [1]. The VBAP experiment was conducted in 2 runs and with 16 BCI-naïve users in [1]. The electrode positions were the same as in our current experiments. The sound stimuli were presented with small ear-fitting headphones in both the modalities. The ISI was set to 500 ms in the VBAP experiment, and to 150 ms in the HRIR experiment. In the VBAP modality, the average ITR score was 1.05 bit/min and the best was 1.78 bit/min. In the HRIR modality, the average ITR was 1.35 bit/min and the best was 2.40 bit/min. The ITR scores of the HRIR experiment were recalculated for 2 runs,

User	Run							ITR [bit/min]	
	1	2	3	4	5	Average	Best	Average	Best
#1	60%	80%	40%	20%	40%	48%	80%	2.26	9.60
#2	0%	40%	100%	80%	100%	64%	100%	5.27	18.58
#3	20%	20%	0%	40%	80%	32%	80%	0.46	9.60
#4	20%	20%	20%	20%	40%	24%	40%	0.06	1.21
#5	0%	40%	40%	80%	60%	44%	80%	1.70	9.60

Table 1: Vowel spelling accuracies and ITRs of each user obtained in the EEG experiments

the same as for the VBAP experiment. HRIR based modality produced better results than the VBAP based modality for both the average and the best score.

## 4 Conclusions

The EEG results presented confirm the P300 responses of BCI-naive users. The mean accuracy was not very good owing to the short ISI, but the accuracy tends to improve when the number of run increases. Therefore, more attention training or interface using practice may be necessary for BCI-naive users. The ITR scores were higher compared (no significance analysis due to different user groups) with our previous study using HRIR stimuli, and also compared with the previously reported VBAP-based spatial auditory BCI. Nevertheless, the current study is not able to compete with the faster visual BCI spellers. Furthermore, it is necessary to improve the ITR for a more comfortable spelling. We plan to continue research with larger numbers of sound stimuli, a better suited ISI, and more complex spatial sound patterns.

## References

- [1] M. Chang, N. Nishikawa, Z. R. Struzik, K. Mori, S. Makino, D. Mandic, and T. M. Rutkowski. Comparison of P300 responses in auditory, visual and audiovisual spatial speller BCI paradigms. In *Proceedings of the Fifth International Brain-Computer Interface Meeting 2013*, page Article ID: 156, Asilomar Conference Center, Pacific Grove, CA USA, June 3-7, 2013. Graz University of Technology Publishing House, Austria.
- [2] T. M. Rutkowski, A. Cichocki, and D. P. Mandic. Spatial auditory paradigms for brain computer/machine interfacing. In *International Workshop On The Principles and Applications of Spatial Hearing 2009 (IWPASH 2009) - Proceedings of the International Workshop*, page P5, Miyagi-Zao Royal Hotel, Sendai, Japan, November 11-13, 2009.
- [3] J. Schnupp, I. Nelken, and A. King. *Auditory Neuroscience - Making Sense of Sound*. MIT Press, 2010.
- [4] S. Amano, S. Sakamoto, T. Kondo, and Y. Suzuki. Development of familiarity-controlled word lists 2003 (FW03) to assess spoken-word intelligibility in Japanese. *Speech Communication*, 51(1):76–82, 2009.
- [5] V. R. Algazi, R. O. Duda, D. M. Thompson, and C. Avendano. The CIPIC HRTF database. In *2001 IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics*, pages 99–102. IEEE, 2001.
- [6] J. Wolpaw and E. Winter Wolpaw, editors. *Brain-Computer Interfaces: Principles and Practice*. Oxford University Press, 2012.
- [7] M. Schreuder, B. Blankertz, and M. Tangermann. A new auditory multi-class brain-computer interface paradigm: Spatial hearing as an informative cue. *PLoS ONE*, 5(4):e9813, 04 2010.

# Single-Trial Classification and Evaluation of Hemodynamic Responses during Passive and Active Exercises for Neurorehabilitation

Chang-Hee Han<sup>1\*</sup>, Han-Jeong Hwang<sup>2</sup>, Jeong-Hwan Lim<sup>1</sup>  
, and Chang-Hwan Im<sup>1†</sup>

<sup>1</sup> Department of Biomedical Engineering, Hanyang University, Seoul, Republic of Korea.

<sup>2</sup> Department of Computer Science, Technical University of Berlin, Berlin, Germany.

zeros8706@bme.hanyang.ac.kr\*, ich@hanyang.ac.kr†

## Abstract

In this study, we recorded and evaluated hemodynamic responses induced by conventional passive exercise for neurorehabilitation and combined exercise strategy (passive exercise with active motor execution or motor imagery). Functional near infrared spectroscopy (fNIRS) was recorded while eight healthy subjects conducted three different tasks (passive motor execution alone, passive motor execution with motor imagery, and passive motor execution with active motor execution). From the results, stronger and broader activation around the sensorimotor cortex was observed when subjects performed the combinatory strategies. Results of single-trial pattern classification showed the classification accuracy of more than 70%, demonstrating that fNIRS could be used as a potential tool to monitor how actively the users engaged in the combinatory neurorehabilitation strategy.

## 1 Introduction

Patients with motor impairment due to central nervous system diseases such as stroke and Parkinson's disease have difficulties in their daily-life activities, and many of them are not able to perform basic body movements without caregivers or assistive devices. To now, many researchers have developed various neurorehabilitation strategies for recovering their damaged motor functions.

Recently, researchers have shown that a combinatory neurorehabilitation strategy that simultaneously uses both passive and active exercise can dramatically enhance the outcomes of conventional passive exercises. For example, Joa et al. compared brain activations induced by the passive exercise (passive motor execution by functional electrical stimulation (FES)), the active exercise (voluntary contraction), and combination of both using functional magnetic resonance imaging (fMRI) (Joa, 2012), and demonstrated that the combination of the passive and active exercises may be more effective for successful neurorehabilitation.

Nevertheless, this combinatory strategy has a critical limitation in that the practitioners cannot know how actively the patients perform the active motor execution or motor imagery during passive exercise. Therefore, it is needed to monitor patients' active engagement in the rehabilitation program.

---

\* Created the first draft of this document

† Masterminded and created the first stable version of this document

In this study, we investigated whether functional near-infrared spectroscopy (fNIRS) could be used to evaluate whether the users actively performed active motor execution or motor imagery during passive exercise. To this aim, we compared the concentration changes of oxygenated hemoglobin (oxy-Hb) induced by “passive exercise alone” and “combinatory strategies”. We then classified the “passive exercise alone” and the “combinatory strategies” using single-trial fNIRS data.

## 2 Methods

### 2.1 Subjects

A total of eight healthy subjects (six males and two females, 24-30 years old) were enrolled in our study. All subjects were right-hander and had normal or corrected-to-normal vision. None of them had a previous history of neurological, psychiatric, or other severe diseases that might affect the experimental results.

### 2.2 Experimental setup and procedure

We used a commercial multi-channel NIRS instrument (FOIRE-3000 from Shimadzu Corporation, Kyoto, Japan) for recording cortical activity. The absorptions of three wavelengths (780, 805, and 830 nm) of near-infrared light were acquired at a sampling rate of 10 Hz and then transformed into concentration changes of oxygenated hemoglobin (oxy-Hb), deoxygenated hemoglobin (deoxy-Hb), and total hemoglobin (total-Hb) using the modified Beer-Lambert law (Delpy 1988). The system detected changes in the cortical concentration levels (mM cm) of oxy-, deoxy- and total-Hb. In this study, we used 40 channels with twelve illuminators and thirteen detectors. The international 10-20 system was adopted to locate the optodes to cover whole motor-related areas such as premotor cortex, supplementary motor cortex (SMA), primary motor cortex (M1), and posterior parietal cortex. During the experiments, all subjects performed three different tasks. First task was “passive exercise alone” (passive motor execution: PME). In this task, subject’s right index finger was automatically moved by a finger rehabilitation device regardless of the subject’s movement intention. Second task was “combinatory strategy 1” (passive motor execution with motor imagery: PME+MI). In this task, the subjects simultaneously performed PME and motor imagery of their right index finger. Third task was “combinatory strategy 2” (passive motor execution with active motor execution: PME+AME). The subjects performed both passive motor execution and active motor execution of their right index finger at the same time. One trial consisted of a task period of 10 seconds and a rest period of 10-15 seconds, and all subjects repeated each task 40 times.

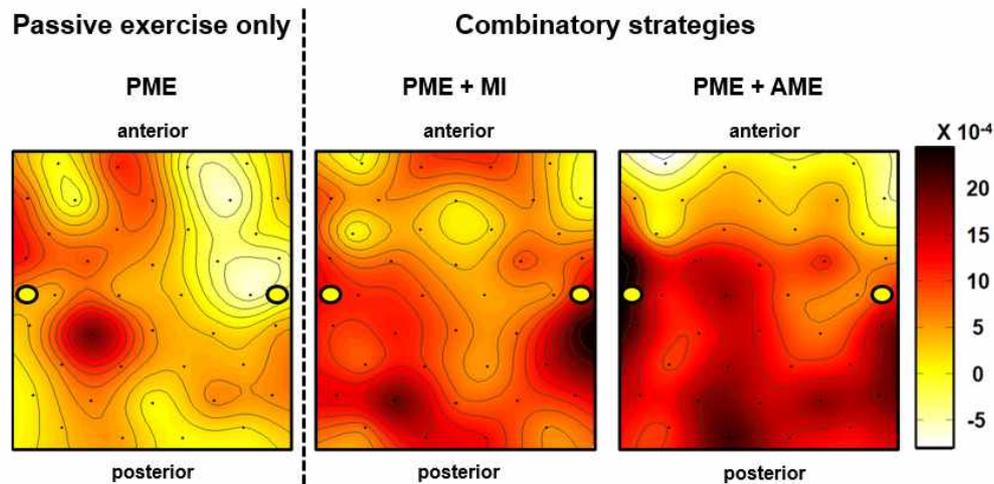
### 2.3 Data analysis

The concentration changes of oxy-, deoxy-, and total-Hb were processed using various preprocessing methods. First, we applied a common average reference (CAR) spatial filter to remove respiratory and cardiac noises. Then, the preprocessed data were band-pass filtered using 4<sup>th</sup>-order zero phase Butterworth filter with a pass-band of 0.01-0.1 Hz to reduce physiological noises. The filtered data were divided into 20-s epochs, including the task period from 0 to 10 s and the rest period from 10 to 20 s. For the baseline correction, we used a base reset method, which adjusts a first sample to the zero point (Lee, 2012). We then conducted statistical analysis and single-trial pattern classification. First, to evaluate the statistical differences among hemodynamic responses elicited by the “passive exercise alone (PME)” or the “combinatory strategy 1 (PME+MI) and 2 (PME+AME)”, we performed Friedman test (significant level:  $p < 0.05$ ) and Wilcoxon signed rank test (post-hoc analysis using Bonferroni corrected p-value). Through this procedure, we selected only statistically

meaningful channels among 40 channels. Second, we classified the “passive exercise alone” and the “combinatory strategies” using the hemodynamic responses of the selected channels. We used five moving windows with different windows sizes of 1, 2, 4, 5, and 10 seconds and extracted features by moving the windows along the time axis with 50% overlap. Five features (mean, variance, kurtosis, skewness, and slope) were evaluated for each moving window. The linear bayes normal classifier implemented in the PRtools package was used for the binary classification (Duing, 2000). The 10 by 10 cross-validation was performed for the estimation of classification accuracy.

### 3 Results

Figure 1 shows topographic maps of oxy-Hb concentration changes induced by the “passive exercise alone” and the “combinatory strategies”. It could be observed from the figures that the oxy-Hb concentration changes during both “combinatory strategies” were stronger and broader than that during “passive exercise alone”.



**Figure 1:** Topographic map of oxy-Hb concentration changes in each task. Left and right yellow circles represent locations of C3 and C4, respectively.

Table 1 shows the results of the single-trial pattern classification. The “passive exercise alone” and the “combinatory strategies” could be classified with classification accuracy larger than 70%, which is high enough to be used as an indicator to evaluate whether the users actively performed active motor execution or motor imagery during passive exercise (Choularton, 2004).

Subject No.	Classification accuracy (%)	
	PME vs PME + MI	PME vs PME + AME
1	69.4	70.3
2	73.4	74.1
3	77.5	72.3
4	70.9	68.4
5	73.4	73.4
6	67.0	73.4
7	75.0	73.4
8	67.4	68.5
Mean	71.8	71.7

**Table 1:** Results of single-trial pattern classification

## 4 Discussion and Conclusion

In the present study, we investigated whether fNIRS could be used to evaluate how actively users performed active motor execution or imagery during passive exercise. To the best of our knowledge, the classification of neural responses elicited by “passive exercise” and “combinatory strategies” was not reported before. Our experimental results showed stronger and broader neural activation during the “combinatory strategies” than the conventional “passive exercise alone”. In addition, two different neurorehabilitation strategies could be classified with the classification accuracy larger than 70%, demonstrating that the user’s active engagement in the neurorehabilitation might be successfully monitored using fNIRS.

## References

- K.-L. Joa, Y.-H. Han, C.-W. Mun, B.-K. Son, C.-H. Lee, and Y.-B. Shin. (2012). *Evaluation of the brain activation induced by functional electrical stimulation and voluntary contraction using functional magnetic resonance imaging*. Journal of neuroengineering and rehabilitation
- D. T. Delpy, M. Cope, P. van der Zee, S. Arridge, S. Wray, and J. Wyatt. (1988). *Estimation of optical pathlength through tissue from direct time of flight measurement*. Physics in medicine and biology
- S. Lee, D. Koh, A. Jo, H. Y. Lim, Y.-J. Jung, C.-K. Kim, Y. Seo, C.-H. Im, B.-M. Kim, and M. Suh. (2012). *Depth dependent cerebral hemodynamic responses following direct cortical electrical stimulation (DCES) revealed by in vivo dual-optical imaging techniques*. Optics express
- R, P. W. Duin. (2000) *PRTools version 3.0: A Matlab toolbox for pattern recognition*. Delft University of Technology

# Is heart rate variability a predictor for neurofeedback effects?

Elisabeth V. C. Friedrich<sup>1</sup>, Hristos Courellis<sup>1</sup>, Alexandra L. Tonnesen<sup>1</sup>,  
Richard Gevirtz<sup>2</sup> and Jaime A. Pineda<sup>1</sup>

<sup>1</sup>University of California San Diego, La Jolla, CA, USA

<sup>2</sup>Alliant International University, San Diego, CA, USA  
efriedrich@ucsd.edu, jpineda@ucsd.edu

## Abstract

The goal of this study was to investigate the relationship between heart rate variability (HRV) and performance in a clinical application for brain-computer interfaces, namely Neurofeedback training (NFT) for children with autism spectrum disorder (ASD). HRV parameters in an initial pre-training test predicted performance as well as performance improvements in social cognition following NFT, confirming a relationship between the autonomic nervous system and social cognition. Furthermore, HRV improved after each NFT session in comparison to pre-session levels. However, no direct relationship between resting HRV and NFT performance was found. The decrease in HRV over sessions might be explained by an increase in mental effort due to an increase in difficulty of the NFT protocol. Nonetheless, HRV can serve as a predictor for NFT effects on social dysfunction in autism and HRV-biofeedback might lead to further improvements in social cognition for autistic children.

## 1 Introduction

Recent studies have suggested that heart rate and its variability (HRV) can be used as predictor for mental effort in a brain-computer interface (BCI) (Pfurtscheller et al. 2013) and can predict P300-based BCI performance (Kaufmann et al. 2012). Additionally, a positive relation between HRV and social cognition was recently shown in able-bodied individuals using the ‘Reading the Mind in the Eyes-test’ (RMET; Baron-Cohen et al. 2001; Quintana et al. 2012). These links between brain and body could be especially important for clinical applications of BCIs such as Neurofeedback training (NFT) for social dysfunction in children with autism spectrum disorder (ASD).

In this study, an EEG-based NFT for children with ASD was implemented as in previous work (Pineda et al. 2008, 2014). ASD is characterized with deficits in social and communicative skills such as imitation and empathy. Besides neurophysiological abnormalities, deficits in the social engagement system have been linked to the regulation of heart rate (Porges 2007). The goal of this study was to investigate if the reported positive relation between HRV and social cognition as well as BCI performance can be replicated in children with ASD. Therefore, (1) resting HRV was correlated with performance in the RMET which was used as an indicator for social cognition and NFT success and (2) HRV in a resting baseline before and after each NFT session was correlated to NFT performance.

## 2 Methods

Thirteen children (aged 6-17, one female) with a confirmed diagnosis of ASD participated in a pre- and posttest and in sixteen 1-hour NFT sessions twice a week. During the NFT sessions, EEG

was recorded from one electrode over the right sensorimotor cortex (C4), sampled at 256 Hz and filtered for mu (8-12 Hz), theta (3-8 Hz) and high beta (18-30 Hz) frequency bands. Before and after each NFT session, a 5-min baseline was recorded for HRV analyses (see last paragraph of methods). The children were trained to control a video game involving social interactions by modulation of their mu frequency band. They were not provided with specific control strategies but learned by operant conditioning via feedback. The beta and theta frequency bands inhibited positive feedback in the game if the amplitudes in these frequencies exceeded a certain threshold. The thresholds for mu, beta and theta were set as a function of an initial preceding resting period of baseline activity. The amplitude value for mu was set in the first session and then shaped to get higher in the following sessions. In contrast, the amplitude values of theta and beta were shaped to get lower in the subsequent sessions.

Performance during the NFT was calculated as:  $Performance = Hitrate \times Difficulty$  (1)

A hit was defined as fulfilling all threshold criteria (e.g. above mu and below beta and theta) and thus triggering positive feedback in the game. In order to make different parts of the game comparable, the hitrate was defined as:  $Hitrate = \frac{Hits}{Minutes}$  (2)

Due to the continuous shaping of the threshold, the possibility of fulfilling all threshold criteria decreased and thus difficulty increased over sessions. In order to adjust the hitrate for the level of difficulty, the distances ( $\Delta_{obs}$ ) between the shaped threshold and the preceding baseline were considered. Distances were calculated so that positive numbers reflected the threshold being set easier than the baseline values and negative numbers reflected the threshold being set more difficult than the baseline values. The observed distances were then normed ( $\Delta_{norm}$ ) to the defined standard distance ( $\Delta_{std}$ ). The standard distance was set in a way that the mu ( $\mu$ ) threshold was 50% lower and the beta ( $\beta$ ) and theta ( $\theta$ ) threshold was 50% higher than the preceding baseline value which resulted in a difficulty of 1 for the first session. In order to ensure that a distance of zero and negative distance values still show reasonable results, a logarithmic transformation was used:

$$Difficulty = \frac{1}{10^{\Delta_{norm}}} \quad \Delta_{norm} = \frac{(\sum_{i=\mu\beta\theta} \Delta_{obs_i} - \sum_{i=\mu\beta\theta} \Delta_{std_i})}{(\sum_{i=\mu\beta\theta} \Delta_{std_i})} \quad (3)$$

In the pre- and posttest, two 6-min baselines with open and closed eyes were recorded before the children completed the RMET. In this test, pictures of individuals' eye regions were shown. Based on the eyes, children had either to determine what the individual is thinking or feeling out of 4 possible choices presented at the four corners of the display, or what gender (male or female) the individual is. The percentage of correct responses (Corr%), as well as the reaction time (RT), was calculated.

The electrocardiogram (ECG) was recorded at a sampling rate of 2048 Hz from electrodes attached to the left wrist and the right side of the neck. HRV parameters were analyzed in all 5-min resting baselines before and after each NFT session as well as in the pre- and posttest. The ECG data was down-sampled to 512 Hz and the interbeat interval (IBI) detection and artifact correction were made with the software ARTiiFACT (Kaufmann et al. 2011). The statistical parameters for HRV were calculated with KUBIOS (Tarvainen et al. 2014) and included the SDNN in ms (i.e. standard deviation of IBIs) in the time domain and the power ( $ms^2$ ) and percentage (%) of high- (HF; 0.15-0.4 Hz) and low-frequency (LF; 0.04-0.15 Hz) measures in the frequency domain derived using Fast Fourier Transformation. SDNN and HF have a positive, LF (%) a negative association with HRV.

### 3 Results

First, the relationship between HRV and the RMET was investigated. A normal distribution was found for all variables. Children had more Corr% in the gender (M=84, SE=5) than in the emotion recognition (M=52, SE=6) task ( $F_{1,12}=74.5, p<.01$ ) confirming the differences in difficulty of the tasks, and more Corr% in the post- (M=70, SE=5) than in the pretest (M=65, SE=6) by trend ( $F_{1,12}=3.9, p<.1$ ) suggesting improvement as a function of training. There were neither significant effects for RT in the RMET nor for HRV parameters between pre- and posttest.

HRV parameters in the resting baseline of the pretest - but not of the posttest - correlated with Corr% and RT in the RMET of the pre- and posttest (Table 1). In order to adjust the correlations for possible influence of age and gender of the participants, partial correlations were calculated. All correlations above  $r=\pm.3$  were in the expected direction: The higher SDNN and HF, the higher the Corr% and the shorter the RT in the RMET. For the LF-percentage, the opposite occurred.

Second, HRV parameters during the baselines before and after each NFT session were correlated with hitrate, difficulty and performance during each NFT session. However, the correlations failed to show any consistent pattern. Hitrate decreased over sessions ( $F_{7,84}=9.1, p<.01$ ), whereas difficulty and performance increased ( $F_{7,84}=3.7/1.9, p<.05/.1$ ). The parameters SDNN and HF power showed significantly higher HRV in the baselines after than before each NFT session ( $F_{1,12}=44.4/13, p<.01$ ). However, HF ( $ms^2, \%$ ) decreased and LF ( $\%$ ) increased over the training sessions ( $F_{7,84}=2.3-3.4, p<.05$ ).

		Pretest				Posttest			
		Gender		Emotion		Gender		Emotion	
Pretest		Corr%	RT	Corr%	RT	Corr%	RT	Corr%	RT
	SDNN	.49	-.29	<b>.61*</b>	-.41	.47	-.40	<b>.66*</b>	-.44
Eyes	HF ( $ms^2$ )	.44	-.31	<b>.57<sup>(*)</sup></b>	-.47	.44	-.42	<b>.61*</b>	-.40
open	HF (%)	.50	.25	.20	-.19	<b>.68*</b>	-.47	.48	-.40
	LF (%)	<b>-.55<sup>(*)</sup></b>	-.25	-.22	.10	<b>-.70*</b>	.42	-.40	.40
Eyes	HF (%)	<b>.63*</b>	-.07	.36	<b>-.53<sup>(*)</sup></b>	<b>.67*</b>	-.47	<b>.63*</b>	-.47
closed	LF (%)	<b>-.62*</b>	.05	-.37	.51	<b>-.71*</b>	<b>.54<sup>(*)</sup></b>	<b>-.63*</b>	.51

**Table 1: Bivariate partial correlation coefficients controlled for age and gender between parameters of HRV and RMET.** The significance level is indicated with the asterisks (\*  $p<.05$ , <sup>(\*)</sup>  $p<.1$ ).

### 4 Discussion and conclusion

First, our results confirm a positive correlation between HRV and the performance in the RMET, which was proposed by Quintana et al. (2012). Moreover, this relationship between the autonomic nervous system and social cognition can be extended to an autistic population of children. As individuals with ASD have deficits in social cognition, the correlations between HRV in the pretest and emotion recognition in the pre- and posttest suggest that training HRV before starting NFT might boost performance in social cognition. Although participants improved in social cognition as a function of the NFT (i.e. more Corr% in the post- compared to the pretest), this improvement was not evident in HRV (i.e. no difference in HRV between pre- and posttest), which should be linked to the social engagement system (Porges 2007). This suggests that while these systems may occasionally function in a connected or coupled way, they are distinct and orthogonal systems.

Second, we could not find consistent correlations between HRV and performance in our NFT, which was controlled by children on the autism spectrum by modulating frequency bands in the EEG. This is in contrast to Kaufmann et al. (2012) who reported that performance in a P300-based BCI

could be predicted by HRV in able-bodied individuals. In the present study, HRV improved after each NFT session compared to before, which suggests that the NFT-game was successful in inducing stress-free social interactions. However, across training sessions, HRV decreased, whereas performance increased. As the increase in performance was the product of increasing difficulty and decreasing hitrate, we could explain the decrease in HRV as an increase in mental effort (Pfurtscheller et al. 2013; Cowley et al. 2013). Analyzing HRV during game play itself and not only in the baseline before and after each NFT might reveal more consistent correlations with performance parameters.

To conclude, although we could not show consistent correlations between resting HRV and NFT performance, this study extends the idea that HRV is a good predictor for behavioral performance in social cognition in individuals with ASD, and can be used as an indicator of mental effort rather than performance during NFT. Moreover, the results suggest that HRV-biofeedback might improve NFT effects in social cognition for children on the spectrum.

### Acknowledgements

This research was supported by a fellowship provided by the Max Kade Foundation to the Department of Cognitive Science, UCSD and by grant funding from the ISNR Research Foundation. The authors thank Kaitlin Huennekens and Anna Christina Macari for help with the data analyses and all colleagues from the CNL and HWU for their support.

## References

- Baron-Cohen S, Wheelwright S, Hill J, Raste Y, Plumb I (2001). The ‘Reading the Mind in the Eyes’ Test Revised Version: a Study with Normal Adults, and Adults with Asperger Syndrome or High-functioning Autism. *Journal of Child Psychology and Psychiatry* 42: 241–51.
- Cowley B, Ravaja N, Heikura T (2013). Cardiovascular physiology predicts learning effects in a serious game activity. *Computers & Education* 60: 299–309.
- Kaufmann T, Sütterlin S, Schulz S M, Vögele C (2011). ARTiiFACT: a Tool for Heart Rate Artifact Processing and Heart Rate Variability Analysis. *Behavior Research Methods* 43: 1161–70.
- Kaufmann T, Vögele C, Sütterlin S, Lukito S, Kübler A (2012). Effects of Resting Heart Rate Variability on Performance in the P300 Brain-computer Interface. *International Journal of Psychophysiology* 83: 336–41.
- Pfurtscheller G, Solis-Escalante T, Barry R J, Klobassa D S, Neuper C, Müller-Putz G R (2013). Brisk Heart Rate and EEG Changes During Execution and Withholding of Cue-paced Foot Motor Imagery. *Frontiers in Human Neuroscience* 7: 379.
- Pineda J A, Brang D, Hecht E, Edwards L, Carey S, Bacon M, Futagaki C, et al (2008). Positive Behavioral and Electrophysiological Changes Following Neurofeedback Training in Children with Autism. *Research in Autism Spectrum Disorders* 2: 557–581.
- Pineda J A, Carrasco K, Datko M, Pillen S, Schalles M (2014). Neurofeedback Training Produces Normalization in Behavioral and Electrophysiological Measures of High Functioning Autism. In *Mirror Neurons: Fundamental Discoveries, Theoretical Perspectives and Clinical Implications*, edited by P F Ferrari and G Rizzolatti. Royal Society Publishing, London.
- Porges, S W (2007). The Polyvagal Perspective. *Biological Psychology* 74: 116–43.
- Quintana D S, Guastella A J, Outhred T, Hickie I B, Kemp A H (2012). Heart Rate Variability Is Associated with Emotion Recognition: Direct Evidence for a Relationship Between the Autonomic Nervous System and Social Cognition. *International Journal of Psychophysiology* 86: 168–72.
- Tarvainen M P, Niskanen J-P, Lipponen J, Ranta-Aho P O, Karjalainen P (2014). Kubios HRV-heart Rate Variability Analysis Software. *Computer Methods and Programs in Biomedicine* 113: 210–20.

# Independent BCI Use in Two Patients Diagnosed with Amyotrophic Lateral Sclerosis

Elisa Mira Holz, Loic Botrel, Andrea Kübler

Institute of Psychology, University of Würzburg, Germany

elisa.holz@uni-wuerzburg.de, loic.botrel@uni-wuerzburg.de,  
andrea.kuebler@uni-wuerzburg.de

## Abstract

A P300 based brain-computer interface (BCI) application for creative expression was implemented at the homes of two artists diagnosed with amyotrophic lateral sclerosis (ALS). Here we present data of independent BCI home use for up to 23 months. Our findings indicate that a P300 BCI can be used in the patients' daily life with high satisfaction and gratification for both BCI end-users.

## 1 Introduction

To date BCIs have hardly been established at the patients' homes. To our knowledge there exist only 3 studies presenting independent BCI use [1-3]. The current study aimed at demonstrating proof-of-principle of independent BCI use and investigating usability of independent BCI use in severely disabled patients. A P300 BCI for creative expression, *Brain Painting*, was implemented at the patients' homes. Here we present data within two patients, who used the BCI at their homes for several months. The study covering data of the first 14 months (200 sessions) for the first patient is published in its extended form in [4]. In this paper we present follow-up data and also data of a second patient, who has been participating in the study for 5 months.

## 2 Method

For a detailed description of methods the reader is referred to [4].

### 2.1 Patients

Both patients (one female), aged 73 and 74, were diagnosed with ALS. Both are artists and with progress of disease related paralysis were no longer able to paint. Patient 1 is in the locked-in state with preserved eye-movements. Patient 2 is tetraplegic, able to talk and to move his head.

### 2.2 Easy to use BCI and set-up

Electroencephalography (EEG) was recorded with 8 active electrodes (gamma cap, g.tec, Graz) from electrode positions Fz, Cz, Pz, P3, P4, Po7, Po8, Oz. *Brain Painting* [5] was facilitated for

independent home use [4], in the way that *Brain Painting* can be started within a few clicks and BCI parameters, such as calibration data, are loaded automatically. In an initial meeting including two consecutive days, BCI was calibrated and the family and/or assistants were trained in setting up the BCI system. BCI is used independent of the help of BCI experts. BCI data is automatically transmitted to a remote-server to the BCI lab. BCI use is monitored and supported in case of e.g. technical problems using a remote desktop application. Patient 1 was visited in further meetings for e.g. improvement of the system or recalibration. At date of submission, patient 1 has been using the BCI for 23 months and patient 2 for 5 months. Both studies are ongoing.

### 2.3 Face valid measures of BCI use and evaluation

A BCI session was considered valid, if the end-user painted for at least several seconds. Total painting time is the summed painting time of all runs excluding beaks per session. After every BCI session end-users indicated their satisfaction and frustration with the BCI session on visual analogue scales (VAS) ranging from zero (not at all satisfied/ frustrated) to 10 (very satisfied/ frustrated). BCI users could comment on the sessions in a comment line. Subjective level of BCI control was rated by the BCI end-users within four classes (percentage of correct selections): (1) zero (0-50%), (2) low (50-70%), (3) medium (70-90%) and (4) high (90-100 %).

## 3 Results

### 3.1 Face valid measures of BCI use

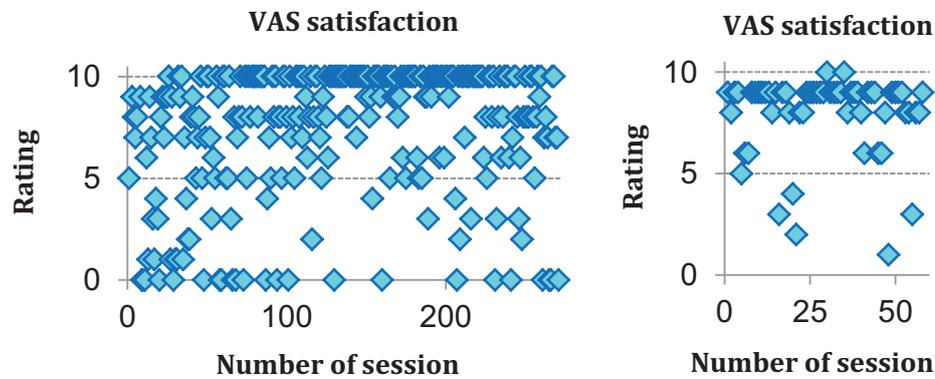
In 23 months patient 1 painted in 271 sessions, with a total painting time of 403.48 hours with a mean painting duration of  $M=89.33$  ( $SD=52.87$ ) minutes. Patient 2 painted with the BCI in 58 sessions within 5 months, resulting in a total painting time of 64.60 hours with a mean painting duration of  $M=66.82$  ( $SD=34.56$ ) minutes.

### 3.2 Subjective level of BCI control

In most sessions subjective level of BCI control was rated as *medium*, in 94 sessions for patient 1 and 35 sessions for patient 2. *Low* was indicated in 75 and 12 sessions, *zero* in 51 and 1 session(s) and *high* in 43 and 10 sessions, for patient 1 and patient 2 respectively.

### 3.3 Satisfaction with BCI sessions

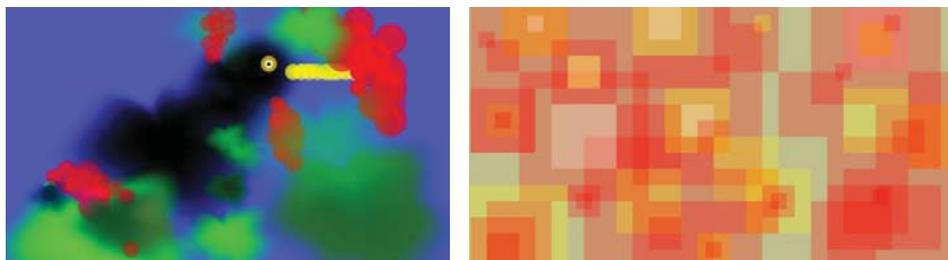
Overall satisfaction was high in both subjects,  $M=7.42$  ( $SD=3.21$ ) for patient 1 and  $M=7.95$  ( $SD=2.0$ ) for patient 2, consequently overall frustration was low,  $M=2.88$  ( $SD=3.06$ ) and  $M=3.02$  ( $SD=2.36$ ) respectively, (**figure 1**). Dissatisfaction occurred due to (1) technical problems with software, mostly at the beginning of the study, or with hardware, e.g. malfunctioning of amplifier or EEG electrodes, (2) low BCI control due to e.g. drying electrode gel, shifting EEG cap and (3) personal factors, such as attention issues (fatigue, concentration) and disease factors (e.g. cough), both influencing BCI control too, and satisfaction with painting.



**Figure 1:** VAS satisfaction for patient 1 (**left**) and patient 2 (**right**) over time (271 and 58 sessions). Zero indicates “not at all satisfied” and 10 “very satisfied”.

### 3.4 Paintings

Patient 1 created more than 120 paintings, patient 2 more than 30. Two paintings are depicted in **figure 2**.



**Figure 2:** Paintings: *Black bird* by patient 1 (**left**) and *Patchwork* by patient 2 (**right**), with permission of the artists.

## 4 Discussion

The present study demonstrates that independent BCI use is possible with high satisfaction and gratification for the patients. BCI use was challenged by technical problems, BCI control and personal factors. Technical problems, occurring mostly at the beginning of the study (in patient 1), could be solved and fully removed. Dissatisfaction with the painting occurred also only at the beginning of the study and disappeared with more practice in using the *Brain Painting* matrix. Attentional issues could be reduced through frequent breaks within sessions and time management, e.g. BCI session is not started, if the patient is too tired. Disease related factors, such as cough, however are less likely to be overcome and will remain in daily-life. In most of the BCI sessions end-users rated BCI control as

being *medium*. BCI control depended on the adjustment of EEG cap and electrodes, but also on personal factors (e.g. concentration and cough). Problems on BCI control could be decreased by improvements on hard- and software, e.g. EEG cap and electrodes or an auto-calibration option [6]. Interestingly patients indicated moderate to high satisfaction even if BCI control is moderate and not 100%. Patient 2 states: “*I do not want to focus of my attention on my cough. I have enough time and I refine until I am satisfied with it [painting]. This is the way I did it in the past as a [hand] painting artist [...]*”. This shows that, even with low(er) BCI control, the patient can nevertheless be satisfied with BCI as assistive technology. Patient 1 points out: “[...] *Brain Painting makes me happy and satisfied everytime again, if it works - if not, then frustration and disappointment. Fortunately this happens in only few cases. BP completely changed my life. I can paint again and be creative again using colours [...]*”.

## 5 Conclusion

BCI controlled painting is accepted well as assistive technology by the severely disabled end-users. *Brain Painting* exactly matches the artists’ need for creative expression and positively impacts the quality of life of the patients. P300-BCIs have come of age and can be used independently in daily-life. To replicate our findings, inclusion of more patients is planned.

## 6 Acknowledges

This work is supported by the European ICT Programme Project FP7- FP7-288566 (Backhome). This paper only reflects the authors’ views and funding agencies are not liable for any use that may be made of the information contained herein.

## References

- [1] Neumann, N., A. Kübler, J. Kaiser, T. Hinterberger, and N. Birbaumer, Conscious perception of brain states: mental strategies for brain-computer communication. *Neuropsychologia*, 2003. 41(8): p. 1028-36.
- [2] Karim, A.A., T. Hinterberger, J. Richter, J. Mellinger, N. Neumann, H. Flor, A. Kubler, and N. Birbaumer, Neural internet: Web surfing with brain potentials for the completely paralyzed. *Neurorehabil Neural Repair*, 2006. 20(4): p. 508-15.
- [3] Sellers, E.W., T.M. Vaughan, and J.R. Wolpaw, A brain-computer interface for long-term independent home use. *Amyotroph Lateral Scler*, 2010. 11(5): p. 449-55.
- [4] Holz, E.M., L. Botrel, T. Kaufmann, and A. Kübler, Long-term independent brain-computer interface home use improves quality of life of a patient in the locked-in state: A case study. *Arch Phys Med Rehabil*, accepted for publication.
- [5] Zickler, C., S. Halder, S.C. Kleih, C. Herbert, and A. Kübler, Brain Painting: usability testing according to the user-centered design in end users with severe disabilities. *Artif Intell Med*, 2013. 59(2): p. 99-110.
- [6] Kaufmann, T., Völker, S., Gunesch, L., and Kübler, A. Spelling is Just a Click Away - A user-centered brain-computer interface including auto-calibration and predictive text entry. *Front. Neurosci.* 2012;6:72.

# Unifeature vs. multifeature oddball and magnitude of deviance effects on the mismatch negativity

Helena Erlbeck<sup>1\*</sup>, Andrea Kübler<sup>1</sup>, Boris Kotchoubey<sup>2</sup>, and Sandra Vesper<sup>2</sup>

<sup>1</sup>University of Würzburg, Germany

<sup>2</sup>University of Tübingen, Germany

helena.erlbeck@uni-wuerzburg.de

## Abstract

The mismatch negativity (MMN) is commonly studied in oddball paradigms comprising only one deviant. Recently, multifeature oddball paradigms comprising multiple deviants have been developed and might represent a valuable tool in clinical assessment where time is a limiting factor. This study addressed whether (1) duration MMNs elicited by physically identical stimuli differ between a unifeature vs. multifeature oddball paradigm, and (2) the magnitude of the standard vs. deviant difference would affect the MMN amplitude. Event-related potentials were recorded in 26 healthy participants who listened to three oddball paradigms (two unifeature, one multifeature). Results showed that (1) duration deviants elicited larger MMNs in the multifeature oddball, and (2) MMN amplitudes increased with increasing deviant-minus-standard distance in terms of the target stimulus feature. Taken together, duration oddballs elicit reliable MMNs in multifeature paradigms and may therefore represent a valuable tool in clinical settings.

## 1 Introduction

The mismatch negativity (MMN) is frequently used to investigate discrimination skills and residual cognitive functioning in patients with disorders of consciousness (DOC). These results may then serve as a basis for the future application of a brain-computer interface (BCI). Traditionally, the MMN is studied in auditory oddball designs using a secondary task to prevent elicitation of an N2b contaminating the MMN effect. It peaks between 150 and 250 ms (Duncan et al., 2009). Typical paradigms comprise one standard and one rare deviant, thus focusing on only one feature of auditory stimuli. However, attention span of DOC patients is short, so paradigms have to be brief, with a maximum on information, but at the same time allow for a reliable detection of MMNs in single subjects. Using oddball paradigms comprising several deviant tones in various dimensions may be a promising tool to record event-related potentials (ERPs) in response to more than one deviating stimulus within a short period of time. These multifeature oddball paradigms were first developed by Näätänen, Pakarinen, Rinne, and Takegata (2004) whose tone sequence comprised deviants in five different dimensions: frequency, intensity, location, duration, and a silent gap in the middle of a tone. They found MMN responses in the multifeature paradigm to be as large as in the traditional unifeature oddball, which was also supported by (Pakarinen et al., 2009).

---

\* Created the first draft of this document

With this study we addressed whether (1) the MMNs elicited by physically identical stimuli differ as a function of paradigm uni- vs. multifeature oddball), and (2) the amplitude of the MMN increases with the increasing deviant-minus-standard distance (DSD) in terms of the target stimulus feature in three different dimensions (frequency, intensity, and duration).

## 2 Methods

Event-related potentials were recorded in 26 healthy participants (mean age = 34.1 years, SD = 9.9, 14 females). EEG was recorded with a 32 active electrodes system (gTec, Graz, Austria) with a sampling frequency of 512 Hz. The data were filtered online between 0.1 to 100 Hz (butterworth). The paradigms were presented while participants watched a silent movie and were required to press a key when a certain scene appeared. A scale for subjectively experienced effort ranging from 0 to 220 (Eilers, Nachreiner, & Hänecke, 1986) was administered after each paradigm.

Auditory stimulation was realized in three different paradigms. All paradigms shared the same harmonic tone as standard with 500+1000+1500 Hz, 75 ms duration (rise/fall 5 ms) and an intensity of 70dB. The paradigms used were the following: (1) a duration oddball called MMNAbsolut comprising 900 standards and 100 odds differing in duration (50 ms), (2) a second duration oddball paradigm called MMNProportion comprising 900 standards and 100 odds differing in duration as well (37 ms) and (3) a multifeature oddball called MMNMultifeature with 600 standards and 600 deviants varying in three dimensions: duration, frequency, and intensity. The deviants in the multifeature paradigm were 37 and 50 ms in the duration domain, 450+900+1350 and 400+800+1200 Hz in the frequency domain, and 65 and 60dB in the intensity domain. Stimulus onset asynchrony was 500 ms in all paradigms.

Offline, the EEG data were band-pass filtered between 0.1 and 25 Hz, automatic artefact correction was performed, and segments from 0 to 500 ms were averaged. Finally, grand averages were obtained. For all calculations, mean amplitudes under the curve for the electrode Fz were entered into analyses. All analyses were carried out in MATLAB (The Math Works, Inc., M.A.), statistics were calculated in SPSS 17.0 (SPSS Inc., IL). Relevant time windows for component analyses in all paradigms were defined by visual inspection and set to 120 to 245 ms post-stimulus in all paradigms.

## 3 Results

Results were obtained in repeated measures ANOVAs including the factors paradigm (unifeature vs. multifeature), DSD (small vs. large), and deviant type (duration, frequency, intensity).

### 3.1 Subjective effort

Ratings for subjective effort did not vary according to the paradigm ( $F(2, 56) = .165, p = .848$ ). Thus, listening to the various paradigms was judged to be equally effortless ( $M_{MMNAbsolut} = 57.58, SD = 42.81, M_{MMNProportion} = 53.79, SD = 38.62, M_{MMNMultifeature} = 54.39, SD = 45.52$ ).

### 3.2 Physiological data

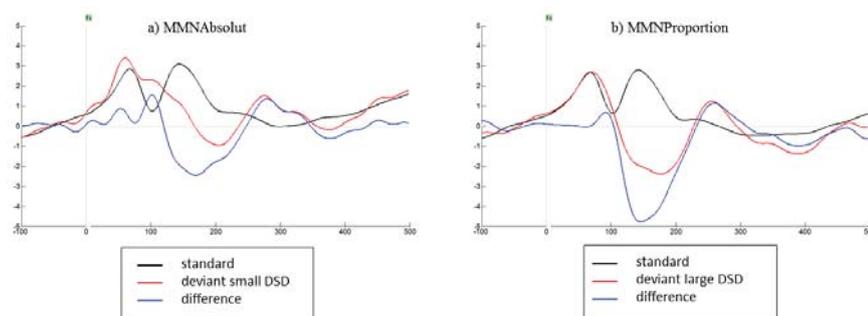
Figure 1 and Figure 2 show the recorded MMN potentials for all paradigms and deviant types at electrode position Fz.

First, we calculated an ANOVA including the factors paradigm and DSD to compare MMN amplitudes elicited by physiologically identical duration deviants of different DSDs in the unifeature

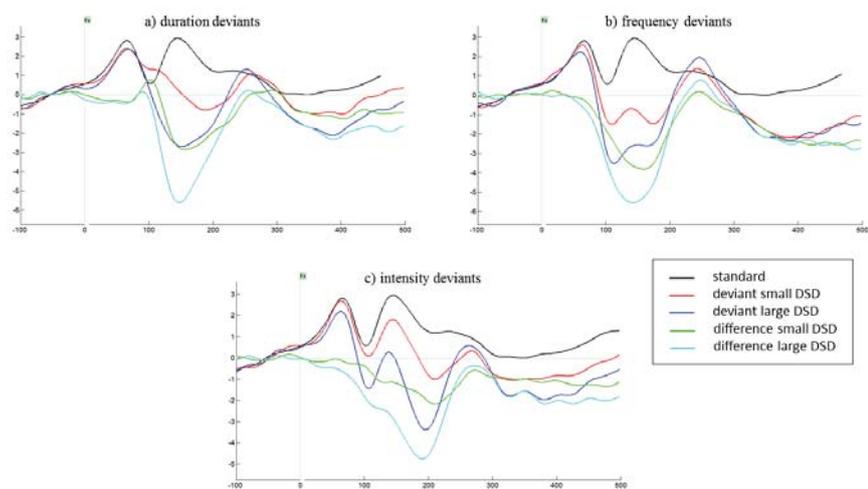
vs. multifeature paradigm. An MMN was elicited in all conditions. The results reveal a significant main effect of paradigm ( $F(1, 25) = 6.29, p = .019$ ) with larger amplitudes in the multifeature paradigm ( $M_{Multifeature} = -2.73, M_{Unifeature} = -2.26$ ) and a significant main effect of DSD ( $F(1, 25) = 22.03, p < .001$ ) with larger amplitudes when DSD was large ( $M_{DSD\ small} = -1.76, M_{DSD\ large} = -3.23$ ). Thus, for duration deviants, larger amplitudes were found in the multifeature paradigm and when the DSD was large.

Secondly, we calculated an ANOVA including the factors deviant type and DSD to compare MMN amplitudes elicited by duration, frequency, and intensity deviants in the multifeature paradigm. An MMN was elicited in all conditions. MMN amplitudes did not differ according to the deviant type, but varied as a function of DSD ( $F(1, 25) = 71.95, p < .001$ ) with larger amplitudes for large DSDs ( $M_{DSD\ small} = -1.97, M_{DSD\ large} = -3.34$ ). Furthermore, DSD effects varied as a function of deviant type ( $F(2, 50) = 3.45, p = .040$ ): The difference in MMN amplitudes between small and large DSD was smallest for frequency deviants ( $M_{Diff\ frequency} = .76$ ), larger for duration deviants ( $M_{Diff\ duration} = 1.48$ ), and largest for intensity deviants ( $M_{Diff\ intensity} = 1.85$ ). All SIDAK corrected post-hoc comparisons between DSDs were significant ( $p < .013$ ).

Thus, also for intensity and frequency deviants, MMN amplitudes increase with larger DSDs.



**Figure 1:** recorded MMN potentials at Fz in the unifeature oddball paradigms MMNAbsolut (a) and MMNProportion (b)



**Figure 2:** recorded MMN potentials at Fz in the multifeature oddball paradigm for duration deviants (a), frequency deviants (b), and intensity deviants (c)

## 4 Discussion and Conclusion

Duration, frequency and intensity oddballs elicited reliable MMNs also in multifeature paradigms. For duration deviants, we can even assert that physically identical deviants elicit higher MMNs in the multifeature than in the unifeature oddball paradigm. Higher amplitudes in multifeature paradigms may be of benefit in patient setting: Firstly, patients often exhibit smaller amplitudes in general. Secondly, testings are based on single subject analysis, and thirdly, time and attentional awareness are a limited resources. These benefits are supported by our finding that listening to a multifeature paradigm is judged to be as effortless as listening to a unifeature oddball.

Like previous studies, our results show an increase in MMN amplitudes for high DSDs (e.g. Jaramillo, Paavilainen, & Näätänen, 2000). However, Horváth and colleagues (2007) postulated that this difference might only be due to a contamination with N1 effects. By minimizing the N1 effect, they did not find a significant variation according to the magnitude of deviance. May and Tiitinen (2010) summarized that any measurable MMN is N1-contaminated. This might also be the case in our study and definite quantification of MMN vs. N1 needs to be further investigated.

Taken together, our results indicate multifeature oddball paradigms to be a suitable tool to investigate auditory discrimination profiles. Eliciting larger duration MMNs than the unifeature oddball paradigm, renders them especially interesting for research in DOC patients. In these patients, successful completion of such basic tasks is a precondition for the application of more complicated tasks like BCIs that also require prolonged attention and the switching of the attentional focus.

## References

- Duncan, C. C., Barry, R. J., Connolly, J. F., Fischer, C., Michie, P. T., Näätänen, R., ... (2009). Event-related potentials in clinical research: Guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400. *Clinical Neurophysiology*, *120*, 1883–1908. doi:10.1016/j.clinph.2009.07.045
- Eilers, K., Nachreiner, F., & Hänecke, K. (1986). Entwicklung und Überprüfung einer Skala zur Erfassung subjektiv erlebter Anstrengung. *Zeitschrift für Arbeitswissenschaft*, *40*(4), 215–224.
- Horváth, J., Czigler, I., Jacobsen, T., Maess, B., Schröger, E., & Winkler, I. (2007). MMN or no MMN: No magnitude of deviance effect on the MMN amplitude. *Psychophysiology*, *45*. doi:10.1111/j.1469-8986.2007.00599.x
- Jaramillo, M., Paavilainen, P., & Näätänen, R. (2000). Mismatch negativity and behavioural discrimination in humans as a function of the magnitude of change in sound duration. *Neuroscience Letters*, *290*, 101–104. doi:10.1016/S0304-3940(00)01344-6
- May, P. J. C., & Tiitinen, H. (2010). Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. *Psychophysiology*, *47*(1), 66–122. doi:10.1111/j.1469-8986.2009.00856.x
- Näätänen, R., Pakarinen, S., Rinne, T., & Takegata, R. (2004). The mismatch negativity (MMN): towards the optimal paradigm. *Clinical Neurophysiology*, *115*, 140–144. doi:10.1016/j.clinph.2003.04.001
- Pakarinen, S., Lovio, R., Huotilainen, M., Alku, P., Näätänen, R., & Kujala, T. (2009). Fast multifeature paradigm for recording several mismatch negativities (MMNs) to phonetic and acoustic changes in speech sounds. *Biological Psychology*, *82*, 219–226. doi:10.1016/j.biopsycho.2009.07.008

# Real-time artifact correction enables EEG-based feedback inside the fMRI scanner

C. Zich<sup>1</sup>, S. Debener<sup>1</sup>, M. De Vos<sup>2</sup>, C. Kranczioch<sup>1</sup>, I. Gutberlet<sup>3</sup>

<sup>1</sup>Neuropsychology Lab, University of Oldenburg, Germany.

<sup>2</sup>Methods in Neurocognitive Psychology, University of Oldenburg, Germany.

<sup>3</sup>BlindSight GmbH, Schlitz, Germany.

catharina.zich@uni-oldenburg.de

## Abstract

The aims of this contribution were (1) to investigate the feasibility of motor imagery (MI) electroencephalogram (EEG)-based online feedback during continuous functional magnetic resonance imaging, and (2) to examine the validity of electric MI signatures by comparing EEG MI data obtained inside and outside the MRI scanner. Participants (N=22) imagined hand movements, which led to an event-related desynchronization (ERD) of the contralateral sensorimotor rhythm. The contralateral ERD of the online corrected as well as of the offline corrected EEG data correlated strongly with the EEG data recorded outside the MRI scanner. We provide a proof-of-principle demonstration that meaningful EEG-based feedback inside the MRI scanner is possible for MI.

## 1 Introduction

As both movement execution and motor imagery (MI) decrease the amplitude of sensorimotor rhythms, termed event-related desynchronization (ERD), MI electroencephalogram (EEG)-based neurofeedback seems a promising intervention supporting the recovery of motor impairments. For the improvement of MI practice it is highly relevant to investigate the neurophysiological correlates of MI EEG-based feedback using high spatial resolution imaging. Combining MI EEG-based feedback with functional magnetic resonance imaging (fMRI) allows addressing several important aspects, such as a) an investigation of the role of feedback; b) a better understanding of individual differences in MI; and c) an identification of the brain patterns underlying the neural correlates of EEG MI feedback.

However, to successfully implement MI EEG-based neurofeedback inside the fMRI scanner good EEG signal quality must be achieved, which requires the online correction of the gradient artifact (GA) and the ballistocardiogram (BCG) artifact. As a first step we investigated the feasibility of real-time EEG-based MI feedback during fMRI, as this would open the door to address the above stated aspects. The validity of the MI EEG data collected inside the MRI scanner was verified by comparing the ERD after online and offline artifact correction with the ERD obtained outside the MRI scanner.

## 2 Methods

### 2.1 Subjects and Task

Twenty-four healthy individuals (11 females, 20-30 years of age, mean age: 23.9 years) with no MI experience participated in this study, which was approved by the local ethics committee. The data from two subjects had to be excluded because of noncompliance with task instructions. Participants were

measured twice with a six-week interval in between. During the first session concurrent EEG-fMRI was recorded and during the second session the same experiment was recorded outside the scanner in the EEG lab. Each session consisted of four blocks and each block comprised 40 trials (20 left and 20 right hand). Subjects were instructed to execute a finger tapping movement in the first block and to imagine the same movement in the last three blocks. Two imagery blocks were performed without feedback and in the last one online feedback was provided. Stimulus presentation was controlled with OpenViBE and followed the standard Graz MI protocol [1], except that we employed longer inter-trial intervals (4.5 to 9 sec) to accommodate the fMRI protocol timing.

### 2.3 EEG and fMRI Data Acquisition

EEG data were collected from 64 equidistant scalp sites, with Cz as reference, Iz as ground and ECG at the lower back, using an MR-compatible BrainAmp system and the BrainVision Recorder software 1.20.0506 (Brain Products GmbH, Gilching, Germany). Raw EEG was sampled at 5kHz, both GA and BCG artifacts were corrected online and corrected data were passed on to OpenViBE via a direct network link. fMRI data acquisition was performed on a 3T Siemens MRI scanner (Siemens AG, Erlangen, Germany). During functional measurements 420 T2\*-weighted gradient echo planar imaging volumes (3.1 x 3.1 x 3.0 mm voxels, 0.75 mm gap,  $TR = 1.5s$ ,  $FoV = 200 * 200$ , Flip Angle =  $90^\circ$ , 27 transversal slices) were obtained within each block. The EEG data obtained outside the scanner were acquired using the same paradigm and EEG setup, except for a lower sampling rate of 500 Hz.

### 2.4 Online processing of EEG Data

To enable real-time EEG-based neurofeedback inside the MRI scanner, EEG data had to be immediately corrected for MRI artifacts in all MI blocks. Online artifact corrections of GA and BCG were performed stepwise using the BrainVision RecView software 1.42 (Brain Products GmbH, Gilching, Germany). The online GA correction method employed is based on the average artifact attenuation method proposed by [2] with the notable difference that online corrections can only incorporate the already recorded data into the GA correction template. An artifact subtraction template of 1500ms was built, based on the MR system TR events and the corrected data were down sampled to 500 Hz using a block down sampling algorithm. Subsequently the data were filtered using a Butterworth low-pass with an edge frequency of 35 Hz (48 dB slope).

Online BCG correction first performed a search for the most suitable BCG template within the first 30 seconds of GA-corrected EEG data. This prototypical BCG template was then used in a moving template matching approach to identify subsequent valid BCG episodes. Minimum correlation threshold for BCG detections was 0.7, with an allowed average amplitude ratio range from 0.6x to 1.4x relative to the prototypical BCG data. The total BCG correction window length varied between 800 (-100 to 700ms post R-wave) and 900 ms (-100 to 800 ms post R-wave) depending on the average inter-beat-interval of the subjects as determined before the main recordings commenced. The BCG artifact subtraction template was based on a moving window comprised of 21 subsequent BCG episodes.

The two MI blocks without feedback were concatenated and filtered from 8-30 Hz. MI intervals were extracted using EEGLAB. Artifactual segments were rejected and a common spatial pattern algorithm was applied to train a linear discriminant analysis classifier [3]. During the feedback block data were preprocessed online in OpenViBE, keeping the parameter settings identical to those used in EEGLAB. The output of the classifier was translated into the length of a light blue horizontal bar (updated at a frequency of 16 Hz).

### 2.5 Offline processing of EEG Data

For offline analysis standard preprocessing steps were applied on the uncorrected raw EEG data. Offline corrections of GA and BCG artifacts were performed in separate steps using the BrainVision Analyzer Professional software package 2.0.4 (Brain Products, Gilching, Germany). The GA correction time window was placed from -35 to 1465 ms relative to the MR system TR marker to accurately capture

all MR gradient activity. All other GA correction parameters were the same as for online GA correction (see above). Analysis of the BCG artifact structure in the (GA corrected) data showed that the ECG channel data was afflicted with strong additional artifacts related to respiratory cycle chest movements. Since even strict filtering of the ECG data (3 to 12 Hz - 48 dB slope Butterworth filter) typically did not remove these artifacts, a spatio-temporal matching approach was employed that detected BCG episodes based on spatial and temporal BCG signatures in all EEG channels. This method resulted in near perfect detection of valid BCG episodes. All other parameters were the same as for online BCG correction, with the notable exception that offline BCG correction automatically determined the optimal position and length of the BCG artifact subtraction window relative to the R-markers detected.

After offline correction, data were re-referenced to the common average. All three MI blocks were concatenated, filtered from 8 and 30 Hz, segmented into consecutive time intervals of one second and gross artifactual segments rejected. The remaining data were submitted to extended infomax ICA to identify and subsequently attenuate eye-blink, eye-movement and heart-beat artifacts. Artifact-corrected EEG data were segmented to extract task-related event-related desynchronization (ERD). The ERD time course was computed following the procedure proposed by [4].

Online and offline analysis of the EEG data recorded outside the MRI scanner was exactly the same except for the necessity of MRI artifact correction in the former.

### 3 Results

As reported previously MI caused an ERD prominent above the contralateral sensorimotor areas [e.g. 1, 5, 6]. This effect was evident for all three conditions (Figure 1), online, offline and outside the MRI scanner. Descriptively, the relative power decrease of the contralateral ERD at electrode sites corresponding to C3 and C4 was on average  $-18.5\% \pm 15.1\%$  for the online corrected EEG signal,  $-31.1 \pm 14.9\%$  for the offline corrected signal and  $-28.7\% \pm 13.5\%$  for the data obtained outside the scanner. Analysis of the ERD time course using a one-way ANOVA with factor condition (online corrected, offline corrected, EEG recorded outside the MR) revealed a significant main effect ( $F_{2,42}=12.47$ ,  $p<.001$ ,  $\eta^2=.37$ ). Subsequent paired sample t-tests revealed no significant difference in ERD amplitude between offline corrected data and those recorded outside the MRI scanner ( $t_{21}=-.83$ ,  $p=.42$ ) but a significantly lower ERD amplitude in the online compared to offline corrected ( $t_{21}=5.83$ ,  $p<.001$ ) data. The standard deviation within the MI interval across trials was significantly higher for online corrected compared to offline corrected data,  $t_{21}=-3.5$ ,  $p=.003$ .

Examination of the association between the online as well as the offline corrected EEG data recorded inside and outside the MRI scanner revealed strong correspondences for the contralateral ERD ( $r=.56$ ,  $p=.007$ ;  $r=.53$ ,  $p=.01$ ). Analysis of the overall spatial ERD pattern was performed for the 30 electrodes

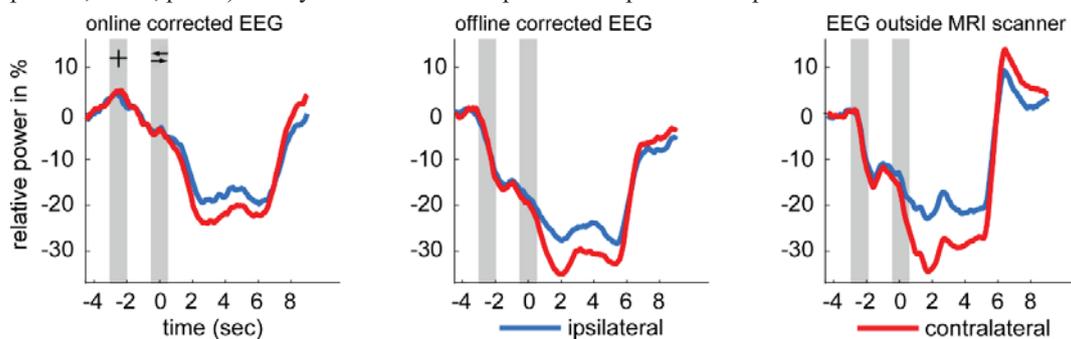


Figure 1: Grand average ERD from N=22 for online and offline corrected EEG data and data obtained outside the scanner at C3 and C4. Blue indicates contra- and red ipsilateral electrode sites, with respect to stimulus direction.

on the top of the scalp. Between the offline corrected data and the data acquired outside the scanner high associations for right and left hand MI ( $r=.87$ ,  $p<.001$ ;  $r=.88$ ,  $p<.001$ ) were found, weather no significant correspondences could be obtained for online corrected data and data obtained outside the MRI scanner.

## 4 Discussion

Motivated by our interest in the spatial underpinnings of MI EEG-based neurofeedback, the possibility of EEG-based feedback during fMRI acquisition for MI was explored and the validity of MI EEG data after online and offline artifact correction examined. Although the procedure is technically challenging, it offers the unique possibility to investigate in detail the brain patterns involved in EEG-BCI operation. This may help to further develop effective MI EEG-based feedback.

For the first time we provide a proof-of-principle demonstration that EEG-based feedback during continuous MRI scanning is possible for MI. This is illustrated by high associations of the contralateral ERD, for EEG data obtained inside and outside the MRI scanner. The contralateral ERD was significantly lower for online versus offline corrected data, however as EEG data and artifact correction protocols were identical for online and offline corrections, the differences can only be attributed to the vastly reduced efficacy of the ECG detection method used in online versus offline corrections. A possible adaptation of the spatio-temporal matching algorithm used for the offline ECG detection thus holds excellent promise for a further improvement of the efficacy of online EEG-fMRI based MI protocols. Furthermore the development of more efficient online artifact correction procedures can accelerate the correction process and thereby reduce the delay of feedback inside the MRI scanner. However, even with the currently implemented methods, we were able to show that it is possible to remove MRI related artifacts in real-time, which is necessary to provide online feedback.

The present findings open up new possibilities to investigate the neurophysiological underpinnings of MI, which seems necessary in light of the promising role of MI for stroke recovery.

## Acknowledgments

This work is supported by a travel grant of the Universitätsgesellschaft Oldenburg to CZ.

## References

- [1] G. Pfurtscheller and C. Neuper. Motor imagery activates primary sensorimotor area in humans. *Neuroscience Letters*, 239(2-3):65-68, 1997.
- [2] P.J. Allen, O. Josephs, and R. Turner. A Method for Removing Imaging Artifact from Continuous EEG Recorded during Functional MRI. *NeuroImage*, 12:230-239.2000.
- [3] B.Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K.R. Müller. Optimizing Spatial Filters for Robust EEG Single-Trial Analysis. *IEEE Signal Processing Mag*, 25(1):41-56, 2008.
- [4] G. Pfurtscheller and F.H. Lopes da Silva. Event-related EEG/MEG synchronization and desynchronization: basic principles. *Electroenceph clin Neurophysiol*, 110(11):1842-1857, 1999.
- [5] B. Blankertz, F. Losch, M. Krauledat, G. Dornhege, G. Curio, and K.R. Müller. The Berlin Brain-Computer Interface: accurate performance from first-session in BCI-naive subjects. *IEEE Trans Biomed Eng*, 55(10):2452-2462, 2008.
- [6] C. Zich, M. De Vos, C. Kranczoch, and S. Debener. Wireless EEG with individualized channel layout enables efficient motor imagery training. In press.

# Decoding of Executed Movements and Source Imaging

Patrick Ofner<sup>1,2</sup>, Joana Pereira<sup>3</sup> and Gernot R. Müller-Putz<sup>1,2</sup>

<sup>1</sup> Graz University of Technology, Graz, Austria

<sup>2</sup> BioTechMed-Graz, Graz, Austria

<sup>3</sup> University of Lisbon, Lisbon, Portugal

patrick.ofner@tugraz.at, gernot.mueller@tugraz.at

## Abstract

A brain-computer interface (BCI) in combination with a neuroprosthesis can be used to restore movement in paralyzed persons. Usually, the control of such BCIs by the deliberate modulation of brain oscillations is unnatural and unintuitive. Recently, low-frequency brain signals have been found to encode movement information and can be used to decode movement trajectories. These signals are potential candidates for new types of BCIs which can be used to naturally control neuroprosthesis by imagining movement trajectories. We analyzed the contributing brain areas in the source space and found motor areas but also parietal and lateral areas encoding movement information.

## 1 Introduction

A brain-computer interface (BCI) allows the control of devices through brain-signals. In combination with a neuroprosthesis, e.g., functional electrical stimulation (FES), a BCI can be used to restore motor functions in paralyzed persons [5]. One type of non-invasive BCIs – so called sensorimotor rhythm (SMR) BCIs – are usually based on the deliberate modulation of brain oscillations [7] by movement imagery (MI). However, the control of SMR-based BCIs is not natural and intuitive, because MIs are assigned artificially to control functions (e.g., a foot MI controls the elbow function). Furthermore, SMR-based BCIs do not decode the imagined movement trajectories, but the general activity at sensorimotor areas during MI. As opposed to non-invasive BCIs, invasive BCIs were already used to decode trajectories of imagined movements and to control robotic arms [2, 3]. On the downside, invasive BCIs have the drawback that they require a major surgical intervention with the risk of infection. Gratifyingly, Bradberry et al. [1] discovered that low-frequency electroencephalography (EEG) signals can be used to decode executed movement trajectories, and also our group decoded 3D hand positions from EEG signals [6]. In this work we decoded frontal and lateral hand movements from brain sources reconstructed from the EEG. Notably, we did not train a decoder using all sources simultaneously and interpreted the decoder weights. This can lead to wrong interpretations as these weights must be seen as a filter and not as a pattern. Instead, we decoded the movements from each brain source separately and calculated the correlation coefficients with the measured movements, and obtained maps showing the involved brain regions when decoding movements.

## 2 Methods

### 2.1 Subjects

We recruited 8 right-handed and 1 left-handed subjects who got compensated for their participation (5 males, 4 females). Most of them had already participated in BCI experiments. Subjects sat comfortably in a chair with their arms supported by arm rests.

## 2.2 Paradigm

We recorded 3 runs. In the first run (frontal run) subjects moved their right arm in front of them while the gaze was fixated on a cross on a screen. The second run (lateral run) was similar to the first run, except that subjects moved their right arm laterally. In the third run (ball run), subjects moved their right arm in front of them again, but now observed and followed with the eyes a moving ball on the computer screen. Subjects were instructed to execute the arm movements independently from the ball movements. In all runs, we asked subjects to execute round, natural movements (not jaggy) with the extended right arm, and to keep the hand closed with the thumb being on the upper side. Each run comprised of 8 65s long trials, with subject specific breaks between the trials to avoid muscle fatigue (usually around 1 minute). A trial started with the presentation of a cross (run 1 and 2) or a ball (run 3), respectively, on the computer screen. Two seconds later a beep chimed indicating to the user to start with the movement. In run 3, also a ball on the screen started to move. Additionally to the arm movement trials, we recorded two trials where we instructed the subjects to follow a moving ball on the screen with the eyes, but avoid any arm movement. These two trials were used to remove the influences of eye movements from the EEG with a linear regression method [8], and to calculate the noise covariance matrix used for source imaging (after removal of eye movements).

## 2.3 Recording

We recorded the EEG with 68 passive Ag/AgCl electrodes covering frontal, central and parietal areas, and the electrooculogram (EOG) with 3 electrodes placed above the nasion and the two outer canthi of the eyes. Reference was placed on the left mastoid, ground on the right mastoid. We assured that all impedances were below 5 kOhm. All biosignals were recorded with g.USBamp amplifiers (g.tec medical engineering GmbH, Schiedlberg, Austria). We applied an 8-th order Butterworth bandpass filter with cut off frequencies at 0.01 Hz and 100 Hz, a notch filter at 50 Hz, and then sampled the signals with 512 Hz. The position of the right hand was tracked with a Kinect sensor device (Microsoft, Redmond, US). Here, the x-axis was orientated leftward, y upward, and z backward with respect to the subject. We also recorded the electrode positions with a CMS 20 EP system (Zebris Medical GmbH, Isny, Germany).

## 2.4 Preprocessing

We computed the independent component analysis for each run in the frequency range 0.3 Hz – 70 Hz, using the extended infomax algorithm [4], and removed independent components suspected to be muscle or technical artefacts. Subsequently, we applied a zero-phase anti-aliasing filter and downsampled data to 16 Hz for computational convenience. Then we applied a zero-phase 4-th order Butterworth band-pass filter with cutoff frequencies interesting for decoding at 0.2 Hz and 2 Hz. Afterwards, we removed influences of eye movements on the EEG with a linear regression method [8], and removed samples exceeding a threshold of 5.9 times the median absolute deviation (MAD) of a channel to get rid of remaining artefacts. MAD is a robust deviation measure, and the threshold corresponds to 4 times the standard deviation when the data are normally distributed. Furthermore, we filtered the measured positions with the same band-pass filter as used for the EEG (i.e. 0.2 Hz – 2 Hz), and centered and scaled them to a standard deviation of one. Finally, we omitted the first 5 seconds of each trial to exclude possible existing movement onset effects.

## 2.5 Source Imaging

To transform the data from the sensor space into the source space we used the software Brainstorm [9]. We calculated the head model using the Colin27 model included in Brainstorm, and coregistered the electrode positions. Using the head model and a noise covariance matrix, we calculated 15028 brain sources with the weighted minimum norm estimation (wMNE) method. The noise covariance matrix was calculated from the two trials without arm movements after we removed eye movements from them.

## 2.6 Decoding

We analyzed all runs separately using a 10-fold cross-validation. For this purpose we divided each run in segments of 10s length and assigned these segments to train and test sets. The decoder [6] itself comprised of 2 multiple linear regressions between a brain source (voxel) and the 2D position in the movement plane (frontal or lateral, respectively). We used the current time step and time lags at ca. 60, 130, 190ms of the EEG as the input for the decoder. X and y positions were decoded in the frontal and ball runs, and y and z positions in the lateral run. We decoded the positions from the test sets, and calculated the Pearson correlation coefficients between the decoded and the measured 2D positions. Subsequently, we calculated the average of the correlation coefficients across the 10 test sets. Finally, we averaged the correlation coefficients over the movement plane dimensions, i.e. x/y or y/z, respectively. This procedure was performed for every single voxel, and we got one correlation coefficient for each voxel, run and subject.

To assess the chance level, we randomly permuted the coordinate segments and performed a 10-fold cross-validation as described above and repeated this procedure 50 times. Thus, we got 50 correlation coefficients for each voxel, run and subject, and then, fitted a normal distribution to these 50 chance correlation coefficients (this is reasonable as the correlation values are usually around 0 and not at the limits 1/-1). Subsequently, we calculated the p-values of the correlation coefficients based on the chance level distributions.

## 2.7 Results

The maximum correlation coefficient reached by a subject averaged over all subjects were (mean value/standard deviation)  $0.47 \pm 0.09$  (frontal),  $0.52 \pm 0.13$  (lateral), and  $0.47 \pm 0.10$  (ball). The corresponding chance level correlations were  $0.12 \pm 0.03$  (frontal),  $0.12 \pm 0.02$  (lateral), and  $0.11 \pm 0.02$  (ball). Figure 1 shows the subject averaged correlations of each voxel in each run and their average over the runs. Before averaging, all non-significant correlations were set to 0. Observable are higher correlations on the central and left motor cortex, parietal areas, and right lateral correlations.

## 3 Discussion

We have successfully decoded executed movements from frontal and lateral arm movements on a per brain source (voxel) basis. The subject averaged correlations indicate a contribution of the primary motor cortex. This was expected, as subjects executed movements. However, also contributions from parietal and lateral areas are observable. These contributions can be external sources projected onto the margins of the head model. Such an external source could be muscle activity, although muscle activity is thought to be most prominent in higher

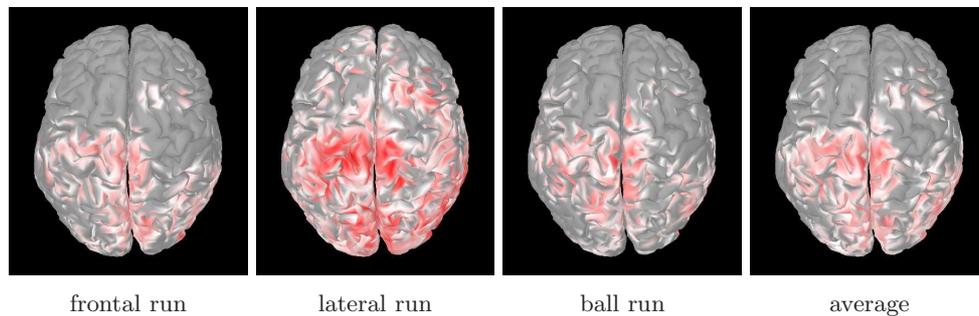


Figure 1: Subject averaged correlations on a per voxel basis for all 3 runs and the average of them. Red corresponds to the maximum value of 0.33 in the lateral run, white to 50 % of the maximum value, and correlations below 50 % of the maximum are not shown.

frequency ranges. To summarize, the findings indicate that indeed brain sources carry decodable movement information, but the measurements are potentially contaminated by external sources.

## References

- [1] Trent J Bradberry, Rodolphe J Gentili, and José L Contreras-Vidal. Reconstructing three-dimensional hand movements from noninvasive electroencephalographic signals. *Journal of Neuroscience*, 30:3432–3437, 2010.
- [2] Jennifer L Collinger, Brian Wodlinger, John E Downey, Wei Wang, Elizabeth C Tyler-Kabara, Douglas J Weber, Angus J C McMorland, Meel Velliste, Michael L Boninger, and Andrew B Schwartz. High-performance neuroprosthetic control by an individual with tetraplegia. *Lancet*, 381(9866):557–564, 2013.
- [3] Leigh R Hochberg, Daniel Bacher, Beata Jarosiewicz, Nicolas Y Masse, John D Simeral, Joern Vogel, Sami Haddadin, Jie Liu, Sydney S Cash, Patrick van der Smagt, and John P Donoghue. Reach and grasp by people with tetraplegia using a neurally controlled robotic arm. *Nature*, 485:372–375, 2012.
- [4] T W Lee, M Girolami, and T J Sejnowski. Independent component analysis using an extended info-max algorithm for mixed sub-gaussian and super-gaussian sources. *Neural Computation*, 11(2):417–441, 1999.
- [5] G R Müller-Putz, R Scherer, G Pfurtscheller, and R Rupp. EEG-based neuroprosthesis control: a step towards clinical practice. *Neuroscience Letters*, 382:169–174, 2005.
- [6] Patrick Ofner and Gernot R Müller-Putz. Decoding of velocities and positions of 3d arm movement from eeg. In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pages 6406–6409, 2012.
- [7] Gert Pfurtscheller and Fernando H Lopes Da Silva. Event-related eeg/meg synchronization and desynchronization: basic principles. *Clinical Neurophysiology*, 110(11):1842–1857, 1999.
- [8] A Schlögl, C Keinrath, D Zimmermann, R Scherer, R Leeb, and G Pfurtscheller. A fully automated correction method of eeg artifacts in eeg recordings. *Clinical Neurophysiology*, 118(1):98–104, 2007.
- [9] François Tadel, Sylvain Baillet, John C Moshier, Dimitrios Pantazis, and Richard M Leahy. Brainstorm: A user-friendly application for meg/eeg analysis. *Computational Intelligence and Neuroscience*, 2011.

# BCI-controlled Wheelchair; Audio-cued Motor Imagery-based Paradigm

Francisco Velasco-Alvarez<sup>1</sup>, Sergio Varona-Moya<sup>1</sup>, María José Blanca-Mena<sup>2</sup>, Salvador Sancha-Ros<sup>1</sup> and Ricardo Ron-Angevin<sup>1</sup>

<sup>1</sup>Dept. of Electronic Technology, University of Malaga, Spain.

<sup>2</sup>Dept. of Psychobiology and Methodology of Behavioral Sciences, University of Malaga, Spain.

fvelasco@dte.uma.es, rra@dte.uma.es

## Abstract

In this study we present a control paradigm that enables subjects to drive a real wheelchair using four navigation commands. The control is achieved through the discrimination of two mental tasks (relaxed state versus one active Motor Imagery task), in order to reduce the risk of misclassification. After a short training, the paradigm becomes only audio-cued, thus avoiding the need of a graphical interface that could distract subjects' attention.

## 1 Introduction

A brain-computer interface (BCI) is based on analysis of the brain activity, recorded during certain mental activities, in order to control an external device. It helps to establish a communication and control channel for people with serious motor function problems but without cognitive function disorder (Wolpaw, Birbaumer, McFarland, Pfurtscheller & Vaughan, 2002). Amyotrophic lateral sclerosis (ALS), brain or spinal cord injury, cerebral palsy and numerous other diseases impair the neural pathways that control muscles or impair the muscles themselves. Among others ways to control such a system, Sensorimotor rhythm-based BCIs (SMR-BCI) are based on the changes of  $\mu$  and  $\beta$  rhythms. These rhythms correspond to specific features of the EEG signals characterized by their frequencies that can be modified by voluntary thoughts: when a person performs a movement (or imagines it), it causes a synchronization/desynchronization in this activity (event-related synchronization/desynchronization, ERS/ERD) which involves a rhythm amplitude change (Neuper & Pfurtscheller, 1999).

Recently, a survey on BCI-controlled devices has been published (Bi, Fan, & Liu, 2013) which shows a wide variety of systems controlling wheelchairs or mobile robots. Most of the SMR-BCI there reported match the number of commands to the number of mental tasks. However, incrementing the number of mental tasks can reduce the classification performance (Kronegg, Chanel, Voloshynovskiy & Pun, 2007; Obermaier, Neuper, Guger, & Pfurtscheller, 2001). In order to provide different commands without making the BCI performance worse, our group proposed a paradigm based on the discrimination of only two classes (one active mental task versus any other mental activity), which enabled the selection of four navigation commands (Ron-Angevin, Díaz-Estrella & Velasco-Álvarez, 2009). The evolution of this paradigm included its adaptation to a self-paced BCI enabling a Non-Control (NC) state (Velasco-Álvarez & Ron-Angevin, 2009), the support for continuous movements (Velasco-Álvarez, Ron-Angevin & Blanca-Mena, 2010) or its conversion into an audio-cued paradigm (Velasco-Álvarez, Ron-Angevin, da Silva-Sauer, Sancha-Ros & Blanca-

Mena, 2011). This control paradigm has already been tested for virtual and real environments (Velasco-Álvarez, Ron-Angevin, da Silva-Sauer & Sancha-Ros, 2013); once it has proved to be valid for controlling a virtual wheelchair or a mobile robot, our aim is to adapt it to the use of a real wheelchair. The use of an audio-cued paradigm is important, as a graphical interface could limit the subjects' field of view and, at the same time, distract them from the task of controlling the wheelchair by requiring them to look at the computer screen. Furthermore, for disabled people without gaze control a graphical interface may not be useful.

## 2 Methods

### 2.1 Data Acquisition, Initial Training and Signal Processing

In our system, the EEG is recorded from ten active electrodes, combined to result two laplacian channels around C3 and C4 positions (right and left hand sensorimotor area, respectively) according to the 10/20 international system. The ground electrode is placed at the FPz position. Signals are amplified and digitized at 200 Hz by an actiCHamp amplifier (Brain Products GmbH, Munich, Germany).

Subjects have to follow an initial training that consists of two sessions: a first one without feedback and a second one providing continuous feedback. These two training sessions are used for calibration purposes. The training is the same as the one used in (Ron-Angevin, 2009) and is based on the paradigm proposed by the Graz group (Leeb, Settgast, Fellner & Pfurtscheller, 2007). It consists of a virtual car that dodges a water puddle in the road, by moving left or right according to the mental task carried out (relaxed state or right hand MI). An offline processing of the first session determines the parameters for the feedback session. The same parameters are used to calibrate the system for the navigation sessions. This processing is based on the procedure detailed in (Guger, Edlinger, Harkam, Niedermayer & Pfurtscheller, 2003), and consists of estimating the average band power of each channel in predefined, subject-specific reactive frequency (manually selected) bands at intervals of 500 ms. In the feedback session, the movement of the car is computed on-line every 25 ms as a result of a Linear Discriminant Analysis (LDA) classification. The trial paradigm and all the algorithms used in the signal processing are implemented in MATLAB 2013b (The MathWorks Inc., Massachusetts, USA).

### 2.2 Navigation Paradigm

The training procedure to control the paradigm consists of a first navigation session in which subjects control a virtual device with the help a graphical interface showed simultaneously with presence of the audio cues. Once the users become familiar with it, it is replaced by an only audio-cued interface.

The procedure to control the device is similar to the one used in (Velasco-Álvarez et al., 2013): the system waits in a NC state in which an NC interface is shown: a semi-transparent vertical blue bar placed in the centre of the screen. By extending the bar carrying out the MI task (above a subject-dependant selection threshold for a given selection time), subjects can switch to the Intentional Control (IC) state, where the control is achieved through the IC interface. The IC interface consists of a circle divided into 4 parts, which correspond to the possible navigation commands (move forward/backward, turn right/left), with a bar placed in the centre of the circle that is continuously rotating clockwise. The subject can extend the bar by carrying out the MI task in order to select a command when the bar is pointing at it (same selection threshold and time). Subjects receive audio cues while they interact with the system. When the state switches from IC to NC (which occurs when

the rotating bar completes two turns without selecting any command), they hear the Spanish word for "wait"; the reverse switch is indicated with "forward", since it is the first available command in the IC state. Finally, every time the bar points to a different command, they can hear the correspondent word ("forward", "right", "back" or "left").

### 2.3 Robotic Wheelchair

We customized a robotic Invacare Mistral Plus electric wheelchair. The wheelchair is modified in such a manner that its direction and speed can be controlled with a computer, and sensors are included that give the software information about the wheelchair's status and environment.

A custom electronic board that emulates an analog 2-axis joystick is needed, which is attached to the wheelchair control board replacing its actual joystick. Via a USB/Bluetooth interface, an external computer can communicate with that board and set the virtual joystick position in real time. Besides, the board includes an IIC bus that allows a number of sensors, such as a magnetometer, an accelerometer and a set of sonars, to be attached to the main system. The status of these sensors can also be read from the USB/Bluetooth interface.

This board is controlled by an application (running on the same computer that processed the EEG signal, or on a different one connected via TCP). This application is responsible for translating high-level navigation commands from the BCI interface (move forward, turn right, etc.) into virtual joystick positions, which in turn makes the wheelchair move. The attached sensors allow for safe and reliable navigation. After an initial calibration stage, the magnetometer can be used as a digital compass which the application use to instantly correct small variations of the direction when moving the wheelchair forward or backwards, or to perform discrete turns to a specific angle. The accelerometer, for its part, is used to detect steep slopes and reduce the wheelchair speed, or to stop it when a potentially dangerous tilt was reached.

The set of sonars allow to create an estimated real-time map of the environment. The application keeps a discrete grid which describes the likelihood of a certain small area around the wheelchair being occupied. When one of the sonars detects an obstacle at a given distance, all the grid cells within its detection cone are updated according to a model of the sensor.

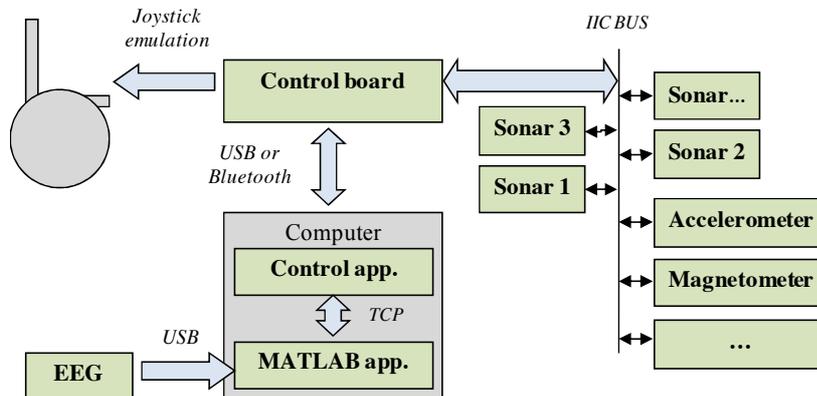


Figure 1: Module distribution for the BCI-controlled wheelchair

### 3 Work in Progress

Our group is nowadays preparing experiments to test the viability of the proposed system, so the given description of the system corresponds to a work in progress. As long as the navigation paradigm has already been validated, we expect that, with the proper training and adaptation sessions, the change from a mobile robot to a real wheelchair will not severely affect the performance.

### 4 Acknowledgements

This work was partially supported by the Spanish Ministry of Economy and Competitiveness, project TEC 2011-26395, the European fund ERDF and the University of Málaga.

### References

- Bi, L., Fan, X. -, & Liu, Y. (2013). EEG-based brain-controlled mobile robots: A survey. *IEEE Transactions on Human-Machine Systems*, 43(2), 161-176. doi:10.1109/TSMCC.2012.2219046
- Guger, C., Edlinger, G., Harkam, W., Niedermayer, I., & Pfurtscheller, G. (2003). How many people are able to operate an EEG-based brain-computer interface (BCI)? *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11(2), 145-147.
- Kronegg, J., Chanel, G., Voloshynovskiy, S., & Pun, T. (2007). EEG-based synchronized brain-computer interfaces: A model for optimizing the number of mental tasks. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 15(1), 50-58.
- Leeb, R., Settgast, V., Fellner, D., & Pfurtscheller, G. (2007). Self-paced exploration of the austrian national library through thought. *International Journal of Bioelectromagnetism*, 9(4), 237-244.
- Neuper, C., & Pfurtscheller, G. (1999). Motor imagery and ERD. In G. Pfurtscheller, & F. H. Lopes da Silva (Eds.), *Event-related desynchronization handbook of electroencephalography and clinical neurophysiology, revised series* (pp. 303-325). Amsterdam: Elsevier.
- Obermaier, B., Neuper, C., Guger, C., & Pfurtscheller, G. (2001). Information transfer rate in a five-classes brain-computer interface. *IEEE Trans Neural Syst Rehabil Eng*, 9(3), 283-288.
- Ron-Angevin, R. (2009). Changes in EEG behavior through feedback presentation. *Annual Review of CyberTherapy and Telemedicine*, 7(1), 184-188.
- Ron-Angevin, R., Díaz-Estrella, A., & Velasco-Álvarez, F. (2009). A two-class brain computer interface to freely navigate through virtual worlds. *Biomedizinische Technik*, 54(3), 126-133.
- Velasco-Álvarez, F., & Ron-Angevin, R. (2009). *Asynchronous brain-computer interface to navigate in virtual environments using one motor imagery*, LNCS, 5517 (I), 698-7058.
- Velasco-Álvarez, F., Ron-Angevin, R., & Blanca-Mena, M. J. (2010). Free virtual navigation using motor imagery through an asynchronous brain-computer interface. *Presence: Teleoperators and Virtual Environments*, 19(1), 71-81.
- Velasco-Álvarez, F., Ron-Angevin, R., da Silva-Sauer, L., Sancha-Ros, S., & Blanca-Mena, M. J. (2011). *Audio-cued SMR brain-computer interface to drive a virtual wheelchair*. LNCS, 6691 (I), 337-344.
- Velasco-Álvarez, F., Ron-Angevin, R., da Silva-Sauer, L., & Sancha-Ros, S. (2013). Audio-cued motor imagery-based brain-computer interface: Navigation through virtual and real environments. *Neurocomputing*, 121(0), 89-98. doi:<http://dx.doi.org/10.1016/j.neucom.2012.11.038>
- Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain-computer interfaces for communication and control. *Clinical Neurophysiology*, 113(6), 767-791.

# Switching Characters between Stimuli improves P300 Speller Accuracy

Thibault Verhoeven, Pieter-Jan Kindermans, Pieter Buteneers and Benjamin Schrauwen

Ghent University, Ghent, Belgium  
thibault.verhoeven@ugent.be

## Abstract

In this paper, an alternative stimulus presentation paradigm for the P300 speller is introduced. Similar to the checkerboard paradigm it minimizes the occurrence of the two most common causes of spelling errors: adjacency distraction and double flashes. Moreover, in contrast to the checkerboard paradigm, this new stimulus sequence does not increase the time required per stimulus iteration. Our new paradigm is compared to the basic row-column paradigm and the results indicate that, on average, the accuracy is improved.

## 1 Introduction

The P300 speller, as first described by Farwell & Donchin in [1], is a BCI for spelling text. In the original speller in [1] the six rows and six columns of the grid are highlighted in a random order. The target element is the element on the intersection of the row and the column that elicited a P300 as response when highlighted. A sequence containing every stimulus once is called an ‘iteration’. This row-column paradigm (RC) thus has an iteration length of 12 stimuli.

Two common phenomena limit the speller accuracy. In ‘adjacency distraction’ the patient is distracted by the intensification of a row or column next to the actual target. ‘Double flash’ (DF) errors occur when the target element is intensified twice in rapid succession. The second P300 wave generated has a lower amplitude [2] which can result in a wrong target selection.

To alleviate these effects, several stimulus sequence paradigms have been examined, for example the checkerboard paradigm [4]. Unfortunately, the improvements to the accuracy were diminished by an increased time required per iteration. New paradigms always seem to make this compromise between accuracy and speed of spelling. The goal of this paper is to present a paradigm that achieves a higher accuracy than RC with a speed of spelling as high as in RC.

In what follows, the new paradigm is explained in detail together with the simulation results. The performance of this paradigm is compared with RC in a first online experiment.

## 2 The Switching Paradigm

For every character to spell, the new paradigm constructs a pre-determined sequence of stimulus patterns through an optimization algorithm. This algorithm takes a full sequence of stimuli (in this paper: a sequence of 15 iterations constructed by the RC paradigm) as initialisation and tries to reduce the potential causes of spelling errors by swapping the highlighted elements between stimuli. If the initial sequence, for example, contains the subsequence of stimuli  $S_1$ - $S_2$ - $S_3$  in which the sets of characters  $\{A, B, C, D, E, F\}$ ,  $\{E, K, Q, W, 3, 9\}$  and  $\{C, I, O, U, 1, 7\}$  are respectively highlighted, there is clearly a DF of character E in the consecutive stimuli  $S_1$  and  $S_2$ . Swapping the characters E and U between the stimuli  $S_2$  and  $S_3$  changes the sequence to  $\{A, B, C, D, E, F\}$ - $\{U, K, Q, W, 3, 9\}$ - $\{C, I, O, E, 1, 7\}$  and eliminates the DF.

Double flashes are thus easily avoided by preventing elements to be highlighted twice in rapid succession. Adjacency distraction itself can not be avoided. However, when the target element is simultaneously highlighted with a neighbouring element, the generated P300 wave on this target stimulus is also linked to this neighbour. A single distraction by a following intensification of this neighbour is enough to wrongly select this neighbour as the target. Adjacency distraction errors thus can be partly avoided by preventing an element to be highlighted simultaneously with a neighbouring element in the grid. The subject then needs to be distracted twice to a neighbour before wrong selection, which is less likely to occur. This results in the reported decrease of adjacency distraction errors in the checkerboard paradigm [4].

Not every two elements from every two stimuli can be swapped. If two elements are highlighted in the same set of stimuli in the sequence and these stimuli appear to be target stimuli, then it is impossible to determine which one of the two is the target element. The sequence is said to be ‘undecodable’. A swap that does not harm this decodability is a ‘legal swap’. Only these swaps will be considered further on.

The algorithm works in two phases. The first phase tries to remove the most severe causes of mistakes, while the second phase optimizes the sequence further on. As each optimization in one phase creates new optimization opportunities in the other, these phases are alternately executed until no more optimizations are found. We allow a maximum of 5 alternate executions to prevent the algorithm from getting stuck in a loop, constantly swapping the same elements.

For every iteration in the sequence separately, every possible pair of stimuli is examined in random order. For every pair of stimuli, every legal swap of highlighted elements between these stimuli is examined in random order. Two criteria are used to verify if the swap is indeed optimizing the sequence. In the first phase these criteria are:

1. For both elements: does the swap reduce the number of DFs? If this number is not altered, does the swap reduce the second order DFs (DFs with one stimulus in between)?
2. For both stimuli: does the swap reduce the number of direct horizontal/vertical neighbouring elements in the grid simultaneously intensified? If this number is not altered, does the swap reduce the number of direct diagonal neighbours?

In the second phase these criteria are:

1. For both elements: does the distance in the sequence to the closest stimulus in which this element is highlighted increase?
2. For both stimuli: is the distance in the grid of the swapped element to the closest highlighted element bigger, i.e. does the spread of highlighted elements in the grid increase?

If at least one of these questions has a strict positive answer and none of the others is strictly negative, then the swap is executed. After considering all legal swaps, the procedure is repeated all over again until no more optimizations are found and the algorithm moves to the next phase.

A total of 500 row-column sequences of 15 iterations each are created and compared to their counterparts after applying the algorithm. The number of target double DFs seen by the subject in a sequence of 15 iterations is shown in the histograms in Figure 1a. On average, 2.71 ( $\sigma = 1.49$ ) double flashes are noticed in a standard RC sequence. After applying the algorithm, this reduces by 98% to only 0.05 ( $\sigma = 0.23$ ) double flashes on average.

The probability of adjacency distraction is reflected in the number of times the subject sees the target being flashed simultaneously with at least one of its direct (horizontal or vertical) neighbours in the grid. In the RC paradigm the target is always flashed with a neighbour. After applying the algorithm only 1.32% of the target intensifications occur together with a neighbour.

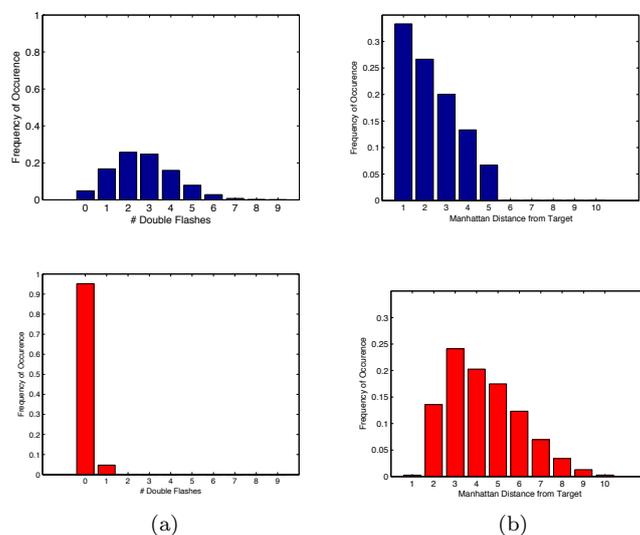


Figure 1: Simulation results. Left: normalized histogram of the number of double flashes seen in the sequence for the RC paradigm (top) and the new paradigm (bottom). Right: normalized histogram of the Manhattan distance in the grid from the target to every other element for RC (top) and the new paradigm (bottom).

Figure 1b gives a histogram of the Manhattan distance in the grid from the target to every other element simultaneously highlighted. After application of the algorithm the highlighted characters are clearly spread more over the grid.

### 3 Experimental Setup & Online Experiment

After presentation of a stimulus sequence to the subject, the recorded brain signals are preprocessed. A Common Average Reference filter is applied, followed by a bandpass filter with cutoff frequencies 0.5 Hz and 15 Hz. Each EEG channel is normalized to zero mean and unit variance. Dimensionality reduction retains 10 samples per stimulus, centered around the expected time step of the P300 wave and uniformly distributed over the range between 150ms and 450ms after stimulus onset. Finally, we add a bias term. The stimuli are classified as target/non-target by detecting the presence/absence of a P300 wave in the response. The classifier used in this work is based on the unified probabilistic model proposed by Kindermans et al. in [3].

9 subjects (6 male, 3 female) with an average age of 22.56 ( $\sigma = 0.73$ ) used the original RC paradigm and the switching paradigm (SP) to spell the 36 characters of the grid in random order. Both the order of the characters and the order in which the paradigms are used are altered between subjects. In the calibration procedure, 10 characters are spelled. Calibration is done separately for both paradigms. The stimulus and interstimulus time are 62.5 ms and 125 ms respectively. Every sequence of stimuli contains 15 iterations. Apart from subject 7, the two tests were done on different days. Only subjects 5 and 8 had previous experience with the P300 speller. The Emotiv EPOC headset is used to record the EEG.

The accuracy of the speller is defined as the fraction of characters correctly spelled. Accu-

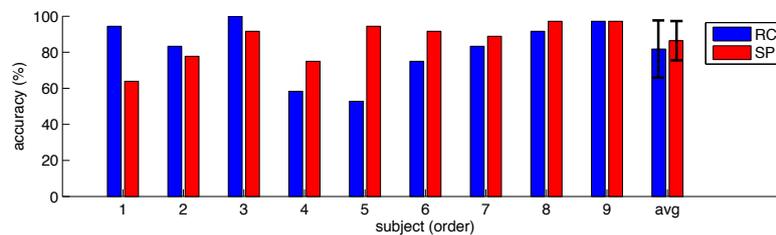


Figure 2: Spelling accuracies obtained with both paradigms for the 9 subjects separately, and their average result. For the first 4 subjects the RC paradigm is used first, for the other 5 SP.

racies for both paradigms can be found in Figure 2. Subject 1 to 4 used the RC paradigm first, the others started with the SP paradigm. A mixed-design ANOVA shows the larger effect of the order in which the paradigms are used ( $F_{between-subjects} = 0.81$ ) compared to the effect of the paradigm itself ( $F_{within-subjects} = 0.337$ ). A possible cause is that, once a subject gets used to a paradigm, it may be harder for him to adapt to a new paradigm. A more thorough examination will be done in the future in which the effect of the order is reduced by letting the subjects get more familiar with the speller and the paradigms before passing on to the actual experiment. The average accuracy over all subjects when using the RC paradigm is 81.79% compared to 86.42% when using SP. This indicates that the newly developed paradigm can potentially achieve a higher accuracy than RC while the time per iteration is still 2.25 s.

## 4 Conclusion

The first online results with the new paradigm are promising, although not convincing. On average, the accuracy is increased while the iteration length and thus the speed of spelling remain the same. Starting from these pilot tests, extensive testing with real EEG caps will have to show whether the new paradigm significantly improves the accuracy of the classifier.

## Acknowledgment

This project is funded by The Special Research Fund of Ghent University.

## References

- [1] 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*, 70:510–523, 1988.
- [2] C.J. Gonsalvez and J. Polich. P300 amplitude is determined by target-to-target interval. *Psychophysiology*, 39:388–396, 2002.
- [3] P. Kindermans, H. Verschore, and B. Schrauwen. A unified probabilistic approach to improve spelling in an event-related potential based brain-computer interface. *IEEE transactions on biomedical engineering*, 60(10):2696–2705, 2013.
- [4] G. Townsend, B.K. LaPallo, C.B. Boulay, D.J. Krusienski, G.E. Frye, C.K. Hauser, N.E. Schwartz, T.M. Vaughan, J.R. Wolpaw, and E.W. Sellers. A novel P300-based brain-computer interface stimulus presentation paradigm: moving beyond rows and columns. *Clinical Neurophysiology*, 121(7):1109–1120, July 2010.

# Influence of P300 latency jitter over (c)overt attention BCIs

Pietro Arico<sup>1,2\*</sup>, Fabio Aloise<sup>1†</sup>, Francesca Schettini<sup>1,2‡</sup>, Serenella Salinari<sup>2§</sup>,  
Donatella Mattia<sup>1\*\*</sup> and Febo Cincotti<sup>2††</sup>

<sup>1</sup> Neuroelectrical Imaging and BCI Lab, Fondazione Santa Lucia IRCCS, Rome, Italy.

<sup>2</sup> Department of Computer, Control, and Management Engineering, Sapienza University of Rome, Italy.

p.arico@hsantalucia.it

## Abstract

Several ERP-based BCIs that can be controlled even without eye movements (covert attention) have been recently proposed. However, when compared to similar systems based on overt attention, they displayed a significantly lower accuracy. In the current interpretation, this is ascribed to the absence of the contribution of short-latency visual evoked potentials (VEPs) in the tasks performed in the covert attention modality. This study aims to investigate if this decrement (i) is fully explained by the lack of VEP contribution to the classification accuracy; (ii) correlates with lower temporal stability of the single-trial P300 potentials elicited in the covert attention modality. We evaluated the latency jitter of P300 evoked potentials in three BCI interfaces exploiting either overt or covert attention modalities in 20 healthy subjects. Results highlighted that the P300 jitter is higher when the BCI is controlled in covert attention and classification accuracy negatively correlates with jitter.

## 1 Introduction

The Farwell and Donchin's P300 Speller (1988) is among the most widely validated Brain Computer Interface (BCI) paradigms for communication applications. Brunner and colleagues (2010) have recently shown that the P300 Speller recognition accuracy was significantly decreased if the subject was not allowed to gaze at the target stimulus. Several user interfaces designed to be used in covert attention modality have been implemented and tested with the overall result of a lower system performance in covert with respect to overt attention usage. The observed superiority in the system performances under overt usage modality was mainly ascribed to the contribution of visual evoked potential (VEP) components recorded at occipital and parieto-occipital sites (Aloise et al., 2012; Treder and Blankertz, 2010). Another important contribution regards the P300 latency jitter that occurs when the lag between each target stimulus onset and the related potential peak is not constant for the different stimulus repetitions. Thompson and colleagues (2013) demonstrated that the accuracy

---

\* Wrote the manuscript and followed the experimental protocol

† Developed the stimulation interface used in the study

‡ Participated to the performed analyses

§ Revised the document before the submission

\*\* Revised the document before the submission

†† Revised the document before the submission

achieved with the P300 Speller was strongly correlated with the jitter in the P300 latency. In this study we addressed the issue of whether the accuracy of BCIs used in covert attention modality i) is fully explained by the lack of VEP contribution to the classification accuracy and/or ii) is correlated with a lower stability of the P300 potential elicited in the covert attention with respect to the overt attention modality.

## 2 Materials and methods

Twenty healthy volunteers (14F and 6M,  $28 \pm 5$  years) were requested to complete a spelling task using a BCI. For this purpose, visual stimuli containing 36 alphanumeric characters for the GeoSpell (used in covert attention, Aloise et al., 2012) and the P300 Speller (Farwell and Donchin, 1988) interface (used in overt attention), and 2 characters for a simple Visual Oddball interface used in overt attention, were delivered in different arrangements, through three alternative visual interfaces. For all interfaces, the frequency of target stimuli was 16.7% (i.e. 1/6).

Scalp EEG signals were recorded (g.USBamp, gTec, Austria) from 8 Ag/AgCl electrodes (Fz, Cz, Pz, Oz, P3, P4, PO7 and PO8, referenced to the right earlobe and grounded to the left mastoid) at 256 Hz. Visual stimulation and acquisition were operated by means of the BCI2000 software. At the beginning of each trial the system suggested to the subject the character to be written before the stimulation started. Recordings took place in two sessions on separate days. Each session consisted of 3 runs for each interface and 6 trials (i.e. characters) per run. Subjects were required to spell 6 words (3 words per session) using both the GeoSpell and the P300 Speller interfaces; subjects were required to spell the sequence “OOOOOO” (all ‘rare’ stimuli) using the Visual Oddball interface. This latter sequence was repeated for six runs. Each trial consisted of 8 stimulation sequences and corresponded to the selection of a single character displayed on the interface. Each character was intensified for 125ms (Stimulus duration), with an Inter Stimulus Interval (ISI) of 125ms.

The EEG signals were segmented into 800 ms overlapping epochs following the onset of each stimulus. Two runs of each recording session were considered as training set while the remaining run provided the data for the testing set, exploring all possible permutations.

*P300 latency jitter evaluation:* To evaluate the influence of the P300 latency jitter on the classification accuracy, it was necessary to reconstruct the P300 potential waveform for each single epoch. In this regard, we applied a method based on the Continuous Wavelet Transform (CWT) and the estimation of the empirical Cumulative Distribution Function (CDF), in order to enhance the signal (P300) to noise (spontaneous EEG) ratio (Aricò et al, 2014). At this point, we calculated the inverse CWT for each epoch, and we estimated the latency of the P300 potential as the highest peak of the signal into the epoch. The latter had been manually selected from the averaged waveforms, to embrace the whole P300 shape. We quantified the jitter of the P300 latency as the difference between the 3<sup>rd</sup> and the 1<sup>st</sup> quartile of each distribution for each testing run. These analyses were performed on the Cz channel, where the P300 is most prominent.

*BCI accuracy evaluation:* For each participant, we assessed the BCI accuracies offline, as a function of the number of stimulation sequences averaged during each trial. We used a Stepwise Linear Discriminant Analysis (SWLDA, Aloise et al., 2012) to select the most relevant features that allowed to discriminate between target and non-target stimuli. We performed a three-fold cross-validations exploring all possible combinations of training (2 runs) and testing (1 run) data set for each session and interface. We evaluated the performance of the subjects for each interface considering i) [*Whole epoch*] the entire time length of the epoch (0-800ms), ii) [*Whole epoch decimated*] same epoch length as above, reducing by a factor of 12 the number of time samples (each new sample is the average of 12 original samples), iii) [*P300 epoch non-realigned*] only the epoch segment containing the P300 potential thus disregarding those VEPs components influenced by

gazing at the target stimuli, iv) [P300 epoch realigned] same epoch length as above, using potentials obtained after realignment of the single epochs, whose time courses were shifted according to the estimated P300 latency values.

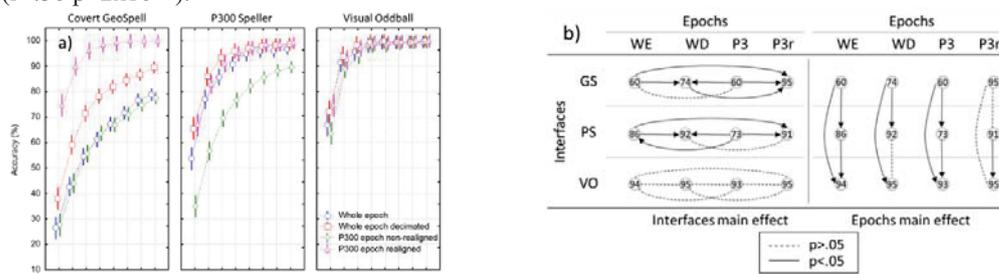
*Correlation between P300 latency jitter and performance:* The information transfer rate (ITR, bit/min) was calculated at each fold of cross-validation as a function of the number of sequences in the trial. In particular, we calculated the mean value of the ITR along the 8 stimulation sequences, in order to have a synthetic measure of the system’s performance. To assess the correlation between the  $ITR_{Mean}$  and the P300 latency jitter, we estimated the non-parametric Spearman’s rank correlation coefficient between these variables.

### 3 Results

*P300 latency jitter:* Significant differences of P300 latency jitter elicited by the 3 interfaces were explored by means of one-way repeated measures ANOVAs (Confidence Interval = .95) where interface was considered as factor and the P300 latency jitter as dependent variables. The analysis revealed a significant difference across the interfaces for the jitter ( $F(2, 357)=52.58; p=9 \times 10^{-6}$ ). Post-hoc analysis (Duncan test) showed that the GeoSpell produced a latency jitter significantly larger than the P300 Speller and the Visual Oddball (mean:  $108 \pm 24ms$ ,  $76 \pm 24ms$ , and  $74 \pm 38ms$ , respectively;  $p < 10^{-4}$ ).

*BCI accuracy:* Differences in the classification accuracy achieved with each of the 3 visual interfaces and each of the 4 conditions introduced previously (Figure 1a). A two-way repeated measures ANOVA (Confidence Interval = .95) was performed with interfaces and conditions as factors and the accuracy per stimulation sequences as dependent variables. The analysis revealed a significant interaction between the factors ( $F(6, 1428)=42.57; p=10^{-9}$ ). The Duncan's multiple range test was used for post hoc comparison. The differences in the epoch choices and the interfaces are summarized in Figure 1b.

*Correlation between P300 latency jitter and performance:* The non-parametric Spearman’s rank correlation coefficient was used to evaluate the correlation between the classification accuracy as expressed by the  $ITR_{Mean}$  values and the P300 latency jitter obtained for each interface. We found a significant negative correlation between the latency jitter and the accuracy achieved by the subjects with all 3 interfaces (GeoSpell:  $r=.17 p=.04$ ; P300 Speller:  $r=.35 p=10^{-4}$ ; Visual Oddball:  $r=.18 p=.03$ ). Considering all the interface together, we found a significant negative correlation as well ( $r=.50 p=2 \times 10^{-23}$ ).



**Figure 1:** (a) Mean and confidence intervals (CI = 0.95) of the cross-validation target classification accuracies achieved with the three interfaces, relative to each epoch choice as a function of the number of stimulations; (b) Graphical representation of the differences between the epochs (WE: Whole epoch; WD: Whole epoch decimated; P3: P300 epoch non-realigned; P3r: P300 epoch realigned) and the interfaces (GS: GeoSpell; PS: P300 Speller; VO: Visual Oddball) in terms of accuracy, highlighted by the post hoc test. Numbers in the circles indicate the percent mean accuracy value.

## 4 Discussion

The overall aim of this study was to investigate the influence of the P300 latency jitter evoked during (c)overt attention based BCI tasks on the accuracy achieved. The results proved that when the user operates a BCI using covert attention, the latency jitter is greater than using overt attention. Particularly, the P300 latency jitter was significantly greater when using the GeoSpell interface than using the other interfaces. In line with previous studies (Brunner et al., 2010), our findings on the first phenomenon clearly indicate the significant contribution of the early VEPs to the classification accuracy only for the overt (i.e. P300 Speller) interface. Also, removing the VEP contribution from ERPs elicited using the P300 Speller and the GeoSpell interface, the latter still performed significantly worse than the former, suggesting that the lack of VEPs is not the only reason for the performance decrement in the tasks performed in covert attention modality. In addition, a significant correlation was found between the latency jitter and the BCI performances for all the interfaces.

## 5 Conclusion

We found that (i) even canceling the contribution of short latency VEPs, the P300 Speller interface (used in overt attention modality) remains more accurate than the GeoSpell (used in covert attention); (ii) the P300 latency jitter is negatively correlated with the accuracy of the BCI classifier; (iii) a compensation of the P300 latency jitter makes the GeoSpell (used in covert attention) than the P300 Speller.

## Acknowledgments

The work is partly supported by the EU grant FP7-287774 ABC and NINA research project, co-financed by EUROCONTROL. This paper only reflects the authors' views and funding agencies are not liable for any use that may be made of the information contained herein.

## References

- Aloise, F., Schettini, F., Aricò, P., Salinari, S., Babiloni, F., Cincotti, F., (2012). *A comparison of classification techniques for a gaze-independent P300-based brain-computer interface*. Journal of neural engineering 9, 045012. doi:10.1088/1741-2560/9/4/045012
- Aricò, P., Aloise, F., Schettini, F., Salinari, S., Mattia, D., Cincotti, F., 2014. *Influence of P300 latency jitter on event related potential-based brain-computer interface performance*. J Neural Eng 11, 035008. doi:10.1088/1741-2560/11/3/035008
- Brunner, P., Joshi, S., Briskin, S., Wolpaw, J.R., Bischof, H., Schalk, G., (2010). *Does the "P300" speller depend on eye gaze?* J Neural Eng 7, 056013. doi:10.1088/1741-2560/7/5/056013
- Farwell, L.A., Donchin, E., (1988). *Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials*. Electroencephalogr Clin Neurophysiol 70, 510–523.
- Thompson, D.E., Warschausky, S., Huggins, J.E., (2013). *Classifier-based latency estimation: a novel way to estimate and predict BCI accuracy*. J Neural Eng 10, 016006. doi:10.1088/1741-2560/10/1/016006
- Treder, M.S., Blankertz, B., (2010). *(C)overt attention and visual speller design in an ERP-based brain-computer interface*. Behav Brain Funct 6, 28. doi:10.1186/1744-9081-6-28

# Towards Cross-Subject Workload Prediction

Carina Walter<sup>1</sup>, Philipp Wolter<sup>1</sup>, Wolfgang Rosenstiel<sup>1</sup>, Martin Bogdan<sup>2,1</sup>  
and Martin Spüler<sup>1</sup>

<sup>1</sup> Department of Computer Engineering, Eberhard-Karls University Tübingen, Tübingen, Germany  
Carina.Walter@uni-tuebingen.de

<sup>2</sup> Department of Computer Engineering, University of Leipzig, Leipzig, Germany

## Abstract

Developing an electroencephalogram (EEG) based workload (WL) prediction could be utilized in learning environments to adapt instructional material to the students individual WL and thereby support the learning process successfully. In an adaptive learning environment it is not feasible to use data from the same subject and the same task for classifier training and testing. Therefore, the present study examined cross-subject WL prediction, in which one subject's individual WL is predicted based on data from other subjects. EEG-data was recorded from 10 subjects who had to solve math problems with increasing level of difficulty. By applying linear regression within- and cross-subject a precise WL prediction could be reached. For the within-subject WL prediction an average correlation coefficient ( $CC$ ) of  $CC = 0.88$  was achieved. However the cross-subject WL prediction leads to  $CC = 0.84$  on average. Since both prediction methods achieved good WL prediction results the cross-subject method is a feasible approach that can be used in an online adaptive system.

## 1 Introduction

In accordance with Cognitive Load Theory (CLT) [5] the type and amount of workload (WL) during learning is crucial for successful learning and should be held in the individual optimal WL range for each learner. Thus it seems advisable to provide technological support for learners, which adapts instructional materials and tasks to their level of expertise and WL capacity.

As it is not feasible to use data from the same subject and the same task for classifier training and testing in an adaptive learning environment, one can either use data from the same subject but different tasks (cross-task) or from the same task, but different subjects (cross-subject) to calibrate the classifier.

In a previous study [7] we used cross-task classification, where a SVM was trained on electroencephalogram (EEG)-data, recorded while participants had to solve working memory tasks. Subsequently the SVM was tested on EEG-data recorded while solving complex mathematical tasks. This approach led to classifier accuracies around chance level. Thus we introduced a cross-subject WL prediction in the present study based on linear regression, for a group of 10 subjects. As most of the EEG-based WL classifications are subject-specific, apart from a few exceptions [8], this work is a step towards the development of EEG-based WL classifiers as it would enable training a classifier once that could handle multiple subjects.

## 2 Materials and Methods

### 2.1 Participants and Task Design

A total of 10 subjects (17 - 32 years) voluntarily participated in the EEG experiment. The experiment had a within-subject design and comprised two tasks, a WL as well as a vigilance

task. In this article we will merely report the results of the WL task. The subjects had to solve 240 math problems with an increasing level of difficulty, by typing the solution in a given time. The presented problems varied in difficulty as measured by the information content (Q) according to Thomas [6], with a Q value ranging from 0.6 (easy) to 7.2 (difficult). As postulated from Kantowitz [2] increasing task difficulty ( $\hat{=}$  increasing Q-value) always increases WL, since by definition, an increase of task difficulty demands additional capacity. Further, an increase of WL is characterized in EEG-data by changes in the theta-, alpha- and beta-frequency bands [3]. To avoid the classifier of being based on perceptual-motor confounds the time windows used for analyzing EEG-data should not contain motor events. Therefore, each trial was divided into two phases: First, the calculation phase occurred (= phase for analyzing the EEG-data), where the problem to be solved was shown for 5 sec. Subsequently subjects had 3.5 sec to type in their result, followed by an inter-trial interval of 1.5 sec.

## 2.2 EEG Recording

A set of 29 active electrodes (actiCap, BrainProducts GmbH), attached to the scalp placed according to the extended International Electrode 10 - 20 Placement System, was used to record EEG-signals. Three additional electrodes were used to record an electrooculogram (EOG); two placed horizontally at the outer canthus of the left and right eye to measure horizontal eye movements and one placed in the middle of the forehead between the eyes to measure vertical eye movements. EOG- and EEG-signals were amplified by two 16-channel biosignal amplifier systems (g.USBamp, g.tec). The sampling rate was 512 Hz. EEG-data was high-pass filtered at 0.1 Hz and low-pass filtered at 100 Hz during the recording. Furthermore a notch-filter was applied between 48 - 52 Hz to filter power line noise.

## 2.3 Data Processing and Analysis

For further analysis we reduced the number of channels to 17 (FPz, AFz, F3, Fz, F4, FC3, FCz, FC4, C3, Cz, C4, CPz, P3, Pz, P4, Oz, POz), to lower the influence of possible artifacts, which are most prominent on the outer channels. As frequency bands were not consistent and varied between subjects, we used a wide frequency range of 4 - 30 Hz [3].

The power spectra were calculated for each 5 sec window (= calculation phase) by using autoregressive models based on Burgs maximum entropy method [1], using a model order of 32. In addition, the data was z-score normalized along the channels, meaning for each trial the mean of each frequency bin equals zero. The feature selection was conducted during the 10-fold or leave-one-subject-out cross-validation only on the training data. Each feature was correlated with the Information Content (Q) and the 125 features with the highest  $r^2$ -values were selected for the regression analysis. For WL prediction we used a linear ridge regression with the regularization parameter set to a fixed value of  $\lambda = 0.001$ .

As criterion for performance evaluation of the presented prediction method we used the correlation coefficient (CC) to observe the statistical relationship between the actual and the predicted Q-values, as well as the root-mean-squared error (RMSE) to examine the difference between the actual and the predicted Q-values.

Since single-trial prediction is not necessary in a learning environment, the regression output was additionally smoothed to improve prediction accuracy at the expense of increased delay of the system. We used a window-size of 6 trials, which still guaranteed a response time  $\leq 1$  min of the system, which is feasible for the detection of WL.

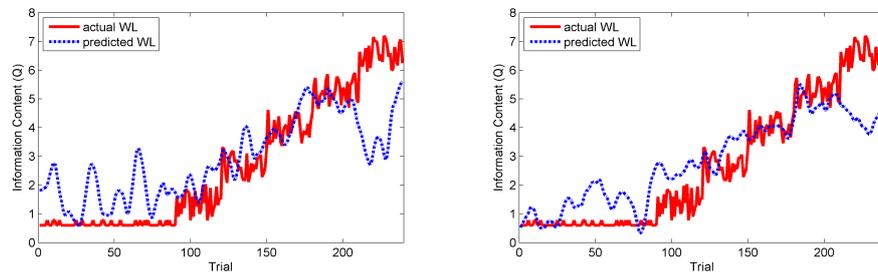


Figure 1: Distribution of actual and predicted WL for subject 1. The continuous red line represents the actual WL and the smoothed dashed blue line the predicted WL for each trial. Left: Within-subject - Using a 10-fold cross-validation; Right: Cross-subject - Regression trained on data of nine subjects and tested with data of the remaining subject 1.

### 3 Results

The performance for the within-subject WL prediction was calculated by using a 10-fold cross-validation on 240 trials (see Table 1). For the evaluation of the cross-subject WL prediction the regression was trained on data of nine subjects (= 240 · 9 trials) and tested with the complete data-set of the remaining subject (see Table 2). The within-subject WL prediction reached on average a correlation coefficient of  $CC = 0.66$  and a  $RMSE = 1.70$ . Using the smoothed regression output for within-subject WL prediction leads to more robust and even better results with an average over all subjects of  $CC = 0.88$  and  $RMSE = 1.02$  (see Table 1).

As stated in Table 2 the average results over all 10 subjects for cross-subject WL prediction slightly decreased. Applying linear regression to unsmoothed data leads on average to a  $CC = 0.58$  and a  $RMSE = 1.95$ . Using the smoothed data for cross-subject WL prediction improved the  $CC$  by 0.26 and the  $RMSE$  by 0.37.

The distributions of actual and predicted WL for within- and cross-subjects are exemplary shown for subject 1 in Fig. 1. In both cases the increasing WL was successfully predicted by using the linear regression method. It can be noticed that during the first 90 trials ( $Q \leq 0.9$ ) and the last 30 trials ( $Q \geq 6.0$ ) the actual and predicted WL deviate strongest from each other. This may be due to subjects being unchallenged or overburdened while solving these tasks.

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	mean
$CC$	0.43	0.69	0.68	0.58	0.84	0.58	0.60	0.72	0.65	0.78	0.66
$CC$ (smooth)	0.69	0.93	0.93	0.89	0.94	0.87	0.76	0.91	0.93	0.93	0.88
$RMSE$	2.20	1.61	1.71	2.09	1.11	1.95	1.63	1.60	1.82	1.30	1.70
$RMSE$ (smooth)	1.39	0.86	0.99	1.21	0.68	1.15	1.04	1.04	1.08	0.73	1.02

Table 1: Performance results of the within-subject WL prediction using 10-fold cross-validation.

### 4 Discussion

The results show that the amount of WL can be predicted using a cross-subject regression method. Compared to earlier studies, where cross-task classification [7] leads to classification

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	mean
<i>CC</i>	0.67	0.50	0.65	0.63	0.63	0.59	0.39	0.62	0.52	0.57	0.58
<i>CC</i> (smooth)	0.89	0.91	0.88	0.91	0.89	0.86	0.74	0.84	0.71	0.79	0.84
<i>RMSE</i>	1.68	2.56	1.69	1.73	1.72	1.77	2.65	1.87	2.04	1.81	1.95
<i>RMSE</i> (smooth)	1.13	2.00	1.25	1.39	1.60	1.44	2.11	1.58	1.57	1.68	1.58

Table 2: Performance results of the cross-subject WL prediction. Regression was trained on data of nine subjects and tested with data from the remaining subject.

accuracies merely around chance level, cross-subject prediction seems to be more robust and recommendable. As the math problems were presented in a fixed order, at advancing levels of difficulty, EEG-signals might change due to non-stationarity (e.g. caused by fatigue) over time. Actually non-stationarity within a session should not have a great influence on the EEG-data while using cross-subject methods, since non-stationarities between subjects are assumed to be larger than within a session. Since we used a very simple method for normalization and thereby alleviated non-stationarities, we expect to further improve the results by using more advanced methods for reduction of non-stationarities [4]. In an upcoming study, these results will be used in an online learning environment to predict the user's WL and adapt the presented exercises accordingly, to support students successfully in their learning process.

## Acknowledgments

The research was supported by ScienceCampus Tübingen and the LEAD Graduate School at the University of Tübingen, which is funded by the German Research Foundation.

## References

- [1] T. M. Cover and J. A. Thomas. *Elements of information theory*. Hoboken, NJ: Wiley-Interscience, 2006.
- [2] B. H. Kantowitz. *Human Factors Psychology*, chapter Mental Workload, pages 81–122. North-Holland, 1987.
- [3] M. Pesonen, H. Hämäläinen, and C. M. Krause. Brain oscillatory 4-30 Hz responses during a visual n-back memory task with varying memory load. *Brain Research*, 1138:171–177, 2007.
- [4] W. Samek, M. Kawanabe, and C. Vidaurre. Group-wise stationary subspace analysis - a novel method for studying non-stationarities. In *Proc. 5th Int. Brain-Computer Interface Conf.*, pages 16–20, 2011.
- [5] J. Sweller, J. J. G. Van Merriënboer, and F. Paas. Cognitive architecture and instructional design. *Educational Psychology Review*, 10(3):251–296, 1998.
- [6] H. B. G. Thomas. Communication theory and the constellation hypothesis of calculation. *Quarterly Journal of Experimental Psychology*, 15(3):173–191, 1963.
- [7] C. Walter, S. Schmidt, W. Rosenstiel, P. Gerjets, and M. Bogdan. Using cross-task classification for classifying workload levels in complex learning tasks. In *Proceedings of the 5th IEEE Humaine Association Conference on Affective Computing and Intelligent Interaction*, pages 876–881, Geneva, Switzerland, 2013.
- [8] Z. Wang, R. M. Hope, Z. Wang, Q. Ji, and W. D. Gray. Cross-subject workload classification with a hierarchical bayes model. *NeuroImage*, 59(1):64–69, 2012.

# Implicit Learning of SSVEP-based Brain Computer Interface

Jonathan Giron, Miri Segal and Doron Friedman

The Interdisciplinary Center, Herzliya, Israel  
doronf@post.idc.ac.il

## Abstract

Our study shows that subjects can implicitly learn to use a steady-state visually-evoked potential (SSVEP) based brain computer interface (BCI). The SSVEP stimuli were presented in an immersive star field virtual environment. Within the star field the SSVEP stimuli appeared in a pseudo random order. Participants' attention to the stimuli resulted in stars moving within the immersive space. Participants were asked to view four short clips of the scene and try to explain why the stars were moving, without being told that they are controlling a BCI. Two groups were tested: one that interacted implicitly with the interface, and a control group in which the interaction was a sham (i.e., the interface was activated independently of the participants' attention). Following the exposure to the immersive scene the participants' BCI accuracy was tested, and the experiment group showed higher accuracy results. This finding may indicate that implicit SSVEP BCI interactions are sufficient in inducing a learning effect for the skill of operating a BCI.

## 1 Introduction

We demonstrate the possibility that BCI can be learned by implicit feedback reinforcement; i.e., without instructions, such as in many other tasks, including motor tasks. Moreover, the subjects in the present study were not even instructed that there was an opportunity to learn, but were led to believe that they are taking part in a passive experiment. The study is based on a generic system we have developed that allows turning any 3D object in a virtual environment into an SSVEP target with reliable flicker rates.

## 2 Method

### 2.1 Participants

Twenty one subjects took part in the experiment (10 males, 11 females) with a mean age of 24 (range 19-40). All subjects had normal or corrected-to-normal vision. They showed no signs of neurological or psychiatric disorders and all gave written, informed consent. Each subject was paid the equivalent of 12 Euro for participating in the experiment, and the study was approved by the institutional ethics committee.

### 2.2 Apparatus

EEG signals were recorded using 8 g.LADYbird sintered Ag/Cl crown active ring electrodes located on the subjects occipital lobe at pO7, PO3, POz, PO4, PO8, O1, Oz and O2 locations according to the international 10-20 system (Figure 1). A reference electrode was positioned on the subjects right ear lobe and a ground electrode was placed at Fpz location according to the international 10-20 system. EEG signals were recorded at 256 Hz sample rate and amplification

and analog filtering (5-100 Hz) and notch filtering at 50Hz were performed using a g.USBamp amplifier (Guger Technologies, Schiedlberg, Austria).

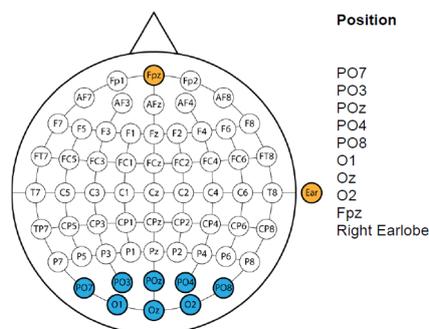


Figure 1: EEG electrode montage

We used a system developed in our lab for turning 3D objects in the Unity 3D game engine (Unity Technologies, USA) into SSVEP targets with reliable flicker rates. The flickering stimuli (at 8.57Hz, 12Hz, 15Hz, and 20Hz flickering rates) were projected on a 182cm (height) by 256cm (width) screen in a dark room that created an immerse virtual environment. Participants were asked to sit on an office arm chair positioned 180cm from the screen. The application was displayed using a 3D ViewSonic 120 screen refresh rate projector at 1280\*768 resolution using a high-end graphics card. SSVEP classification was calculated using the algorithm by [1] analyzed using Matlab (MathWorks, US) and implemented by g.tec (Guger Technologies, Schiedlberg, Austria). A linear discriminant analysis (LDA) classification method with a zero class and a 1% confidence interval was applied and one harmonic frequency for each frequency was taken into account for classification.

### 2.3 Procedure and Data analysis

The experiment included three stages. In the first training stage the system computed a classifier model. The training sessions included 20 stimuli, 5 times of each frequency in pseudo-random order and location on the screen. The first stage might have been repeated until the first successful session, indicated by model fit of 90% accuracy or more to the training data. The second stage was the task stage, including 4 sessions of immerisve experience. The experience included a star field moving towards the subject, with a pseudo random collection of larger stars, which were the SSVEP targets. The instruction to all subjects was to attend to the stimuli and speculate about the reason that the stars were moving. Eleven subjects experienced the experimental group, in which they were only told about their BCI control after 3 of the 4 task trials. After 1 second of continous SSVEP classification of an on screen star that star began moving towards the subject, possibly ‘exploding’ on screen if attention was kept to it long enough. Ten subjects participated in the sham control group, in which the motion of the stars in the first 3 trials was random. The third stage, a validation trial intended to determine BCI accuracy, was identical to the training trial.

Classification data was collected during each training session. For each frequency 5 trials each 7 seconds long were entered into the algorithm for classifier training, while comparing these 7 seconds with the 6 second trial interval rest periods for zero class classification. During the online classification of the signals at the second stage of the experiment, a 2-second running

window buffer was used for classification. The algorithm produced a classification result 5 times per second (i.e., every 200ms). In order to decrease false positive classification errors a filter was applied to the stream of classification results; a total of 6 consecutive results were required of the same class in order to activate a game object (stars) command. This corresponds to one second of consistent classification. Since a running buffer of 2 seconds was also used, this entails that the minimum response time of the stars (game objects) was 3 seconds.

### 3 Results

The hypothesis that implicit reinforcement induces a learning effect was tested using the two groups accuracy results obtained at the validation stage using T-test analysis. First, optimal time point in terms of classification accuracy was obtained; this time point for all subjects was exactly 6 seconds post stimuli onset, and thus the accuracy per session in the following analysis was based on the accuracy between seconds 5 and 6. The experiment group achieved an average of 82.5% accuracy while the sham group achieved an average of 55.55%. The difference between the groups was found to be significant ( $t(19) = 2.23, p < 0.05$ ; Figure 2).

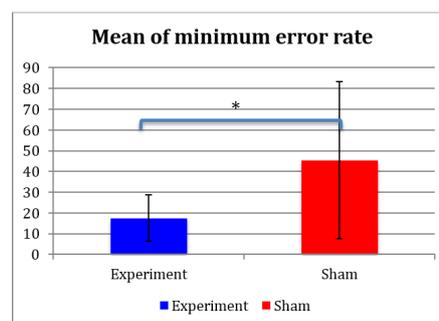


Figure 2: Minimum error rates between groups at optimal accuracy time stamp (5-6 seconds post stimuli onset).

In order to further confirm this finding an ANOVA with repeated measures analysis was used (Figure 3). The within variable considered was time, sampled once per second, the between variable was group, and the dependent variable was accuracy rate. A significant main effect was found for time ( $F(9,171) = 36.43, p < .001, \text{partial } \eta^2 = 0.66$ ). A significant effect for group was not found ( $F(1, 19) = 1.05, \text{n.s.}$ ), but taking into account time a significant interaction was found ( $F(9,171) = 4.05, p < .05, \text{partial } \eta^2 = .17$ ): after five seconds of stimulus presentation, the accuracy between the groups changed, and the sham group displayed lower accuracy rate past that time point, while the experiment group accuracy rates continued to rise. Note that in the first five seconds post stimuli the classification is approximately random. We do not know why the classification in the sham group was slightly higher in this early duration. The optimum performance in this study was around 6 seconds after the trigger, in contrary to 3-4 seconds as reported in most SSVEP studies; we suspect that the result is the combination of a 2 second buffer and a 1 second filter that we have applied. This, together with the minimal training, can also explain the relatively low accuracy rates. When comparing false positive results, a significant effect for time was also found ( $F(9,171) = 12.21, p < .05$ ), but no significant effect was found between groups ( $F(1,19) = 0.91, p > .05, \text{n.s.}$ ), and no interaction was found

( $F(2,171) = 1.88, p > .05, n.s.$ ). These results indicate significantly higher accuracy rate for the experiment group, but only in the duration when classification was not random.

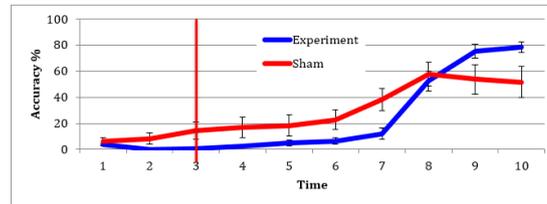


Figure 3: Linear comparison between the groups error rates; a significant interaction was found between time and group BCI accuracy ( $F(9,171) = 4.05, p < .05, \text{partial } \eta = .17$ ).

## 4 Discussion

The present findings support the hypothesis that the developed interface is easy to use and that its activation can be accomplished without prior knowledge. Moreover, it was found that implicit SSVEP BCI interactions are sufficient in inducing a learning effect for the skill of operating a BCI. Middendorf et al. [3] have shown that subjects can voluntarily learn to regulate their SSVEP amplitude by explicit training. Guger et al. [2] demonstrate that very short SSVEP training is often enough to reach high performance. In this study we show that an implicit reinforcement method can result in improved accuracy rates. This suggests the possibility that BCI control may be an acquired ability similar to motor abilities; this has both theoretic implications for understanding how BCI skills are acquired, as well as practical implications, e.g., for locked in patients who cannot learn to control a BCI that is based on explicit control strategies.

## 5 Acknowledgements

This project was supported by EU projects BEAMING (248620) and VERE (257695). The authors wish to thank Gilan Jackont, Keren-OR Berkers and Yuval Kalgony for programming the Unity environment, and Christoph Hintermuller of gTec GMBH for support with the SSVEP module. The authors also wish to thank Beatrice Halser for her suggestions on the experimental design.

## References

- [1] Henrik Gollee, Ivan Volosyak, Angus J McLachlan, Kenneth J Hunt, and Axel Gräser. An SSVEP-based brain-computer interface for the control of functional electrical stimulation. *IEEE transactions on bio-medical engineering*, 57(8):1847–55, August 2010.
- [2] Christoph Guger, Brendan Z Allison, Bernhard Großwindhager, Robert Prückl, Christoph Hintermüller, Christoph Kapeller, Markus Bruckner, Gunther Krausz, and Günter Edlinger. How many people could use an ssvp bci? *Frontiers in neuroscience*, 6, 2012.
- [3] Matthew Middendorf, Grant McMillan, Gloria Calhoun, Keith S Jones, et al. Brain-computer interfaces based on the steady-state visual-evoked response. *IEEE Transactions on Rehabilitation Engineering*, 8(2):211–214, 2000.

# Spatial Tactile Brain-Computer Interface Paradigm Applying Vibration Stimuli to Large Areas of User's Back

Takumi Kodama<sup>1,\*</sup>, Shoji Makino<sup>1,\*</sup> and Tomasz M. Rutkowski<sup>1,2,\*</sup>,<sup>†</sup>

<sup>1</sup> Life Science Center of TARA, University of Tsukuba, Tsukuba, Japan

<sup>2</sup> RIKEN Brain Science Institute, Wako-shi, Japan

tomek@bci-lab.info

<http://bci-lab.info>

## Abstract

We aim at an augmentation of communication abilities of *locked-in syndrome* (LIS) patients by creating a *brain-computer interface* (BCI) which can control a computer or other device by using only brain activity. As a method, we use a stimulus-driven BCI based on vibration stimuli delivered via a gaming pad to the user's back. We identify P300 responses from brain activity data in response to the vibration stimuli. The user's intentions are classified according to the P300 responses recorded in the EEG. From the results of the psychophysical and online BCI experiments, we are able to classify successfully the P300 responses, which proves the effectiveness of the proposed method.

## 1 Introduction

Recently, vibrotactile-based somatosensory modality BCIs have gained in popularity [1, 4, 2]. We propose an alternative tactile BCI which uses P300 brain responses to a somatosensory stimulation delivered to larger areas of the user's back, defined as a back-tactile BCI (btBCI). We conduct experiments by applying vibration stimuli to the user's back, which allows us to stimulate places at larger distances on the body in order to test our hypothesis of the full body-based tactile BCI paradigm validity. The stimulated back areas are both shoulders, the waist and the hips. In order to do so, we utilize a haptic gaming pad "ZEUS VYBE" by Disney & Comfort Research. An audio signal pad's input allows for the delivery of a sound pattern activating spatial tactile patterns of vibrotactile transducers embedded within the device. In the experiments reported in this paper, the users lay down on the gaming pad and interacted with tactile stimulus patterns delivered in an oddball-style paradigm to their backs, as shown in Figure 1. The reason for using the horizontal position of the gaming pad, developed for a seated setting, is that bedridden users, with intact afferent neural fibers [5], could easily utilize it and it could also serve as a muscle massage preventing the formation of bedsores.

## 2 Methods

In the research project reported in this paper the psychophysical and online EEG experiments were carried out with able-bodied, BCI-naive users. Seven healthy users participated in the study (three males and four females) with a mean age of 25 years (standard deviation of 7.8 years). The users were paid for their participation. All the experiments were performed at the Life Science Center of

---

\*The author was supported in part by the Strategic Information and Communications R&D Promotion Program no. 121803027 of The Ministry of Internal Affairs and Communications in Japan.

<sup>†</sup>The corresponding author.

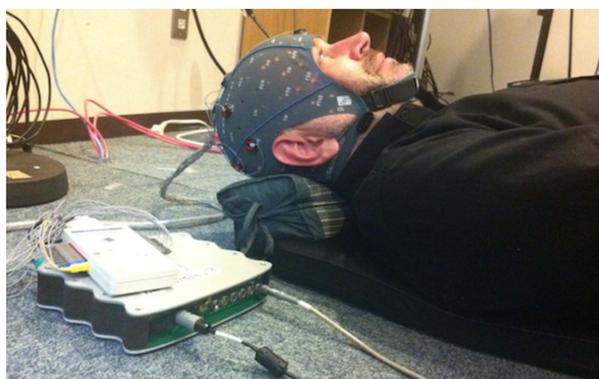


Figure 1: A btBCI user lying on the gaming pad as in the experiments reported in this paper. The g.USBamp by g.tec with g.LADYbird electrodes is also depicted. The photograph is included with the user’s permission.

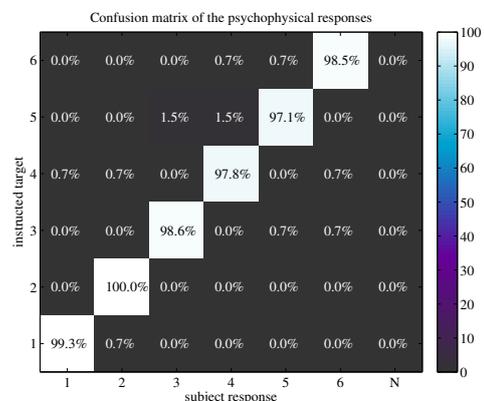


Figure 2: Psychophysical experiment grand mean averaged accuracies with scores above 97% for each of the six commands. “N” stands for no-response cases.

TARA, University of Tsukuba, Japan. The procedures for the psychophysical experiments and EEG recordings for the BCI paradigm were approved by the Ethical Committee of the Faculty of Engineering, Information and Systems at the University of Tsukuba, Tsukuba, Japan. Each participant signed to give informed consent to taking part in the experiments. The psychophysical experiments were conducted to investigate the recognition accuracy and response times to the stimuli delivered from the gaming

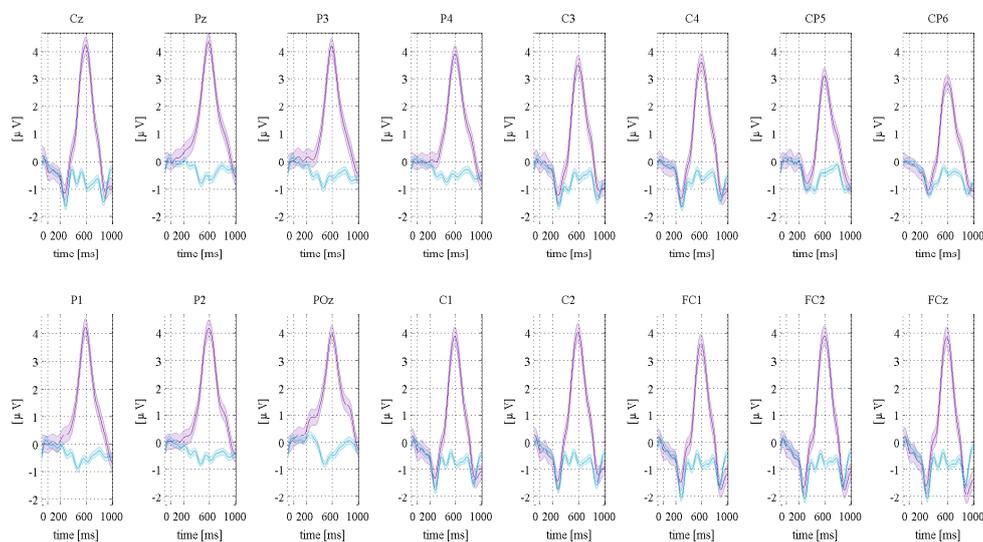


Figure 3: Grand mean averaged ERP for target (purple line) and non-target (blue line) stimuli. The very clear P300 responses are easy to notice for each EEG electrode in the latencies of 200 ~ 1000 ms. Eye blink contaminated ERPs were rejected in the process of creation of this figure, with a threshold of 80  $\mu V$ .

pad. The behavioral responses were collected as keyboard button presses after instructed targets. In the psychophysical experiment, each single trial was comprised of a randomly presented single target and five non-target vibrotactile stimuli (120-targets and 600-non-targets in a single session). The stimulus duration was set to 300 ms and the inter-stimulus-interval (ISI) to 700 ms. In the btBCI online experiments, the EEG signals were captured with a bio-signal amplifier system g.USBamp by g.tec Medical Instruments, Austria. Active EEG electrodes were attached to the sixteen locations *Cz, Pz, P3, P4, C3, C4, CP5, CP6, P1, P2, POz, C1, C2, FC1, FC2 and FCz*, as in 10/10 international system. A reference electrode was attached to the left mastoid, and a ground electrode to the forehead at the *FPz* position. The EEG signals were captured and classified by BCI2000 software [6] using a stepwise linear discriminant analysis (SWLDA) classifier [3] applied to sixteen times downsampled 0 ~ 800 ms ERP latencies. The EEG recording sampling rate was set at 512 Hz, and the high and low pass filters were set at 0.1 Hz and 60 Hz, respectively. The notch filter to remove power line interference was set for a rejection band of 48 ~ 52 Hz. In each trial, the stimulus duration was set to 250 ms and the ISI to random values in a range of 350 ~ 370 ms in order to break rhythmic patterns of presentation. Each online experiment comprised of 10 trials used for epochs averaging. Users conducted three runs of which the first one was used for SWLDA classifiers truing and the remaining for testing. The vibrotactile spatial pattern stimuli (stimulated body locations used as cues) in the two experimental settings above were generated using the same *MAX 6* program, and the trigger onsets were generated by *BCI2000* EEG acquisition and ERP classification software [6]. During the training run the cues were given in form of vibrations of the target location before each sequence. In testing phase the spelled digits (representing the body locations) were given visually on a computer display.

### 3 Results

The psychophysical experiment results are summarized in the form of a confusion matrix depicted in Figure 2, where the behavioral response accuracies to instructed targets and marginal errors are depicted together with no-response errors, which were not observed with the users participating in our experiments. The grand mean behavioral accuracies were above 97%, which proved the easiness of the back vibrotactile stimuli discrimination. The behavioral response times did not differ significantly among classes as tested with the Wilcoxon rank sum test for medians, which further supported the choice of the experiment set-up with vibrotactile stimuli to the back. The EEG experiment results are summarized in Figures 3 and 4 in the form of grand mean averaged ERPs and classification accuracies. The grand mean averaged ERPs resulted in very clear P300 responses in latency ranges of 200 ~ 1000 ms. The SWLDA classification results in online btBCI experiments of six digit spelling are shown in Figure 4, depicting each user's averaged scores in a range of 16.7% ~ 62.45% and the best grand mean results in the range of 57.26% ~ 85.71%, both as a function of various ERP averaging scenarios. The chance level was 16.7%. The best mean scores show very promising patterns for possible improvements based on longer user training. Mean information-transfer-rates ranged from 0.6 bit/min to 3.3 bit/min for 10-trials averaging based SWLDA classification to 0.5 bit/min to 10.9 bit/min for single trial offline analysis cases.

### 4 Conclusions

This paper reports results obtained with a novel six-command-based btBCI prototype developed and evaluated in experiments with seven BCI-naive users. The experiment results obtained in this study confirm the general validity of the btBCI for six command-based applications and the possibility to further improve the results, as illuminated by the best mean accuracies achieved by the users. The EEG

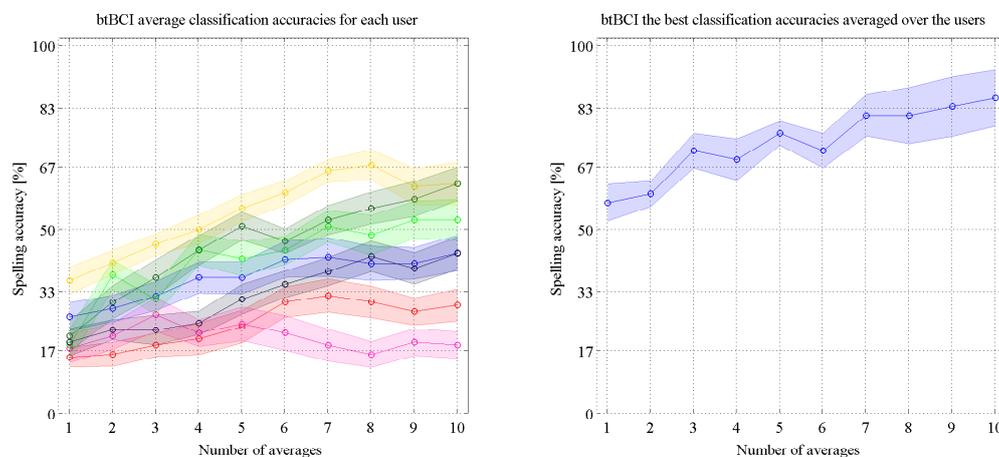


Figure 4: The averaged btBCI accuracies as obtained with a SWLDA classifier with different numbers of averaged ERPs. The left panel presents mean accuracies for each user separately, together with standard error bars calculated over different runs. The right panel presents the averaged maximum scores (the best runs) of all the users together in order to illustrate the potential strength of the paradigm in comparison to average scores shown in the left graph.

experiment with the prototype confirms that tactile stimuli to large areas of the back can be used to spell six-digit (command) sequences. The results presented offer a step forward in the development of somatosensory modality neurotechnology applications. Due to the still not very satisfactory interfacing rate achieved in the case of the online btBCI, the current prototype obviously requires improvements and modifications. These requirements will determine the major lines of study for future research. However, even in its current form, the proposed btBCI can be regarded as a possible alternative solution for locked-in syndrome patients, who cannot use vision or auditory based interfaces due to sensory or other disabilities.

## References

- [1] A-M. Brouwer and J.B.F. Van Erp. A tactile P300 brain-computer interface. *Frontiers in Neuroscience*, 4(19), 2010.
- [2] J.E. Huggins, C. Guger, B. Allison, C.W. Anderson, A. Batista, A-M. Brouwer, C. Brunner, R. Chavarriaga, M. Fried-Oken, A. Gunduz, D. Gupta, A. Kübler, R. Leeb, F. Lotte, L.E. Miller, G. Müller-Putz, T. Rutkowski, M. Tangermann, and D.E. Thompson. Workshops of the fifth international brain-computer interface meeting: Defining the future. *Brain-Computer Interfaces*, 1(1):27–49, 2014.
- [3] D.J. Krusienski, E.W. Sellers, F. Cabestaing, S. Bayouduh, D.J. McFarland, T.M. Vaughan, and J.R. Wolpaw. A comparison of classification techniques for the P300 speller. *Journal of Neural Engineering*, 3(4):299, 2006.
- [4] H. Mori, Y. Matsumoto, Z.R. Struzik, K. Mori, S. Makino, D. Mandic, and T.M. Rutkowski. Multi-command tactile and auditory brain computer interface based on head position stimulation. In *Proceedings of the Fifth International Brain-Computer Interface Meeting 2013*, page Article ID: 095, Asilomar Conference Center, Pacific Grove, CA USA, June 3-7, 2013. Graz University of Technology Publishing House, Austria.
- [5] J R Patterson and M Grabois. Locked-in syndrome: a review of 139 cases. *Stroke*, 17(4):758–64, 1986.
- [6] G. Schalk and J. Mellinger. *A Practical Guide to Brain-Computer Interfacing with BCI2000*. Springer-Verlag London Limited, 2010.

# Reduced Activation at the Cortical Level Following Neurofeedback Treatment is Associated with Reduction in Central Neuropathic Pain Intensity

Muhammad Abul Hasan<sup>1,2</sup>, Aleksandra Vuckovic<sup>1</sup>, Matthew Fraser<sup>3</sup>,  
Bernard Conway<sup>4</sup>, Bahman Nasserolelami<sup>4,5</sup>, David B Allan<sup>3</sup>

<sup>1</sup>University of Glasgow, Glasgow, UK. <sup>2</sup>NED University Karachi, Pakistan. <sup>3</sup>Sothorn General Hospital, UK. <sup>4</sup>University of Strathclyde, Glasgow, UK. <sup>5</sup>Northeastern University, Boston, USA.  
m.hasan.1@research.gla.ac.uk, Aleksandra.Vuckovic@glasgow.ac.uk

## Abstract

The aim of the study was to investigate the effect of neurofeedback (NF) training on Central neuropathic pain in patients with chronic paraplegia. Training consisted of 20-40 NF sessions in which patients were asked to voluntarily reduce the theta (4-8 Hz) and beta (20-30 Hz) band power and to increase the alpha band (9-12 Hz) power over the sensory-motor cortex. A whole head EEG was recorded before the first and after the last training day to assess long-term effect of NF. Four out of five patients reported clinically significant (>30%) reduction in pain. Multichannel LORETA analysis indicated that reduced pain intensity might be associated with reduced cortical and sub-cortical activity in the areas known to be involved in processing of chronic pain.

## 1 Introduction

Central Neuropathic Pain (CNP) is caused by a lesion or disease to the somatosensory system having a prevalence of 40 % in spinal cord injured (SCI) population (Werhagen et al. 2004). CNP is perceived as coming from the body, while it is actually generated in the brain and is mainly associated with the overactivation of the cortex (Samthein et al. 2006). Noninvasive interventions such as repetitive Transcranial Magnetic Stimulation (rTMS) and transcranial Direct Current Stimulation (tDCS) reduce cortical and/or sub-cortical activity thus reducing pain (Moseley and Flor 2012).

Similar to rTMS and tDCS, neurofeedback (NF) is also a neuromodulation technique which is based on voluntary modulation of brain activity. Although NF has been successfully used for different types of chronic pain such as fibromyalgia and migraine (Kayiran et al. 2010), its efficiency on CNP has not been confirmed. In a recent study we showed that paraplegic patients with CNP have increased power in theta band, lower frequency of dominant alpha peak and increased Event Related Desynchronisation (ERD) in the theta, alpha and beta band (Vuckovic et al. 2014). Based on that study, we created a NF protocol for treatment of CNP, presented in this paper.

## 2 Methods

### 2.1 Patients

Seven paraplegic patients with CNP larger than 5 (Visual Analogue Scale, range 0 to 10) participated in the study. All patients had pain in their limbs. Ethical permission was obtained from the NHS Ethical Committee for the Greater Glasgow and Clyde. Experiments were performed at Queen Elizabeth National Spinal Injuries Unit, Glasgow, UK.

### 2.2 Experimental Procedures

The experiment consisted of an initial assessment, NF training, and final assessment. On the initial and final assessment, patients' EEG was recorded (Synamp<sup>2</sup>, Neuroscan, USA) with 61 channels in a relaxed state and during motor imagery of the upper and the lower limbs. In a cue-based experimental paradigm patients were asked to imagine moving their right hand, left hand and feet in a semi-random order. There were 60 repetition of each type of imagined movements (MI). The electrodes were placed according to the standard 10-10 locations using an ear-linked reference and AFz as ground. Sampling frequency was 250 Hz and impedance was kept below 5k $\Omega$ . For MI data, noise was removed using Independent Component Analysis (ICA). ERD and Event-related synchronization (ERS) was calculated based on wavelets (Makeig 1993). For spontaneous EEG recording, after manually removing sections with extensive noise (>100  $\mu$ V), at least 3 min of recording was left for the analysis. The sLORETA (Pascual-Marqui 2002) computed cortical map/image (estimated current density) for 6239 voxels at 5 mm spatial resolution.

For NF training custom made software was designed in Simulink/Matlab and Graphical User Interface was developed in LabView. EEG was recorded using usbamp (Guger technologies, Austria). Prior to each training session patients' EEG (256 sample/sec) was recorded in relaxed, eyes opened state, from C3, C4, P4 and Cz to determine the baseline for training for that day, though a single site was used for training at a time. The relative power, with respect to power in 2-30Hz band was calculated for the 'inhibit' theta (4-8 Hz), 'reward' alpha (9-12Hz) and 'inhibit' beta (20-30Hz) bands, being determined based on (Vuckovic et al. 2014). Patients were trained to reduce a relative power in the inhibit bands and to increase it in the reward band. Alpha  $\geq$ 9Hz was selected in order to shift the dominant frequency towards higher frequencies (Vuckovic et al. 2014). The threshold was set 10% above the reward and 10% below the inhibit bands. The power of both inhibits and a reward band was shown on a computer screen in the form of vertical bars that changed a size and color from red to green. Power was calculated in real time using moving average filter of 5<sup>th</sup> order. Training was based on the operant conditioning and patients were instructed to 'do whatever necessary' to keep the bars green.

For statistical analysis over a full spectrum, the EEG was divided into 4s epochs and power was calculated for each epoch. Following this, a parametric unpaired ttest was applied over epochs to compare power between each two conditions (Pre NF versus during NF, and Pre NF versus Post NF).

## 3 Results

Five out of seven patients completed the study. Two patients dropped out after 3 NF sessions due to family and transport problems. From the remaining five patients, four patients received 40 sessions and the fifth patient received 20 sessions. Patients received therapy 1-3 sessions per week.

### 3.1 Effect of NF Training on Pain Intensity and Power

All five patients achieved a statistically significant ( $p \leq 0.05$ ) reduction of pain (Table 1) and in four patients this reduction was clinically significant ( $>30\%$ ). Patients reported immediate reduction of pain during NF; reduction of pain gradually increased over training sessions and lasted for several weeks after terminating the therapy. Initially we tested several locations over the primary motor cortex which included C3, Cz and C4. In patients with incomplete paraplegia (preserved sensation) we occasionally observed the uncontrolled movements of legs (spasm) which was higher when practicing from C3 than from C4, so we chose C4 as a preferred training site to reduce discomfort. Interestingly, NF training from Cz, located over the motor cortex corresponding to ‘painful’ paralyzed legs resulted in smallest reduction of pain. Patients received 30 min training in total per session.

Figure 1 shows power in one patient before, during and after NF for a single training day. The patient increases the alpha and theta power, and decreases the beta power (black thick line on x-axis: Pre NF versus NF). This effect can also be noticed several minutes after NF training (grey thick line on x-axis: Post NF versus NF). We present PSD of an individual patient rather than the average value across the whole group because the exact frequency bands in which EEG was modulated varied among patients.

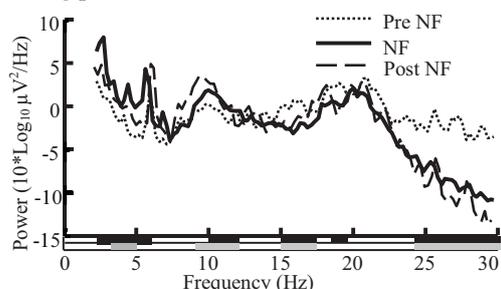


Figure 1: Power of fifth patient before NF training session (dotted line), during NF (solid line) and after NF training session (dashed line).

No	Injury level	Pain (before/after)
1	T5 complete	6/5, p=0.05
2	T6 complete	7/5, p=0.0004
3	T6 incomplete	6/2, p=0.0001
4	T6 incomplete	9/5, p=0.0002
5	T6 incomplete	9/5, p=0.006

Table 1: Patients characteristics and change in pain intensity following long-term NF training. Wilcoxon paired test was applied to find statistically significant change in pain intensity over all training days (Pre NF versus NF).

### 3.2 Long-Term Effect of NF

The reduction of brain activity was noticed in the theta, alpha and beta bands in cortical areas including sensory-motor and limbic cortices. Figure 2 shows the difference in sLORETA estimate of current density in relaxed EO state before and after NF therapy averaged over all 5 patients. A widespread reduction of the activity could be noticed at the Anterior Cingulate Cortex, (BA 24) in the beta band (12-15 Hz). Figure 3 shows ERS/ ERD at Cz site during MI of legs. Reduced activity after NF can be noticed mainly in the theta band and in beta band.

## 4 Conclusion and Discussion

The paper presents the effect of NF training on reduction on CNP and on related neurological measures. Although patients learned to modulate brain activity and reported reduction in pain that lasted for several weeks, a regular NF therapy on large number of patients would be needed to further confirm the effect of NF therapy. The reduced cortical activity in all frequency bands mainly in the theta and beta bands shows that reduction in pain is associated with reduced power in relaxed state

and reduced ERD. These results are in accordance to previous studies on EEG signatures of CNP (Sarnthein et al. 2006; Vuckovic et al. 2014).

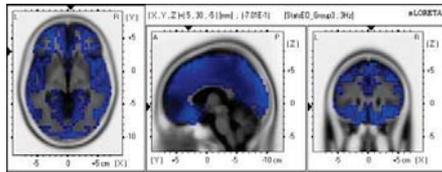


Figure 2: LORETA images showing reduction of activity (after-before NF) in 12-15 Hz band averaged over 5 patients, in ACC (BA 24).

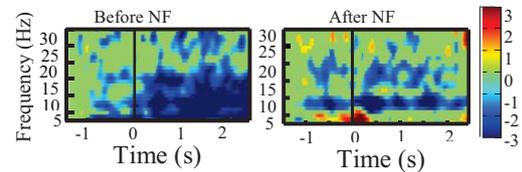


Figure 3: Patient 1, ERS/ERD at Cz before (left) and after treatment (right). Significance level set  $p=0.05$  with false discovery rate.

## 5 Acknowledgements

This work has been partially supported by the MRC grant G0902257/1, the GRPE, NED University of Pakistan, and GU 68. We thank Dr Purcell and Dr Mclean and to all participants.

## 6 References

- Kayiran Sadi et al. 2010. "Neurofeedback Intervention in Fibromyalgia Syndrome; a Randomized, Controlled, Rater Blind Clinical Trial." *Applied psychophysiology and biofeedback*. 35(4):293–302.
- Makeig, S. 1993. "Auditory Event-Related Dynamics of the EEG Spectrum and Effects of Exposure to Tones." *Electroencephalography and clinical neurophysiology* 86(4):283–93.
- Moseley, G. Lorimer, and Herta Flor. 2012. "Targeting Cortical Representations in the Treatment of Chronic Pain: A Review." *Neurorehabilitation and neural repair* 26(6):646–52.
- Pascual-Marqui, R. D. 2002. "Standardized Low-Resolution Brain Electromagnetic Tomography (sloreta): Technical Details." *Methods Find Exp Clin Pharmacol*. 24 (suppl.):5–12.
- Sarnthein et al. 2006. "Increased EEG Power and Slowed Dominant Frequency in Patients with Neurogenic Pain." *Brain : a journal of neurology* 129(Pt 1):55–64.
- Vuckovic et al. 2014. "Dynamic Oscillatory Signatures of Central Neuropathic Pain in Spinal Cord Injury." *Journal of Pain*.
- Werhagen et al. 2004. "Neuropathic Pain after Traumatic Spinal Cord Injury--Relations to Gender, Spinal Level, Completeness, and Age at the Time of Injury." *Spinal cord* 42(12):665–73.

# A Generic Machine-Learning Tool for Online Whole Brain Classification from fMRI

Ori Cohen<sup>1,2</sup>, Michal Ramot<sup>3</sup>, Rafael Malach<sup>3</sup>, Moshe Koppel<sup>2</sup> and Doron Friedman<sup>1</sup>

<sup>1</sup> The Interdisciplinary Center, Herzliya, Israel  
orioric@gmail.com, doronf@idc.ac.il

<sup>2</sup> Bar Ilan University, Ramat Gan, Israel  
koppel@cs.biu.ac.il

<sup>3</sup> Weizmann Institute of Science, Rehovot, Israel

## Abstract

*Objective.* We have developed an efficient generic machine learning (ML) tool for real-time fMRI whole brain classification, which can be used to explore novel brain-computer interface (BCI) or advanced neurofeedback (NF) strategies. *Approach.* We use information gain for isolating the most relevant voxels in the brain and a support vector machine classifier. *Main results.* We have used our tool in three types of experiments: motor movement, motor imagery and visual categories. *Significance.* We show high accuracy results in real-time, using an optimal number of voxels, with a shorter delay compared to the previous method based on regions of interest (ROI). Finally, our tool is integrated with a virtual environment and can be used to control a virtual avatar or a robot.

## 1 Introduction

Real time fMRI is a promising risk-free non-invasive method for several reasons. Due to superior spatial resolution fMRI may be used to classify a much wider set of mental patterns than EEG. Thus, fMRI-based BCI may facilitate exploring new BCI paradigms. fMRI BCI paradigms can be used to localize underlying brain patterns and transfer the paradigms back to other, more accessible signals, such as EEG and near-infrared spectrography (NIRS) [1, 2]. fMRI-based BCI can also be used for training patients in BCI (e.g., before surgery), for rehabilitation sessions, or for pattern-based neurofeedback (NF) – in all these cases very specific brain areas may be targeted.

We have developed a generic efficient tool for online whole-brain classification from fMRI data. In contrast to other research [3], this tool is designed for real-time processing, normalization and multi-classification which can be extended to other algorithms within the Weka ML API [4]. This tool is also integrated with a virtual environment feedback to allow for engaging subjects in a wide range of scenarios and tasks, and can also be integrated with external devices such as a humanoid robot. In this paper, we show how our approach achieves high accuracy in three different mental tasks: motor motion, motor imagery, and visual categories.

## 2 The System

Imaging was performed on a 3T Trio Magnetom Siemens scanner as described in [5, 6], with a repetition time (TR) of 2000ms. Visual feedback is provided by a mirror, placed 11cm from the eyes of the subject and 97.5cm from a screen, which results in a total distance of 108.5cm from the screen to the eyes of the subject.

Dicom files<sup>1</sup> from the scanner are preprocessed by Turbo Brain Voyager (TBV, Brain Innovation, Netherlands). Since fMRI data tends to have non linear non-homogeneous drifts, we introduce a normalization process that has been verified to perform well online; given a raw value at voxel  $i$  and time  $t$ ,  $r_{i,t}$ , and a sliding window of length  $w$  we derive a new value for each raw value:

$$r'_{i,t} = r_{i,t} - \text{median}(r_{i,t-w+1}, \dots, r_{i,t}) \quad (1)$$

We have empirically established that a  $w$  of 40 TRs (80 seconds) is optimal in our case.

Our tool is integrated with the Unity game engine (Unity Technologies, California); in other studies this allowed subjects to control an avatar in a virtual environment or a humanoid robot [5]. The system allows easily configuring classification and interaction parameters during an experiment and playing back experimental sessions.

Training and applying classifiers in real-time requires that learning be executed faster than is generally done in the application of ML to fMRI. Our system is optimized for memory usage, processing speed, and classification speed. To achieve faster processing, we focus on several areas: (1) feature reduction, (2) feature selection, (3) redundant data reduction, (4) minimization of computational cycles and RAM consumption by using sparse data handling, (5) using RAM without having to access the disk drive, and (6) transferring data between processes using an inter-process communication method.

Our method is divided into two parts. In the first part, the subject is given instructions based on the experimental protocol and several runs are recorded as input to a machine-learning tool. In the second part, the subjects perform a task, our system classifies their intentions in real time, and the classification result is transmitted to an external device. In this study the task was similar to training, in order to evaluate BCI accuracy.

The 3D matrix of the entire brain area is composed of 204,800 voxels, which hold the raw blood-oxygen-level dependent (BOLD) derived values. Our TBV plugin transmits the raw data using an inter-process communication method to our application. The 3D matrix is flattened into a 1-dimensional vector, and the set of voxels is further reduced by setting an activity threshold, which removes voxels that belong to the empty space around the head that and are not part of the brain. We also remove any voxels that belong to the subject's eyes, reducing the number of voxels to approximately 26-30,000 voxels per TR.

For purposes of learning, we select only those voxels with highest information gain (IG) [7]. Labeled training examples, each represented as a vector recording the values of the selected voxels, are passed on to our learning algorithm: Weka's [4] implementation of multiclass [8] SVM [9], using default parameters. The result of the training phase is an SVM classifier that can classify previously unseen vectors.

In the second online stage we classify a vector every TR (2 seconds) and use the same noise reduction, eye filtering and normalization methods as in the training stage, selecting the same voxels based on the IG filtering performed at model training. Finally, the data is passed into the trained SMO model, and the classification result is then transmitted to the external device. The training process takes several minutes and the classification process uses 1-5% of the CPU and lasts approximately 50 milliseconds.

Here we compare our results with our previous study, in which we introduced a simple classification method used to successfully classify intentions of motor movement or imagery [5]. In that method, we manually localized the subject's relevant brain areas corresponding to activation of left and right hand and legs, and the algorithm selected online the class that corresponded to the ROI with maximum Z-score normalized averaged value.

<sup>1</sup><http://medical.nema.org/>

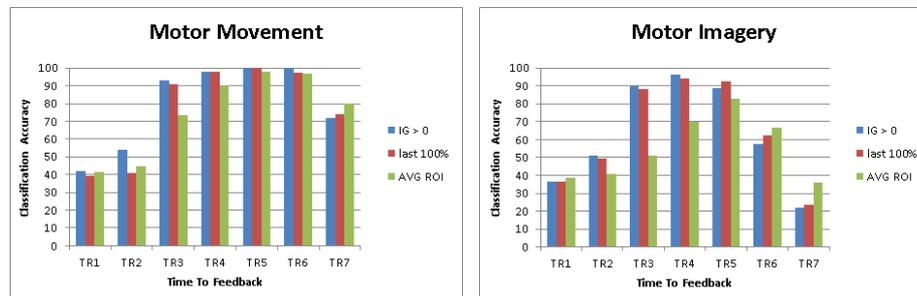


Figure 1: A comparison of MM and MI classification accuracy across six (MM) and three (MI) subjects, between ML and ROI. The ML results were obtained by using (a) all voxels with IG above 0 and (b) the smallest number of voxels that permit perfect classification of all training examples.

### 3 The Experimental Protocols

We tested our system in three areas: i) motor motion (MM, 5 subjects), ii) motor imagery (MI, 3 subjects), and (iii) visual categories (VIS, 5 subjects). In MM and MI, we classified left hand, right hand, and legs movement or imagery; subjects were allowed to move their fingers and toes in the MM condition. In VIS we had subjects watch four visual categories: faces, houses, tools, and a fixation screen that corresponds with idle viewing. In both the MM and MI studies the subjects see an avatar standing in the center of a three-door room. In each run of the cue-based experiment, the subjects received equal amount of triggers from each class, in a pseudo-random order. The total number of triggers was determined based on 10-12 minute run. For MM, MI, and VIS we used 30, 45, and 40 samples per class respectively (in MI one subject had performed the runs earlier in the research and had only 30 samples per class). We used only three runs for model training in MM due to a higher signal to noise ratio, while for MI and VIS we used four runs; the training and testing data were kept in a chronological order. Written informed consent was obtained from all volunteers, and the studies were approved by the Ethics Committee of the Weizmann Institute of Science.

Our results were obtained offline from recorded data. However, the calculation was simulated by our real-time system as if it were an actual real-time experiment; our system received and classified a new DICOM file every 2 seconds, as would be the case in a real-time study.

#### 3.1 Results

Figure 1 compares average classification accuracy across subjects between ML and ROI. Classification accuracy coincides with the hemodynamic response: in TRs 1 and 2 the accuracy is around chance level, then it gradually increases with the best accuracy in TR 3 to TR 5-6 (6-12 seconds after the trigger), and gradually drops to chance level. The results indicate that identifying the most relevant features by using IG and classifying using SMO is superior to the ROI method. In MM above 90% accuracy can be achieved even at TR 3 (6 seconds after a cue), which is better than the 10 seconds delay in the ROI method. In MI, accuracy of 90% in TR3 and 95% in TR4 can be achieved, which is higher than the highest ROI accuracy at TR 5. Figure 2 provides the results of the VIS study, indicating 78% in TR3. Accuracy up to 87% can be reached with a longer delay. Finally, by reducing the voxel count (i.e., raising the IG threshold) to the smallest number of voxels that permit perfect classification of all training examples, we can achieve similar accuracy to that achieved by using all voxels with positive IG, but with greatly reduced computation time.

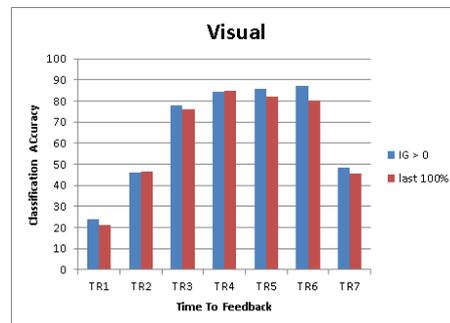


Figure 2: ML results obtained for VIS using (a) all voxels with IG above 0 and (b) the smallest number of voxels that permit perfect classification of all training examples

## 4 Discussion

Our results indicate that the system is robust across subjects, and efficient in classification time and CPU usage. We have identified multi-voxel brain patterns in the primary motor cortex, the fusiform face area (FFA), the parahippocampal place area (PPA), and the lateral occipital object area (LO). Future experiments may lead to identifying more complex brain patterns. We show successful classification in three different tasks in two several brain areas, using the same method. The fact that this system works online opens the door to new paradigms in BCI, NF, and brain rapid brain mapping.

## 5 Acknowledgements

This research was supported by the EU project VERE (number 257695), [www.vereproject.eu](http://www.vereproject.eu).

## References

- [1] S.M. Coyle, T.E. Ward, and C.M. Markham. Brain-computer interface using a simplified functional near-infrared spectroscopy system. *Journal of neural engineering*, 4:219, 2007.
- [2] R. Sitaram, A. Caria, and N. Birbaumer. Hemodynamic brain-computer interfaces for communication and rehabilitation. *Neural networks*, 22(9):1320–1328, 2009.
- [3] S. Lee, S. Halder, A. Kübler, N. Birbaumer, and R. Sitaram. Effective functional mapping of fmri data with support-vector machines. *Human brain mapping*, 31(10):1502–1511, 2010.
- [4] I.H. Witten, E. Frank, and A.M. Hall. *Data mining: Practical machine learning tools and techniques*, 3rd edition. 2011.
- [5] O. Cohen, S. Druon, S. Lengagne, A. Mendelsohn, R. Malach, A. Kheddar, and D. Friedman. fmri robotic embodiment: A pilot study. In *Biomedical Robotics and Biomechanics (BioRob), 2012 4th IEEE RAS EMBS International Conference on*, pages 314–319, June 2012.
- [6] O. Cohen, M. Koppel, R. Malach, and D. Friedman. Controlling an avatar by thought using real-time fmri. *Journal of neural engineering*, 11(3):035006, 2014.
- [7] J.R. Quinlan. Induction of decision trees. *Machine Learning*, 1(1):81–106, 1986.
- [8] T. Hastie and R. Tibshirani. Classification by pairwise coupling. In I.M. Jordan, J.M. Kearns, and A.S. Solla, editors, *Advances in Neural Information Processing Systems*, volume 10. MIT Press, 1998.
- [9] J. Platt et al. Sequential minimal optimization: A fast algorithm for training support vector machines. 1998.

# Online detection and classification of movement kinetics

Mads Jochumsen<sup>1</sup>, Muhammad Samran Navid<sup>1,3</sup>, Rasmus Wiberg Nedergaard<sup>1</sup>, Muhammad Nabeel Anwar<sup>3</sup>, Imran Khan Niazi<sup>1,2</sup> and Kim Dremstrup<sup>1</sup>

<sup>1</sup>Department of Health Science and Technology, Aalborg University, Denmark.

<sup>2</sup>New Zealand College of Chiropractic, Auckland, New Zealand.

<sup>3</sup>Department of Biomedical Engineering & Sciences, National University of Sciences & Technology, Pakistan.

mj@hst.aau.dk, samran@hst.aau.dk, rneder09@student.aau.dk  
nabeel@smme.nust.edu.pk, imrankn@hst.aau.dk, kdn@hst.aau.dk

## Abstract

Over the past years brain-computer interface (BCI) technology has been proposed as a means for neurorehabilitation. To induce Hebbian-associated-like plasticity, the movement-related cortical potential (MRCP) can be detected from the continuous brain activity to trigger timely appropriate inflow of somatosensory feedback from electrical stimulation. The aim of this study was to detect the MRCP online from the continuous brain activity and decode two types of movements that were performed with different levels of force and speed (2x50 movements). 5 healthy subjects and 1 stroke patient performed/attempted to perform the movements. The system correctly detected and classified 65±3 % and 51 % of the movements for the healthy subjects and patient, respectively. The findings suggest that it is possible to detect movements and decode kinetic information online. This may have implications for stroke rehabilitation where task variability may be introduced to optimize the retention of relearned movements.

## 1 Introduction

Stroke is the main cause of adult disability in high-income countries worldwide; therefore, several techniques have been proposed to reverse motoric impairments [1]. Over the past years, brain-computer interfaces (BCIs) have been proposed as tools that can be used in neurorehabilitation [2, 3]. Recently, it was shown that plasticity could be induced by using the movement-related cortical potential (MRCP) as a control signal in a BCI that provides somatosensory feedback from electrical stimulation [4, 5]. The MRCP is a low-frequency brain potential associated with executed and imagined movements [6]. The potential can be observed in the electroencephalogram (EEG) up to two seconds before a voluntary movement; therefore, it has been proposed for BCIs that are used for induction of Hebbian-associated-like plasticity [4, 5]. The MRCP also contains kinetic information of the executed or imagined movement such as the level of force and speed [6, 7]. This kinetic information may be utilized in designing more sophisticated rehabilitation protocols where variety of tasks is required. It has been shown that variety of tasks in a rehabilitation paradigm is good for maximizing the retention of relearned motor skills [8]. To implement this in BCI protocols, we need to detect the movement intention and classify different task parameters (force and speed) to activate

correlated somatosensory feedback (through functional electrical stimulation). The detection and decoding of the MRCP have been performed offline in a previous study [7].

The aim of this study was to implement and test the feasibility of detecting the MRCP and discriminating between fast movements with a high level of force and slow movements with a low level of force.

## 2 Methods

### 2.1 Subjects

Five healthy volunteers (1 female and 4 males:  $29 \pm 5$  years old) and one stroke patient with lower limb paresis (77 years old, male, infarction, right side affected, 46 days since event) participated. All the subjects gave their informed consent before participation, and the procedures were approved by the local ethical committee (N-20130081).

### 2.2 Experimental protocol

Each subject was seated in a chair with the right foot fixed to a pedal with an attached force transducer. The experiment was divided into two sessions; training and testing. The training session started with recording of the maximum voluntary contraction (MVC) force followed by 50 repetitions of cued isometric dorsiflexions of the ankle for each of two movement types. The two tasks were [7]: i) 3 s to reach 20% of MVC and ii) 0.5 s to reach 60% of MVC. The subjects spend ~5 min to familiarize with each task. The order of the two tasks was randomized, but they were not mixed. To assist the subjects in performing the movements with the correct level of force and speed, they were visually cued by a custom made program (Knud Larsen, SMI, Aalborg University), where force was used as input. They were asked to produce force to match a ramp trace.

After the training session, the detector (Section 2.4) and classifier (Section 2.5) were built and the testing session started. In this session, the subjects performed 50 movements of each movement type randomly and in their own pace (they were instructed to separate two consecutive movements with at least 5 s, and the experimenter guided them at the end of the session, so an equal amount of movements were performed). After they performed a movement they verbally expressed the movement type that was performed; this was noted and compared to the outcome of the computer prediction.

### 2.3 Signal acquisition

Ten channels of monopolar EEG were recorded (EEG amplifiers, Nuamps Express, Neuroscan) continuously from FP1, F3, Fz, F4, C3, Cz, C4, P3, Pz, and P4 according to the International 10-20 system (32 Channel Quick-Cap, Neuroscan). The signals were sampled at 500 Hz and analog to digital converted with 32 bits accuracy. The signals were referenced to the right ear lobe. Electrooculography (EOG) was recorded from FP1. The impedance of the electrodes was below 5 k $\Omega$  during the experiment.

### 2.4 Detection, feature extraction and classification

In the training and testing session, the EEG signals were bandpass filtered from 0.05-10 Hz with a 2<sup>nd</sup> order Butterworth filter. A surrogate channel of the nine EEG channels was obtained by applying an optimized spatial filter to improve the signal-to-noise ratio (SNR) as proposed by Niazi et al. [9]. The movements were detected using a template matching technique [7, 9] where a template of the

initial negative phase of the MRCP was matched to the surrogate channel. The template was extracted from an average of the 2x50 movements that were performed in the training session. The length of the template was 2 s, and it was extracted from the peak of maximum negativity and 2 s prior to this point. The detection threshold was obtained for each subject using a receiver operating characteristic (ROC) curve. The ROC curve was generated through 3-fold cross-validation of the training data. The detection threshold was selected to maximize the true positive rate (TPR), but on the expense of more false positive detections (FPs). Detections occurred when the cross-correlation, computed between the template and the surrogate channel, exceeded the detection threshold, and the EOG activity was lower than 125 $\mu$ V. Detector decisions were made every 200 ms. To reduce the number of FPs, the detector was disabled for 3 s after detection. The detection was evaluated through the TPR and number of FPs/min. Six temporal features were extracted from the initial negative phase of the MRCP from the detection onset and 2 s prior this point. These features were: i+ii) slope and intersection of a linear regression fitted to the entire interval (-2 s until the detection onset), iii+iv) slope and intersection of a linear regression fitted to the last 0.5 s of the interval (-0.5 s until the detection onset), v) maximum negative amplitude and vi) mean amplitude.

The features were classified using a support vector machine (SVM) with a linear kernel. All trials from the training session were used to build the classifier that was used in the testing session. The classification accuracy was obtained for the correctly detected movements.

### 3 Results

The results are summarized in Table 1. The TPR was 85 $\pm$ 4 % for the healthy subjects and 85 % for the stroke patient, and less than 1 FP/min was registered. The number of correctly detected and classified movement was higher for the healthy subjects (65 $\pm$ 3 %) compared to the patient (51 %).

Healthy/stroke subject	Detection [%]	Detection and classification	FPs/min
H 1	87	62	0.6
H 2	90	70	0.6
H 3	82	65	0.2
H 4	78	61	0.1
H 5	89	67	2.4
Mean $\pm$ SD	85 $\pm$ 4	65 $\pm$ 3	0.8 $\pm$ 0.8
S 1	85	51	0.9

**Table 1:** Performance of the system for healthy subjects and the stroke patient. 'Detection' is the TPR, and the column to the right is the performance when the detected movement is classified.

### 4 Discussion

In this study, movements were detected and the kinetic information classified for healthy subjects with a performance of 65 $\pm$ 3 % correctly detected and classified movements. The proposed techniques were also tested by a stroke patient where 51 % of the movements were correctly detected and classified.

The detection performance was slightly higher compared to what has been found in previous studies where an online system was simulated [7, 9]; this may be explained by the selection of the

detection threshold. In the previous studies the detection threshold was based on the midpoint of the upward convex part of the ROC curve to obtain a tradeoff between the TPR and number of FPs. In this study, the detection threshold was selected to increase the TPR, but on the expense of more FPs. The number of FPs was accounted for by disabling the detector for 3 s after detection.

The performance of the classifier performed slightly worse than in the offline studies which may be explained by the lower detection threshold leading to an earlier detection of the movements and therefore less kinetic information can be extracted from the 2 s of data extracted prior the detection point. Detection latencies were not calculated in the current study, but it is expected to be in the range of  $\pm 100$  ms [9].

The findings suggest that the BCI system can be used for neuromodulation where task variability can be introduced (two classes). The system performance is in the range of what has been reported to induce plastic changes [4], although it remains an open question what the lower limit is for inducing plasticity [3]. The effect on the induction of plasticity should be investigated to see if current BCI protocols for this purpose can benefit from the introduction of task variability.

## 5 Acknowledgement

The authors would like to thank Physiotherapist Christoffer Wang, Træningsenheden Aalborg, for recruiting the stroke patient.

## References

- [1] P. Langhorne, F. Coupar and A. Pollock, "Motor recovery after stroke: a systematic review," *Lancet Neurology*, vol. 8, pp. 741, 2009.
- [2] J. J. Daly and J. R. Wolpaw, "Brain-computer interfaces in neurological rehabilitation," *The Lancet Neurology*, vol. 7, pp. 1032-1043, 2008.
- [3] M. Grosse-Wentrup, D. Mattia and K. Oweiss, "Using brain-computer interfaces to induce neural plasticity and restore function," *Journal of Neural Engineering*, vol. 8, pp. 025004, 2011.
- [4] I. K. Niazi, N. M. Kersting, N. Jiang, K. Dremstrup and D. Farina, "Peripheral electrical stimulation triggered by self-pace detection of motor intention enhances corticospinal excitability," *IEEE Transaction on Neural Systems and Rehabilitation Engineering*, vol. 20, pp. 595-604, 2012.
- [5] N. Mrachacz-Kersting, S. R. Kristensen, I. K. Niazi and D. Farina, "Precise temporal association between cortical potentials evoked by motor imagination and afference induces cortical plasticity," *J. Physiol. (Lond.)*, vol. 590, pp. 1669-1682, 2012.
- [6] O. F. Nascimento, K. D. Nielsen and M. Voigt, "Movement-related parameters modulate cortical activity during imaginary isometric plantar-flexions," *Experimental Brain Research*, vol. 171, pp. 78-90, 2006.
- [7] M. Jochumsen, I. K. Niazi, N. Mrachacz-Kersting, D. Farina and K. Dremstrup, "Detection and classification of movement-related cortical potentials associated with task force and speed," *Journal of Neural Engineering*, vol. 10, pp. 056015, 2013.
- [8] J. W. Krakauer, "Motor learning: its relevance to stroke recovery and neurorehabilitation," *Curr. Opin. Neurol.*, vol. 19, pp. 84-90, 2006.
- [9] I. K. Niazi, N. Jiang, O. Tiberghien, J. F. Nielsen, K. Dremstrup and D. Farina, "Detection of movement intention from single-trial movement-related cortical potentials," *Journal of Neural Engineering*, vol. 8, pp. 066009, 2011.

# Multiple roles of ventral premotor cortex in BCI task learning and execution

Jeremiah Wander<sup>1</sup>, Rajesh P.N. Rao<sup>2</sup> and Jeffrey Ojemann<sup>3</sup>

<sup>1</sup>Department of Bioengineering, University of Washington, Seattle, WA

<sup>2</sup>Department of Computer Science and Engineering, University of Washington, Seattle, WA

<sup>3</sup>Department of Neurological Surgery, University of Washington, Seattle, WA

jdwander@uw.edu, rao@cs.washington.edu, jojemann@uw.edu

## Abstract

Motor-based brain-computer interface (BCI) use is a learned skill that has been shown to involve multiple cortical areas, though it only explicitly requires modulation of a small area of cortical tissue. The roles being played by these other areas have not yet been determined. In this study, using an electrocorticographic (ECoG) model, we apply a novel computational approach to quantifying the strength and nature of interactions between putative primary motor cortex and ventral premotor cortex (PMv) during BCI use and across BCI skill development. Our findings suggest multiple roles being carried out in PMv that change in strength and nature during skill acquisition.

## 1 Introduction

Brain-computer interfaces (BCIs) show great promise both as a medical technology and a platform from which to investigate the adaptive capacity of the nervous system (Wander & Rao, 2014). Though it has been shown previously that many cortical (Ganguly et al., 2011; Wander et al., 2013) and subcortical (Koralek et al., 2012) areas are active during use of a motor-based BCI, and that as subjects gain experience activity in some of these areas lessens, it is still uncertain what role these areas play in the initial learning and subsequent execution of the neuroprosthetic skill.

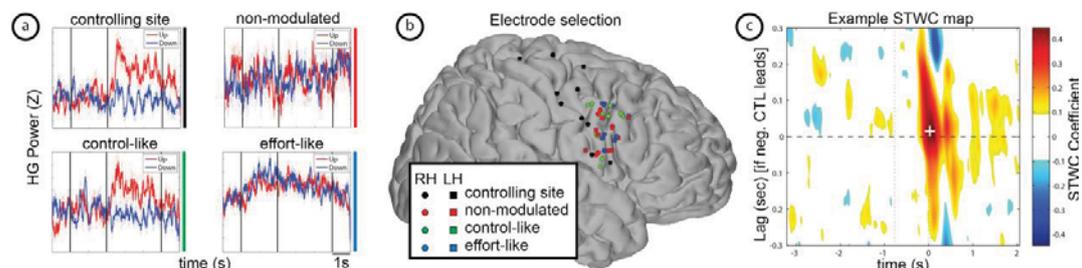
Qualitative review of activity patterns in motor cortical networks during BCI use has revealed that some areas show increased activity only during active execution of motor imagery or overt movement whereas others show increased activity during the task, regardless of the imagery state of the user (Wander et al., 2013). In this study, we employ a newly developed computational method to investigate the roles being played by ventral premotor cortex (PMv) during BCI use. PMv was selected as a cortical area of interest because of its previous implication in effective motor skill execution and adaptation (Xiao et al., 2006). We hypothesize that high-gamma (HG; 70-200 Hz) activity in PMv and HG interactions with the controlling electrode (CTL) will demonstrate multiple roles being carried out by PMv during BCI use. Further, we hypothesize that some of these PMv-to-CTL interactions will lessen in strength as subjects develop automaticity in BCI task execution.

## 2 Methods

We have previously published methods describing our subject population, ECoG recording methods and right justified box (RJB) task structure (Wander et al., 2013), but will briefly summarize and discuss additional points of note.

The RJB is a one-dimensional BCI task where HG power in a single ECoG electrode is mapped to velocity of a cursor in the control dimension. The cursor travels across the workspace at a fixed horizontal velocity while the user controls its vertical velocity to attempt to move it in to a target area. This is a retrospective analysis of previously collected BCI data. Subject selection criteria were as follows: subjects needed to (a) have conducted at least 40 trials of RJB task, (b) performed above chance levels on the task, and (c) have electrode coverage of PMv. This resulted in inclusion of 10 subjects in the current study. Labels for individual electrodes were estimated using the human motor area template (HMAT) (Mayka, et al., 2006). From all the electrodes recorded ( $N=792$ ), we selected only electrodes that were being used for BCI control ( $N=10$ ) and non-control electrodes determined to be over PMv ( $N=39$ ) for subsequent analyses. The classifications of PMv electrodes were as follows: electrodes were considered *control-like* if they exhibited significant HG increases during the feedback period (relative to rest) for up targets, but not for down targets; they were considered *effort-like* if they exhibited significant HG increases during the feedback period for both up and down targets; and they were considered *non-modulated* if they did not exhibit task-positive modulation.

We assessed transient temporal correlations in HG activity between the CTL and PMv electrodes using short-time windowed covariance/correlation (STWC) (Blakely et al., 2014). This method is specifically suited to teasing out amplitude-amplitude interactions (correlations) in neural signals that are not only transient (i.e. event-driven), but also potentially occur at slightly different points in time in each of the two signals. STWC maps were calculated on smoothed (~50 msec FWHM) HG power, using a window width of 500 msec and a maximum lag of 300 msec. STWC maps were then generated by realigning individual maps based on HG onsets and averaging across all trials. To isolate interactions occurring near the time period of HG onsets ( $\pm 500$  msec), we finally extracted the maximum STWC coefficient and corresponding lag from each average map. Significant interactions were selected using a phase-randomized, surrogate-signal bootstrapping approach where STWC coefficients were considered significant if they had less than a 5% chance of occurring in the distributions of maximal STWC coefficients.



**Figure 1 - (a) Example average time series HG activity showing the four electrode classes discussed in this manuscript and (b) anatomical location of all electrodes considered in these analyses. (c) An example STWC map, thresholded at  $\pm 1$  standard deviation from the mean. Lags (y-axis) denote the temporal relationship between the two signals being considered. The white (+) denotes an example peak lag and STWC coefficient that was extracted from the map.**

### 3 Results

**Behavioral performance.** By design, all subjects performed the BCI task at above chance level, as was discussed previously. Mean task performance was 74.2% ( $\sigma=8.57\%$ ,  $N=10$ ). For individual subjects, task performance ranged from 60% to 86.2%.

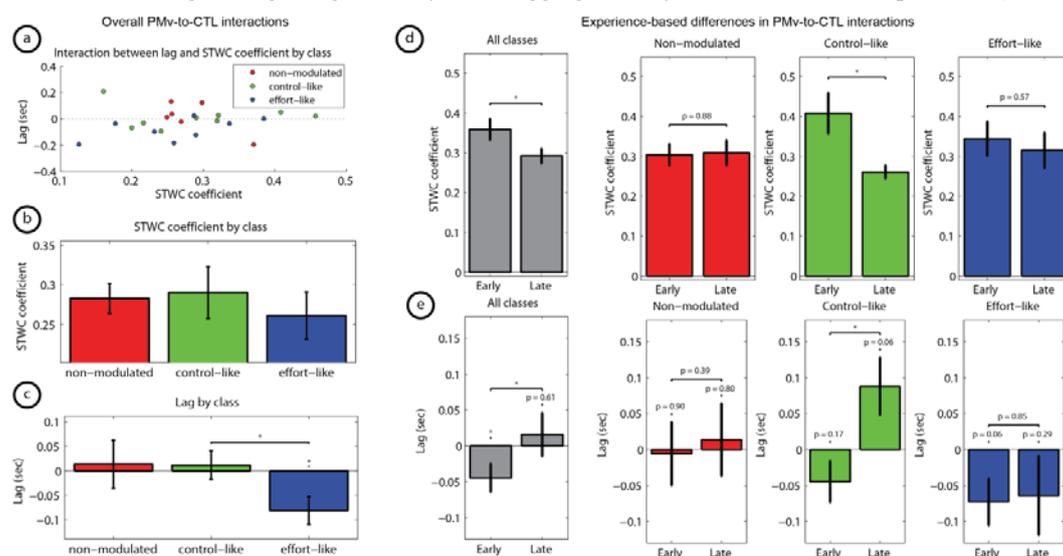
**HG activity at CTL and PMv.** All subjects were able to significantly modulate HG power at CTL for up targets relative to rest periods (two sample student's t-test,  $p<0.05$  in all cases, employing

within-subject false discovery rate [FDR] correction). Each subject had at least one electrode over PMv that exhibited significant modulation during the feedback period relative to rest. All PMv electrodes were classified according to the methodology described above. There were 14 electrodes classified as *non-modulated*, 13 electrodes classified as *control-like*, and 12 electrodes classified as *effort-like*.

**PMv-to-CTL HG interactions.** In looking at the STWC maps generated for the multiple PMv-to-control interactions, it was clear that there is no single role carried out by PMv during BCI task execution. Figure 1 gives a clear example of PMv becoming active before CTL, but other examples exist showing both contemporaneous and lagging interactions. To quantify these relationships, we used the STWC peak-finding approach described above to determine single lag values relative to HG onset for all electrodes meeting the STWC threshold criteria. 23 of the 39 PMv electrodes exhibited significant interactions with CTL ( $p < 0.05$ ; previously described bootstrap approach).

Figure 2 depicts the lags and interaction strengths for each of these 23 electrodes. When considered together, peak covariance lags are not statistically different from zero (one-sample two-tailed t-test,  $N = 39$ ,  $p = 0.3157$ ), however, there is an interesting relationship between PMv electrode classes and covariance lags. Specifically, electrodes classified as being *effort-like* significantly lag CTL by an average 81.1 msec ( $\pm 29$  msec s.e.m.; one-sample two-tailed t-test,  $N=8$ ,  $p=0.0267$ ) and at lags significantly lower than *control-like* electrodes (two-sample two-tailed t-test,  $N=[9, 8]$ ,  $p=0.0398$ ). Neither *control-like* nor *non-modulated* electrodes demonstrated lags that were significantly different from zero, nor were they significantly different from each other. We also observed that *effort-like* electrodes showing significant interactions were typically located more anteriorly and superiorly within PMv, though we recognize that the transformations necessary to map electrode locations to common brain space are approximations and may not be appropriate to uncover spatial relationships on such a small scale.

The above results suggest that there are multiple meaningful interactions that take place between PMv and M1 during the execution of a BCI task, yet it remains an open question as to which of these interactions are involved in the continued execution of the task, and which are specifically associated with the original acquisition of the neuroprosthetic skill. With that purpose in mind, we divided all trials in half chronologically into early and late trials, and re-evaluated STWC maps on these subgroups. Overall, PMv-to-CTL STWC strengths decreased significantly ( $p=0.0301$ ) from early to late trials, and lags changed significantly from lagging CTL by 44.2  $\pm$  19.7 msec ( $p=0.0305$ ) to not



**Figure 2 - STWC results. (a) Scatter plot showing lags and STWC coefficients for all PMv-to-CTL interactions, taken across all trials. (b-c) By-class comparison of distributions for STWC coefficients and lags, respectively. (d) Early-to-late comparison of STWC coefficients overall and by electrode class. (e) Early-to-late comparison of lags overall and by electrode class. See text for summary and discussion.**

statistically lagging. When breaking down electrodes by class, we found significant decreases in STWC for the *control-like* PMv electrodes ( $p=0.0419$ ), a significant change in lags for *control-like* electrodes (from lagging to leading CTL,  $p=0.0139$ ), and no change in lags for *effort-like* electrodes.

## 4 Discussion

In this brief manuscript we have demonstrated that there are multiple statistically significant and meaningfully different roles being performed by PMv during BCI task execution and that the nature of at least one of these roles changes as BCI users gain task experience. These findings suggest that previously demonstrated distributed activity changes seen over the course of BCI skill acquisition may not be solely attributable to optimization of cortical activity, but that they may also play a meaningful role in task learning and execution.

It is important to note that interactions evaluated by the STWC method are correlative in nature and not causal. Interventional studies specifically altering function in one or more of these distributed cortical areas will be necessary to understand the causal interactions between these regions.

Ventral premotor cortex was an excellent first candidate for this type of analysis, but as we have demonstrated previously, other distributed cortical areas also appear to be task-modulated during BCI use. A necessary extension to this work will be to evaluate the activity and interaction patterns taking place in these areas. Additionally, for the sake of brevity, we have limited our analyses simply to HG-HG interactions. There may be an excellent opportunity to uncover some of the neural underpinnings of so-called BCI illiteracy by expanding the field of view to include cross-frequency interactions between HG and the sensorimotor rhythms commonly used in non-invasive BCIs.

## References

- Blakely, T., Ojemann, J. G., & Rao, R. P. N. (2014). Short-time windowed covariance : A metric for identifying non-stationary , event-related covariant cortical sites. *J Neurosci Meth*, 222, 24–33.
- Ganguly, K., et al. (2011). Reversible large-scale modification of cortical networks during neuroprosthetic control. *Nature Neuroscience*, 14(5), 662–7.
- Koralek, A. C., Jin, X., Long II, J. D., Costa, R. M., & Carmena, J. M. (2012). Corticostriatal plasticity is necessary for learning intentional neuroprosthetic skills. *Nature*, 483, 331–335.
- Mayka, M., et al. (2006). Three-dimensional locations and boundaries of motor and premotor cortices as defined by functional brain imaging: a meta-analysis. *Neuroimage*, 31(4), 1453–1474.
- Wander, J. D., et al. (2013). Distributed cortical adaptation during learning of a brain-computer interface task. *Proceedings of the National Academy of Sciences*, 110(26), 10818–10823.
- Wander, J. D., & Rao, R. P. (2014). Brain-computer interfaces: a powerful tool for scientific inquiry. *Current Opinion in Neurobiology*, 25, 70–75.
- Xiao, J., Padoa-Schioppa, C., & Bizzi, E. (2006). Neuronal correlates of movement dynamics in the dorsal and ventral premotor area in the monkey. *Exp Brain Research*, 168(1-2), 106–19.

# Echo State Networks for Brain Computer Interface Classification

Ryan McCardle<sup>1</sup> and Claire Davies<sup>2</sup>

<sup>1</sup> The University of Auckland, Auckland, New Zealand  
ryan.mccardle@auckland.ac.nz

<sup>2</sup> The University of Auckland, Auckland, New Zealand  
c.davies@auckland.ac.nz

## Abstract

Echo state networks are good candidates for use as classifiers in brain computer interfaces. They can also be used to explore the data to some extent by manipulating their training. In this paper similar results were achieved to those in the Brain Computer Interface Competition IV in classifying four-class movements from magnetoencephalography.

## 1 Introduction

Artificial neural networks consist of a series of many *neurons* which are connected to each other in a network. Data is fed into the network and propagates from neuron to neuron. Each neuron generates an output from all the combined inputs according to an activation function built into it. How much influence each input to the neuron has is determined by a *weight* placed on each connection. The internal values of a number of special *output neurons* are the final result of the network. Most neural networks are *feed-forward*, meaning the propagation is in one direction (no feedback). A recurrent neural network (RNN) has a memory because the inputs propagate in all directions, blending with earlier inputs and being slowly drowned out by new inputs.

In a RNN the weights of all the connections between the neurons are modified. This creates a very complex non-linear optimization problem which can be very difficult to solve. RNNs and echo state networks (ESNs) are very similar architecturally, but differ in training. In an ESN the weights of all the connections between the neurons are randomly generated at the start of training and set permanently. This creates a large *reservoir* (i.e. reservoir computing) of randomly connected neurons with random weights. Only the weights leading to the output neurons are modified in training, turning a complex non-linear optimization problem into a simple one solvable with linear regression [4].

ESNs are good candidates for brain computer interfaces (BCIs). They perform well in tasks such as speech and handwriting recognition where there is high variability in the input, and a noisy signal. Their *memory* allows recognition of events over time (e.g. slow rise and fall in electrical potential), and events in sequence (e.g. propagation of brain activity). ESNs naturally produce a time series output, allowing continuous BCI output such as arm velocity over time. Crucially, they are easy to train; a dataset can be plugged into an ESN classifier with little pre-processing, and produce a classification level above chance. ESNs have performed well when compared to more commonly used classification techniques in several BCI studies [2, 1, 5].

ESNs are *black box* classifiers i.e. it is difficult or impossible to determine how the input signal is manipulated to produce the output. This is not much of a problem in practical applications, but can make it difficult to optimise the BCI in research. The classifier can also *cheat*. For example, the classifier could exploit irregularities in the data, instead of recorded signals. Non-brain signals could also be exploited e.g. blinking, or bumping EEG cables.

We propose that with careful manipulation of the training these problems can mostly be avoided. For example, if the classification accuracy is low when trained only on non-brain signals and high when trained on signals from anatomically relevant brain areas, we can infer that cheating is not occurring. The characteristics of the signal the classifier is using can also be approximated by removing parts of the dataset and evaluating the classification accuracy. The purpose of this paper is to determine the effectiveness of ESNs in a 4-class BCI classification problem, also to investigate this manipulation of the training.

## 2 Methods

### 2.1 Training Dataset

Recordings from participants were extracted from one of the datasets distributed as part of the BCI Competition IV [6]. To generate this dataset two right handed subjects moved a joystick 4.5 cm from the centre position in self-chosen cardinal directions. Trial onset was cued with a grey circle. After a variable delay the grey circle would disappear signalling the subject to move. During the trial the subjects were instructed to fixate on a red cross and not to blink.

Magnetoencephalography (MEG) was recorded at 625 Hz, using ten channels (LC21, LC22, LC23, LC31, LC32, LC41, LC42, RC41, ZC01, ZC02) located above the motor areas. The MEG signals were band-pass filtered 0.5 Hz to 100 Hz and re-sampled at 400 Hz. The recordings were then cut into 1 s second trials, with movement onset at 200 ms. The training data set contained 40 trials for each of the four classes per subject, the testing dataset contained 73 trials (true classification unknown during the BCI Competition). As the BCI Competition is over, true labels are available for the testing dataset, and so the training and testing datasets were combined to enlarge the training dataset. The trial order was randomised before training.

### 2.2 Classifier Training

MEG data was processed with the open-source OrGanic Environment for Reservoir computing (Oger) toolbox that provides a robust implementation of ESNs. To prevent saturation of the network the mean was shifted to zero for each channel. It was then normalised to  $\pm 1$ , to keep it in the range of the neuron activation function. Data before movement onset was cut because Waldert et al. [6] found that classification was at chance level until then.

ESNs cannot be used directly for classification as the output neurons produce fluctuating values. By representing the movement direction in the training data as four outputs from the ESN (1.0 for the desired class), the channel with the highest output became the classification. Standard deviation was reported to reflect the varying performance of the randomly generated reservoirs. The accuracy of the trained ESN was determined by ten-fold cross validation.

The time taken to train a network increases exponentially with the reservoir size (Figure 1). To achieve sufficient accuracy with a reasonable training time, a reservoir size of 500 neurons was chosen. The remaining parameters of the ESN were optimised by sweeping through a range of values and choosing those with the highest classification accuracy. This gave final values of 0.2 for the leaking rate, 0.1 for the input scaling, and 0.9 for the spectral radius.

## 3 Results

Figures 1 to 3 show results generated during the optimisation of the classifier, and the final results in 4. Figure 2 shows the accuracy when the network is trained on each electrode indi-

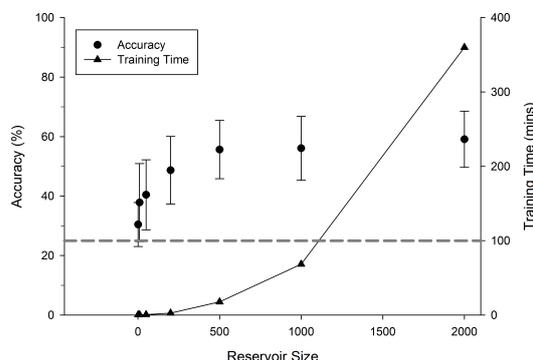


Figure 1: Size of echo state reservoir and its influence on accuracy and training time.

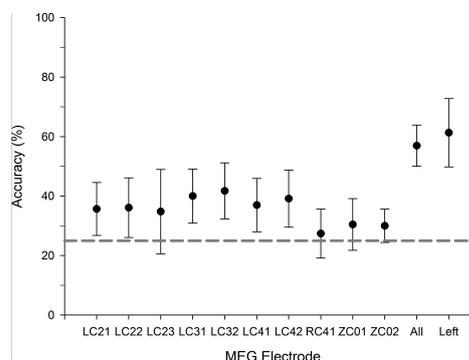


Figure 2: Accuracy for each electrode, all electrodes, and left hemisphere only.

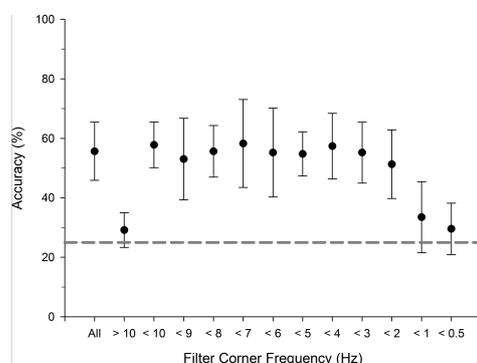


Figure 3: Accuracy of unfiltered data compared to filtering with a 3rd order Butterworth filter at different thresholds.

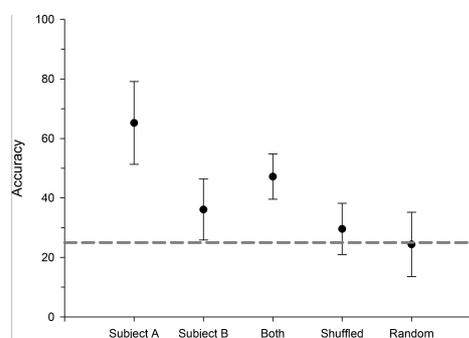


Figure 4: Final accuracy for subjects individually and combined, compared to shuffled and randomised classifiers.

vidually, all the electrodes, and only the electrodes on the left hemisphere. The left electrodes all perform similarly at around 35 % accuracy. The electrodes located along the mid-line, and the right hemisphere (RC41, ZC01, ZC02) perform at close to chance level, suggesting that the classification is not due to EMG from the right wrist. This difference in accuracy also indicates that the classifier is not using EOG artefacts, as these artefacts would effect both hemispheres equally. An accuracy of 57 % is achieved when all electrodes are combined, and with only the electrodes of the left hemisphere accuracy increases to 61 %. A different subset of the electrodes may perform better, but there are too many combinations to test exhaustively.

When the MEG signal is high-pass filtered at 10 Hz the classification accuracy drops off to chance level (Figure 3). When low-pass filtered the accuracy remains the same until it drops off at 2 Hz to 3 Hz. From these drops in accuracy we can infer that the signal being used by the ESN for classification is in the range 1 Hz to 3 Hz. However, this may be because the ESN used is most sensitive to frequencies in this range, and using different parameters when creating the ESN may yield different results.

Using the results described in Figures 1 to 3, only the subjects' left hemisphere electrodes were used, and then bandpass filtered (1 Hz to 3 Hz). The final classification accuracy was compared to shuffled and random classifiers (Figure 4). Shuffling the inputs so they do not match the outputs should make it impossible for the ESN to produce the correct output, and

it should function like a random classifier with a mean accuracy of 25%. In this case the average accuracy for shuffled data is 27% - indicating that there are some irregularities in the dataset which can be exploited by the ESN. Therefore this is a better benchmark for chance level accuracy, which the reported results still exceed.

## 4 Discussion

Using the methods described, a classification accuracy of 65% was achieved for subject 1, and for subject 2 a lower classification accuracy of 36% (Figure 4). This is higher than, but consistent with, the winning BCI Competition IV entrant who achieved 59.5% accuracy for subject 1, and 34.3% for subject 2, with a smaller training dataset and without the benefit of checking their results [3]. The other entrants did not achieve results above chance level.

Manipulating the training dataset showed that brain activity related to the task was localised to the left hemisphere. More training rounds, and a higher electrode density could potentially be used to localise it more precisely. The approximate frequency of the activity was also determined, which shows the potential for using an ESN as a crude way to investigate brain activity with unknown characteristics, or find activity in new frequency bands.

When the ESN was trained with both subjects data combined it achieved an accuracy of 47%. This is unusual as typically BCIs must be trained for each individual due to variation in brain activity. This may suggest that a single ESN BCI can be trained to work with multiple subjects, or be generalised to work with any subject. However it is impossible to tell without data from more subjects. A result of 47% is comparable to the accuracy from simply combining the results from both subjects. This means the ESN may simply have been trained to differentiate between subjects and classify them accordingly (this in itself would be interesting).

## References

- [1] GP Cédric, M Sebag, A Souloumiac, and A Larue. Ensemble learning for brain computer-interface using uncooperative democratic echo state communities. In *Cinquième conférence plénière française de Neurosciences Computationnelles*, 2010.
- [2] Aysegul Gunduz, Mustafa C. Ozturk, Justin C. Sanchez, and Jose C. Principe. Echo State Networks for Motor Control of Human ECoG Neuroprosthetics. *2007 3rd International IEEE/EMBS Conference on Neural Engineering*, pages 514–517, May 2007.
- [3] Sepideh Hajipour Sardouie and Mohammad Bagher Shamsollahi. Selection of Efficient Features for Discrimination of Hand Movements from MEG Using a BCI Competition IV Data Set. *Frontiers in neuroscience*, 6(April):42, January 2012.
- [4] Herbert Jaeger. The "echo state" approach to analysing and training recurrent neural networks - with an erratum note. Technical report, German National Research Center for Information Technology, St. Augustin, 2001.
- [5] Pieter-jan Kindermans, Pieter Buteneers, David Verstraeten, and Benjamin Schrauwen. An Uncued Brain-Computer Interface Using Reservoir Computing. In *Workshop: machine learning for assistive technologies, Proceedings*, page 8. Ghent University, Department of Electronics and information systems, 2010.
- [6] Stephan Waldert, Hubert Preissl, Evariste Demandt, Christoph Braun, Niels Birbaumer, Ad Aertsen, and Carsten Mehring. Hand movement direction decoded from MEG and EEG. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 28(4):1000–8, January 2008.

# BCI-FES hand therapy for patients with sub-acute tetraplegia

Osuagwu, A. Bethel<sup>1</sup>, Vuckovic, Aleksandra<sup>1,2</sup>, Fraser, Matthew<sup>3</sup>, Wallace, Leslie<sup>3</sup>  
and Allan, B. David<sup>2,3</sup>

<sup>1</sup> Centre for rehabilitation engineering, School of Engineering, University of Glasgow, Glasgow, UK  
b.osuagwu.1@research.gla.ac.uk

<sup>2</sup> Scottish Centre for Innovation in Spinal Cord Injury, Glasgow, UK

<sup>3</sup> Queen Elizabeth National Spinal Injury Unit, Southern General Hospital, Glasgow, UK

## Abstract

Brain computer interfaces could be used to control functional electrical stimulation to achieve better results in neurorehabilitation. Four sub-acute tetraplegic patients received brain computer interface controlled functional electrical stimulation (BCI-FES) for rehabilitation of hand functions for a period of twenty days. All the patients showed improvement in the manual muscle test grading after using the BCI-FES. There were also changes in ERD activities that suggested improved hand functions. The results suggest that BCI-FES might be useful for rehabilitation of functions after spinal cord injury.

## 1 Introduction

Brain computer interface (BCI) is a promising new technology with possible application in neurorehabilitation after spinal cord injury [4]. Motor imagery (MI) based BCI coupled with functional electrical stimulation (FES) enables the simultaneous activation of the motor cortices and the muscles they control [4]. When using an MI-based BCI-FES, the subject activates the motor cortex using MI of limb movement. For rehabilitation purposes the MI can be replaced with attempted movement (AM). The BCI system detects the motor cortex activation and activates the FES attached to the muscles of the limb the subject is imaging to move. In this way the afferent and the efferent pathways of the nervous system are simultaneously activated. This simultaneous activation encourages Hebbian type learning [2] which could be beneficial in functional rehabilitation after spinal cord injury (SCI).

FES is already in use in several SCI rehabilitation units but there is currently not enough clinical evidence to support the use of BCI-FES for rehabilitation.

The aim of this study is to assess outcomes in sub-acute tetraplegic patients using MI-based BCI-FES for functional hand rehabilitation.

## 2 Methods

### 2.1 Subjects

Four sub-acute tetraplegic patients ps1, ps2, ps3 and ps4 (male, 70, 25, 35, and 20 years old respectively) participated in this study after giving their informed consent. They all had incomplete cervical injuries respectively at neurological levels, C6, C4, C6, and C5 with ASIA classification levels C, B, B and C. As the result of the injury the patients lost most of their hand functions with minimal elbow and shoulder movement available on both hands. They were recruited into this after about eight weeks following their injuries.

## 2.2 Initial and final assessment

At the start and at the end of the study, an occupational therapist assessed the hand functions of the patients using the manual muscle testing (MMT) [1] grading system. The grades ranges from 0 to 5 for no observable or palpable muscle contraction and normal muscle function respectively. The muscles tested included the extensor digitorum (ED), extensor carpi radialis (ECR), extensor pollicis longus (EPL), flexor carpi radialis (FCR), flexor digitorum profundus (FDP), and intrinsic hand muscles (I).

The initial and final assessment also involved measurement of multichannel EEG of the patients with 48 channels while they were performing AM of the left and right hand. EEG was recorded using the g.USBamp by (GTEC, Austria) at a sample rate of 256 Hz with filters between 0.5 and 60 Hz and a notch at 50 Hz. Event related desynchronisation/synchronisation (ERD/ERS) [3] maps used to quantify movement related cortical activities during MI/AM was compared between 'before' and 'after' the study.

## 2.3 BCI-FES therapy sessions

The patients were scheduled to attend therapy sessions between Monday and Friday until they completed twenty sessions. This was in addition to their regular physiotherapy. Each session lasted for about an hour.

**BCI-FES:** At the beginning of each therapy session, a BCI classifier was computed to enable the control of the BCI-FES system. To obtain the classifier, a patient was instructed to attempt closing and opening of the left and the right hand (20 trials for each hand). During this task EEG was recorded from three pairs of bipolar electrodes to derive C3 (Fc3-Cp3), Cz (Fcz-Cpz) and C4 (Fc4-Cp4). The input signal from the amplifier was bandpass filtered (5<sup>th</sup> order Butterworth) online between 0.5 to 30 Hz. The time domain parameter (TDP) [6] of the filtered data was computed. After extensive test on healthy people, we preferred the TDP to the bandpower feature because it gave better classification accuracy for small number of trials and eliminated the need to select a specific frequency band. The TDP was computed in a similar way as that described by Vidaurre and colleagues [6]. The feature was used to compute linear discriminant analysis classifiers to discriminate the left and right AM from resting state. This gave a 2-class system comprising a hand's AM against the resting state. The classifiers were saved for online use. These steps were all integrated into graphical user interface under MATLAB and Simulink. This process and setting up the EEG on the patient typically took 15 minutes.

In the online classification, TDP was estimated using Simulink's difference blocks to compute derivatives. Each sample of the signal in the feature space was binary classified either as an 'Active' or 'Relaxed' state using the classifier computed offline. The 'Active' state occurred when the subject attempted opening and closing of a hand while the 'Relaxed' state corresponded to a resting period. The classifier output (i.e 'Active' or 'Relaxed') was then buffered for a variable length of time up to a maximum of 3 s or 768 samples. An 'Active' state was detected when the buffer was filled with approximately 95% of 'Active' state. The size of the buffer, usually 1-2 s long, was determined for the patient and optimized to significantly reduce false positives which was reported by the patient. When the patient found it difficult to control the system or reported false positives, the classifier was updated online. The classifier was updated using fixed rate supervised mean and covariance adaptation methods described elsewhere [5].

The FES device (Rehastim by Hasomed, Magdeburg) was used to assist the patients to perform grasp functions. Electrodes were attached to sequentially stimulate the extensor digitorum, extensor pollicis longus, flexor digitorum superficialis and the flexor pollicis brevis

muscles. The frequency of stimulation was set to 26 Hz and the pulsewidth was set to 200  $\mu$ s. The FES current typically 15-35 mA was chosen for each subject to extend and flex the hands or cause visible muscle contraction without discomfort.

The patient was instructed to attempt closing and opening of the hand with the electrodes to activate the FES. Note that this was done on one hand at a time. The FES was activated for 10 s to repeatedly assist in the opening and closing of the hand when the active state was detected. A therapy session lasted for 60 minutes with 30-40 trials on each hand separately.

### 3 Results and discussions

Patients ps1, ps2, and ps4 were able to control the BCI-FES system from the first session because they had ERD of their EEG during AM. Patient ps3 had to be trained for four sessions to produce ERD of his EEG before the patient could have a good control of the system. The patients, ps1, ps2, ps3 and ps4 had a false positive of 10% ,0%,10%,0% respectively at the initial sessions but this decreased to zero at the end of the therapy.

The MMT grading for the right hand of the patients before and after the therapy is shown in Table 1. All the patients had improvement in the MMT with patient ps4 showing the most improvement. Patient ps4 unlike the other patients completed the study in the shortest time and therefore may have received a better effect. Patient ps1 showed the most improvements in the FCR which moved up by one grade. Patients ps2 and ps3 who had more severe injuries according the ASIA scale had improvements in a smaller number of muscles although patient ps3 achieved the maximum MMT grading for ED and ECR.

Patients	Right hand muscles					
	ED	ECR	EPL	FCR	FDP	I
ps1 (b)	4-	4-	3-	3	4-	3
ps1 (a)	4+	4+	3	4	4	3
ps2 (b)	1	1+	0	0	0	0
ps2 (a)	3	3+	0	0	0	0
ps3 (b)	4	4	0	0	0	0
ps3 (a)	5	5	0	0	1	0
ps4 (b)	0	0	0	0	0	0
ps4 (a)	4-	4+	3+	4	4	3+

Table 1: MMT gradings before and after the therapy for the right hand. a, after; b, before; -/+ , -0.5/+0.5.

The ERD/ERS maps computed for ‘before’ and ‘after’ the therapy are shown on Figure 1. The data presented is of the right hand AM. ERD is shown with negative values while ERS is shown with positive values on the figures. The dashed line on the ERD/ERS map show when the patients were given a cue to begin AM.

Visual inspection of the ERD/ERS maps show that there are more ERDs after therapy for all patients except for patient ps2. The improvement in the ERDs are within the characteristic movement related ERD bands typically within 8-12 Hz and 16-24 Hz. These ERD changes are in support of the MMT grades.

For patient ps2, the ERD seen before the therapy was not present after the therapy. This patient had the highest injury level (C4) classified as ASIA B. With such a high level and

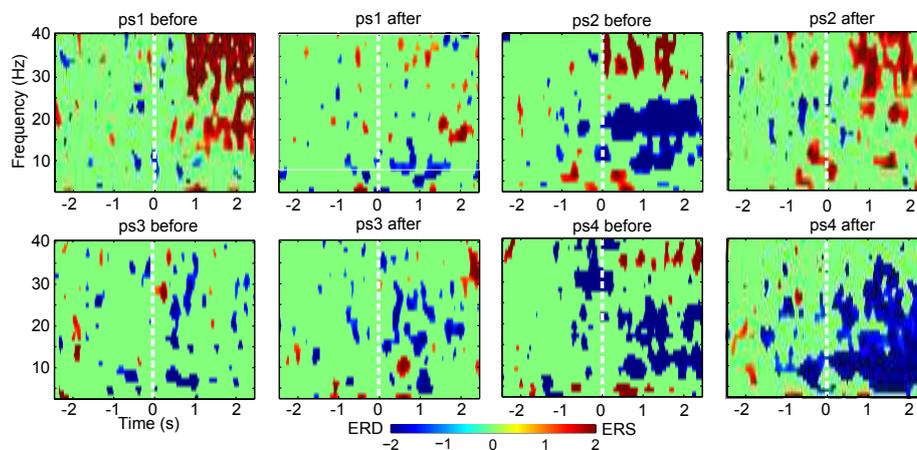


Figure 1: ERD/ERS maps (over channel C3) of right hand AM for before and after the therapy. Movement cue was given at  $t=0$  ms. (Generated with EEGLAB, <http://scn.ucsd.edu/eeglab>).

severe injury, it might be harder for the patient to recover functions. This is reflected in the MMT grades before and after the therapy in which this patient was graded the least. The disappearance of the ERD might reflect cortical reorganisation.

## 4 Conclusions

The MMT grading system showed improvement for all our patients using BCI-FES. This improvement and the increase in ERD activities in three patients were promising results which suggested that BCI-FES might be used for functional rehabilitation in SCI patients. However the improvements may have been a result of natural recovery, conventional therapy and the use of FES. We will therefore require control patients who will receive only FES therapy. The study is ongoing and we are recruiting more patients in order to find the statistical significance of using BCI-FES compared to therapy using only FES for rehabilitation in SCI patients.

## References

- [1] S. C. Cuthbert and G. J. Goodheart. On the reliability and validity of manual muscle testing: a literature review. *Chiropractic & Manual Therapies*, 15(1):4, 2007.
- [2] D.O. Hebb. *The Organization of Behavior: A Neuropsychological Theory*. New York: John Wiley & Sons, 1966.
- [3] G. Pfurtscheller and A. Aranibar. Event-related cortical desynchronization detected by power measurements of scalp EEG. *Electroencephalogr Clin Neurophysiol*, 42(6):817–826, 1977.
- [4] G. Pfurtscheller, G. R. Müller, J. Pfurtscheller, H. J. Gerner, and R. Rupp. 'Thought' - control of functional electrical stimulation to restore hand grasp in a patient with tetraplegia. *Neuroscience Letters*, 351(1):33 – 36, 2003.
- [5] A. Schlögl, C. Vidaurre, and K. Müller. *Brain-Computer Interfaces: Revolutionizing Human-Computer Interaction*, chapter Adaptive Methods in BCI Research An Introductory Tutorial, pages 131–356. Springer, 2010.
- [6] C. Vidaurre, N. Krämer, B. Blankertz, and A. Schlögl. Time domain parameters as a feature for EEG-based brain-computer interfaces. *Neural Networks*, 22(9):1313–1319, 2009.

# Global EEG synchronization as an indicator of emotional arousal and its application for tracking emotional changes during video watching

Jun-Hak Lee<sup>1</sup>, Jeong-Hwan Lim<sup>1</sup>, Chang-Hee Han<sup>1</sup>, Yong-Wook Kim<sup>1</sup>,  
and Chang-Hwan Im<sup>1\*</sup>

<sup>1</sup>Department of Biomedical Engineering, Hanyang University, Seoul, South Korea  
ich@hanyang.ac.kr\*

## Abstract

In the present study, we investigated whether global EEG synchronization can be a new index to track emotional changes during continuous EEG recording. Global synchronization index (GFS), a measure known to reflect human cognitive process, was evaluated to quantify the global synchronization of multichannel EEG data recorded from a group of participants (n=20) while they were watching two short video clips. The two video clips were about 5-min long each and designed to respectively evoke happy and fearful emotions. Other participants (n=21) were asked to select two most impressive scenes in each clip, and the questionnaire results were then compared to the grand averaged GFS waveforms. Our results showed that beta-band GFS value was decreased when people experienced high emotional arousal regardless of the types of emotional stimuli, suggesting that the GFS measure might be used as a new index to track emotional changes during video watching and would be potentially used for evaluating movies, TV commercials, and other broadcast products.

## 1 Introduction

Understanding human emotional process has been regarded as an important research topic in neuroscience as emotion plays a key role in communication or interaction between humans. Many different neuroimaging modalities have been used to study the neural substrates of emotion, such as electroencephalography (EEG), magnetoencephalogram (MEG), and functional magnetic resonance imaging (fMRI) (Peyk et al., 2008). Among these, EEG has been considered the most suitable tool to study the temporal dynamics of emotion thanks to its high temporal resolution and reasonable cost (Millan et al., 2008). In particular, decoding individual emotional states from EEG is attracting increased attention because of the recent popularization of low-cost, wearable EEG systems and their potential applications in affective brain-computer interfaces (BCIs).

Two representative methods have been most widely used to evaluate or recognize individual emotional states, which are frontal alpha-asymmetry and event-related potentials (ERPs) elicited by emotional stimuli (Degabriele et al., 2011). Despite a number of studies based on these methods, only a few studies have attempted to continuously track or monitor the emotional changes using EEG data recorded while continuous and complex emotional stimuli (e.g., movies and video clips) are presented.

In this study, we tried to search for a new index that can effectively estimate temporal changes in emotional states of a group of individuals while they were watching video clips designed to evoke different types of emotions, which can be potentially applied to the evaluation of various cultural

contents including movies, TV commercials, and music videos. Among various possible indices, we tested the feasibility of EEG synchronization as an index of emotional states because some previous studies reported that functional connectivity is generally increased during cognitive processing and decreased during emotional processing especially in beta & gamma frequency bands (Harding et al., 2012, Northoff et al., 2004). Beta-band global field synchronization (GFS), a well-known measure shown strong correlation with human cognitive process (Koenig et al., 2001, Lee et al., 2010), was evaluated to quantify the global synchronization of multichannel EEG data recorded from a group of participants while they were watching two video clips. Our study hypothesis was that decreased GFS values might be observed when stimuli eliciting high emotional arousal are presented because dominant emotional processing would act as a distractor of cognitive processing in the brain. To verify our hypothesis, results of simple questionnaires inquiring two scenes that were most impressive and memorable in each clip and the grand-averaged GFS waveforms were compared.

## 2 Methods

### 2.1 Experimental conditions and paradigm

EEG data were recorded using a multichannel EEG system (ActiveTwo, BioSemi) with 22 active EEG channels (Cz, C3, C4, T7, T8, Pz, P3, P4, P7, P8, Fp1, Fp2, Fz, F3, F4, F7, F8, AFz, AF7, AF8, O1, and O2) and two EOG channels (VEOG and HEOG). A 17-inch LCD monitor was used for the presentation of visual stimuli. The subjects were seated in a comfortable armchair placed in front of the monitor. The distance between the monitor and subjects was set to 70 cm. Two short video clips were presented to the subjects. One clip used to evoke happy emotion was a short video clip titled 'Isaac's Live Lip-Dub Proposal', which was watched by more than 25 million people in YouTube ([http://youtu.be/5\\_v7QrIW0zY](http://youtu.be/5_v7QrIW0zY)) (referred to as a *positive clip*). In the video of 5:13-min long, a man named Isaac surprised his girlfriend with the world's first live lip-dub marriage proposal. The other clip used to elicit negative emotion was edited from a famous movie 'The Grudge'. In the video of 4:36-min long, sudden appearances of ghosts made subjects feel fearful (referred to as a *negative clip*).

Twenty-five healthy subjects initially participated in our experiment. The *positive clip* and *negative clip* were presented to the participants one after another, between which a short break time of about two minutes was given. To investigate the general feeling about each video, another group of participants (n=21) was recruited and asked to select two most impressive scenes in each clip. They watched each video clip twice without recording EEG, and then asked to mark two time points during the second play of each clip.

### 2.2 Data Analysis

Principal components analysis (PCA) was used to remove EOG artifacts from EEG data. The preprocessed data were then filtered using a band pass filter with cutoff frequencies of 1 and 55 Hz. EEG data from five participants were excluded from the analysis because these data were severely contaminated by noises and artifacts. The data were then segmented into 2-s epochs with a 50% overlap to continuously evaluate GFS values along time. The GFS, a method to measure the overall functional connectivity of the brain, has been generally used to investigate cognitive decline in patients with psychiatric disorders (Koenig et al., 2001, Lee et al., 2010). To evaluate the GFS values, EEG signals recorded from different scalp locations are first transformed into frequency domain using the Fast Fourier Transform (FFT), and then the FFT-transformed signals at each frequency are mapped onto a complex plane. The GFS value of a frequency is defined as the normalized difference between two eigenvalues representing the point distribution in the 2-D complex plane (Koenig et al., 2001). According to Lee et al.'s study (2010), the average GFS value in beta band (14~25Hz) showed

strong correlation with Mini-Mental States Examination (MMSE) scores that reflected cognitive process in the brain. The GFS time-series of each participant was filtered using a ten-second moving-average filter, and then grand averaged across all participants.

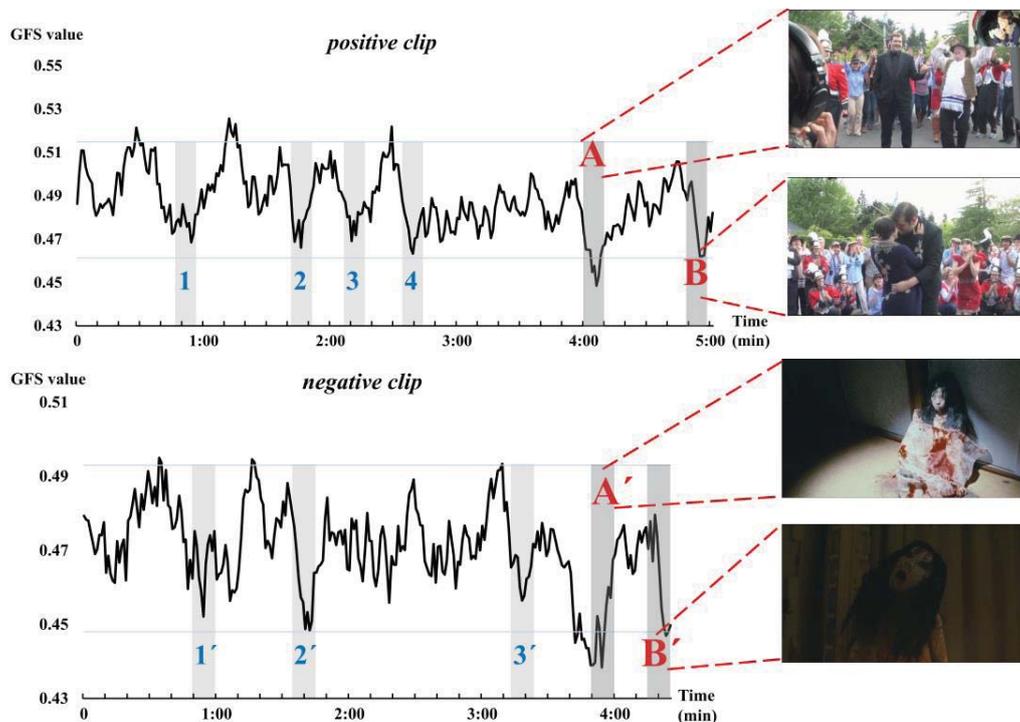


Figure 1: Grand averaged GFS waveforms

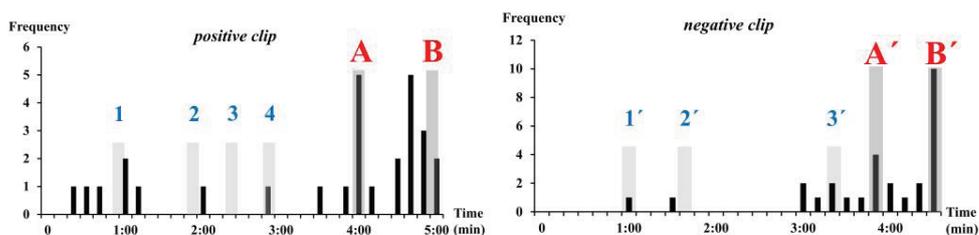


Figure 2: The questionnaire results on two most impressive scenes in each clip

### 3 Results

Figure 1 shows the grand-averaged GFS waveforms of each video clip and Figure 2 shows the results of questionnaire (two most impressive scenes in each clip). Horizontal lines in Figure 1

represent 1.96 standard deviation values ( $p < 0.05$ ). Interestingly, two time periods in each video clip at which the GFS values dropped below the lower horizontal line (denoted by A, B, A', and B') matched fairly well with the time periods most frequently selected in the questionnaire. In the time periods A and B, the woman was finally proposed by the man, and in A' and B', sudden advent of the ghost made the subjects feel most fearful. In Figure 1, there are some other time points showing sudden drops of GFS values to around the lower horizontal line (denoted by 1, 2, 3, 4, 1', 2', and 3', which were chosen objectively), at which time some reported high emotional arousal, as shown in Figure 2. The average GFS values for all emotional events marked in Figure 1 were compared with the GFS values averaged over all time points not marked as emotional events using paired t-test. The statistical analysis results showed significant differences between two conditions (positive clip:  $p = 0.0068$ ; negative clip:  $p = 0.0059$ ).

## 4 Conclusion

The purpose of the study was to continuously track the changes in emotional arousal using a global EEG synchronization measure. As expected in our study hypothesis, decreased GFS values were observed when stimuli eliciting high emotional arousal are presented. It is noteworthy that the decreased GFS values were commonly observed regardless of the types of stimulus valence, which supports our main hypothesis that emotional processing might act as a distractor of cognitive processing and thus the GFS value, an indicator of cognitive processing, would decrease during emotional processing. The results of questionnaires could be elucidated fairly well with the grand-averaged GFS waveforms, suggesting that our study design would be potentially applied to practical applications to evaluate various cultural contents or broadcasting products. Further studies need to be conducted in future studies in order to generalize our hypothesis using more experiments and investigate individual variability in the GFS waveforms.

## References

- Peck P., Schupp H. T., Elbert T., Junghöfer M. (2008). *Emotion processing in the visual brain: a MEG analysis*. Brain Topogr
- Millan J. D., Ferrez P. W., Galan F., Lew E., & Chavarriaga R. (2008). *Non-invasive brain-machine interaction*. International Journal of Pattern Recognition and Artificial Intelligence, 22, 959-972.
- Degabriele R., Lagopoulos J., Malhi G. (2011). *Neural correlates of emotional face processing in bipolar disorder: An event-related potential study*. Journal of Affective Disorders.
- Harding I.H., Solowij N., Harrison B. J., Takagi M., Lorenzetti V., Lubman D. I., Seal M. L., Pantelis C., Yücel M. (2012). *Functional connectivity in Brain Networks Underlying Cognitive Control in Chronic Cannabis Users*. Neuropsychopharmacology.
- Northoff G., Heinzel A., Bermpohl F., Niese R., Pfennig A., Pascual-Leone A., Schlaug G. (2004). *Reciprocal Modulation and Attenuation in the Prefrontal Cortex: An fMRI Study on Emotional-Cognitive Interaction*. Human Brain Mapping
- Koenig T., Lehmann D., Saito N., Kuginuki T., Kinoshita T., Koukkou M. (2001). *Decreased functional connectivity of EEG theta-frequency activity in first-episode, neuroleptic-naïve patients with schizophrenia: preliminary results*. Schizophrenia Research
- Lee S-H., Park Y-M., Kim D-W., Im C-H. (2010). *Global synchronization index as a biological correlate of cognitive decline in Alzheimer's disease*. Neuroscience Research.

# The Influence of Motivation when the Task gets Harder: Visual versus Auditory P300 Brain-Computer Interface Performance

Sonja C. Kleih<sup>1</sup>, Femke Nijboer<sup>2</sup>, Sebastian Halder<sup>1</sup>, Adrian Furdea<sup>3</sup>  
and Andrea Kübler<sup>1</sup>

<sup>1</sup> University of Würzburg, Würzburg, Germany, <sup>2</sup> University of Twente, Enschede, Netherlands,

<sup>3</sup> University of Tübingen, Tübingen, Germany

sonja.kleih@uni-wuerzburg.de

## Abstract

This study investigated the influence of motivation on Brain-Computer Interface (BCI) performance with increased difficulty of the BCI task by changing the sensory modality of the input stimulation from visual to auditory. We found increased brain activation in response to the stimulation during the auditory task when participants were highly motivated as compared to being moderately motivated.

## 1 Introduction

Motivation was shown to have an influence on Brain-Computer Interface performance (e.g. Kleih et al., 2010; Hammer et al., 2012). When using the event-related potential P300 as BCI input signal, it was hypothesized that motivation might increase attentional resource allocation and thereby brain activation and classification accuracy (Kleih et al., 2010). This hypothesis also suggests motivation to be more influential when the task gets more demanding. The allocation of additional attentional resources to the stimulation alone might increase brain activation and thereby balance the effect of less detectable brain signals due to increased complexity of the task. However, this assumption has never been systematically investigated. In the here presented study we systematically increased participants' motivation by monetary reward and compared performance in the visual and auditory sensory input channel with the auditory modality representing the more difficult task. We hypothesized an increase of motivation by monetary reward (H1), an increase of brain activation (P300 amplitude) as a function of motivation and this effect to be more distinct in the auditory as compared to the visual modality (H2) and higher performance when being motivated (H3).

## 2 Methods

### 2.1 Participants

We recruited N=14 students with an average age of 27.00 years ( $SD = 7.30$ , range 19-38). Three males participated in the study and none of the participants reported a history of neurological or psychiatric disorder. Participants were paid 8 € per hour and all were naïve to BCI training prior to participation. Participants gave informed consent to the study, which had been reviewed and approved by the Ethical Review Board of the Medical Faculty, University of Tübingen.

## 2.2 Procedure and Stimulation Parameters

After calibration, participants had to spell the words ‘BRAIN’ and ‘POWER’ with an auditory and a visual P300 BCI under three different reward conditions. In condition “0ct” they received no reward for correctly spelled letters, in condition “5ct” they received 0.05 € for every correctly spelled letter (maximum win: 0.50 €) and in condition “100ct” they received 1 € for each correct letter (maximum win: 10.00 €). We used these rewards as we were interested whether an additional reward would already increase motivation (0.05 €) or whether motivation can only be increased by a relatively high and salient reward (1€). In the visual condition, ten sequences were used, thus each individual character was flashed 20 times. Each stimulus, row or column, flashed for 312.5 ms. Afterwards, the screen was static for another 312.5 ms. Thus, the presentation of the rows and of the columns both had a duration of 31.5 s, which means that one character could be selected every 62.5 s. An interval of 5.625 s was provided before each sequence such that participants could locate the next character in the matrix. For the auditory modality, the auditory speller introduced by Furdea and colleagues (2009) was used. In that auditory version of the P300-speller, letters were represented by a combination of two numbers which indicated its matrix location. Each character therefore could be defined by the coordinates of these number codes. Auditory stimuli consisted of computer-generated numbers spoken by a male voice. To select a particular target character, the participant had to attend to two target stimuli representing the coordinates of the character in the matrix while within one trial only either the number representing the row or the column was presented. Participants viewed the same matrix as in the visual speller to support finding the coordinates of the character-to-select.

All conditions (reward and modality) were counterbalanced. Participants were told before every condition how much money they could earn for every letter correctly spelled and the maximum win they could obtain. The experimenter scored every accurate letter, but the participant did not receive feedback during the run to prevent any effect of success on motivation. Immediately after every condition, motivation was assessed using a visual analogue scale (VAS) ranging from 0 (not motivated at all) to 10 (very high motivation). Only after finishing all six conditions, participants received feedback of their performance and the amount of money they had earned.

## 2.3 Data acquisition and analysis

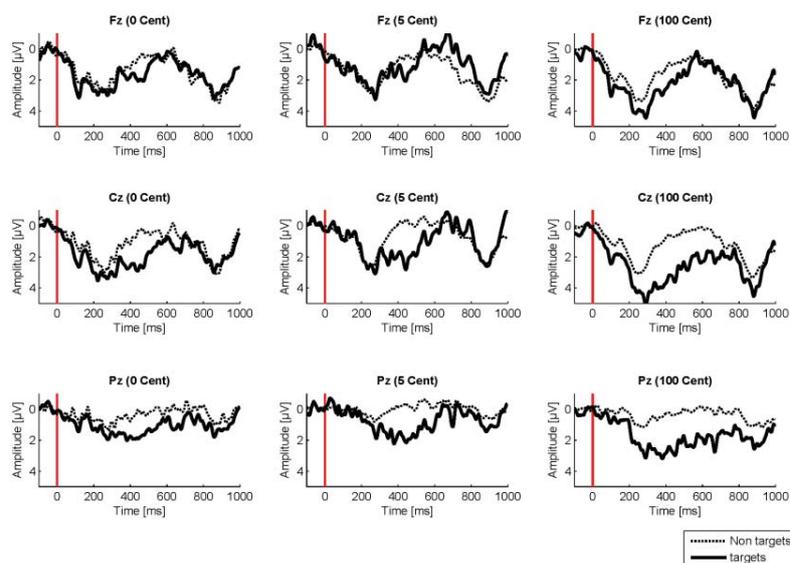
Stimulus presentation and data collection were controlled by BCI2000 software (Schalk et al., 2004). The EEG was recorded using a tin electrode cap (Electro-Cap International, Inc., Eaton, OH) with 16 channels. Electrode locations were F3, Fz, F4, T7, C3, Cz, C4, T8, CP3, CP4, P3, Pz, P4, PO7, Oz, PO8 based on the modified 10-20 system. Each electrode was referenced to the right and grounded to the left mastoid. The EEG was amplified using a 16-channel g.USBamp (Guger Technologies, Austria), sampled at 256 Hz, and bandpass filtered between 0.01 – 30 Hz. Fifty Hz noise was filtered using a notch filter implemented in the BCI2000. Data processing, storage and stimulus presentation was controlled with DELL laptop (Intel Core 2 Duo T5550 1.83 GHz, GB DDR2, Windows XP). Stepwise linear discriminant analysis (SWLDA) was used for classifier weights generation after the initial calibration sessions. The EEG data were corrected for artifacts and baseline (-100 to 0 ms) using Brain Vision Analyzer 2<sup>®</sup> (Brain Products Germany). The P300 was defined as the maximum positive peak occurring between 200 and 700 ms after stimulus onset and was chosen by semiautomatic global peak detection. For statistical analysis IBM SPSS<sup>®</sup> 20 was used. The level of significance was set to  $\alpha = .05$ .

## 3 Results

The first hypothesis that monetary reward would increase participants’ motivation was tested with repeated measures analysis of variance (ANOVA) with the factors *modality* (visual and auditory) and

reward (0ct, 5ct, 100ct) and the dependent variable VAS motivation. We found a main effect for *reward* ( $F_{(2,28)} = 4.94; p < .05$ ). Within-subjects contrasts revealed a significant motivation increase between 5ct and 100ct ( $F_{(1,14)} = 5.34, p < .05$ ). Neither a main effect for modality nor an interaction effect was found.

H2 stated that increased motivation would lead to increased brain activation as reflected in the P300 amplitude. When calculating repeated measures ANOVA with the factors *modality* and *reward* and the dependent variable *P300 amplitude*, we found P300 amplitudes in the visual modality to be significantly higher than the P300 amplitudes in the auditory modality ( $F_{(2,28)} = 72.48, p < 0.01$ ). However, there was no effect of monetary reward on the P300 amplitude ( $F_{(2,28)} = 0.09; p = 0.92$ ) nor an interaction ( $F_{(2,28)} = 1.48; p = 0.25$ ). When comparing brain activation in the auditory modality (see figure 1) between the 5ct and the 100ct conditions, it is apparent that there is higher activation in the 100ct reward condition even though no clear ERP peak is detectable. Therefore, areas under the curve (AUC) in the different auditory conditions were compared. The measure of AUC represents the integral under the curve in a defined time interval. With AUC the distance of each data point from zero is ignored while calculating the integral with reference to the first value of the interval (Pruessner et al., 2003). We compared AUC values with repeated measures ANOVA for two target *electrodes* (Cz, Pz, see figure 1) and the *reward* conditions (0, 5, 100 Cent). A trend towards a main effect of *reward* condition was found ( $F_{(2,22)} = 3.21, p = .06$ ). Within-subjects contrasts revealed a significant difference between reward conditions 5ct and 100ct ( $F_{(1,11)} = 4.99, p < .05$ ) but not between 0ct and 5ct conditions ( $F_{(1,11)} = .09, p = .77$ ). No main effect of *electrode* ( $F_{(1,11)} = .21, p = .65$ ) was found.



**Figure 1:** Brain activation 1 s after stimulus onset in the auditory modality at Fz, Cz and Pz for the motivation conditions 0ct, 5ct and 100 ct reward.

Concerning performance (H3), participants were more successful in the visual ( $M = 95.12\%$  accuracy,  $SD = 9.60$ ) as compared to the auditory task ( $M = 44.88\%$  accuracy,  $SD = 30.66, F_{(1,14)} = 65.91, p < .001$ ). Motivation did not increase accuracy ( $F_{(2,28)} = 2.03, p = .15$ ) and no interaction effect was found between modality and monetary reward ( $F_{(2,28)} = .87, p = .43$ ).

## 4 Discussion

We showed that motivation to perform a BCI task could be successfully manipulated with monetary reward. Compared to the study from Kleih and colleagues (2010), we used higher reward for a correct selection (0.50 € in Kleih et al., 2010 versus 1 € in this study) and the intervals between the monetary rewards were unequal (0, 5, 100 € Cent). We assume this caused a more salient perception of the reward value as compared to the study by Kleih and colleagues (2010). Our hypothesis of motivation to increase and thereby balance potentially reduced brain activation in more difficult tasks was supported by the here presented data. We found an increase of brain activation as measured with the area under the curve in the auditory task in the 5ct and the 100ct motivation condition as compared to the 0ct motivation condition. This result suggests the influence of psychological variables, such as motivation, to be higher when tasks become more challenging while at the same time motivation seems to increase only with the highest possible reward condition. However, the results should be interpreted with caution as spelling accuracy was not influenced by motivation and in the visual version of the P300-speller it was almost twice as high as compared to the auditory. Furdea and colleagues (2009) also found a clear advantage of the visual over the auditory modality. However, in their study no differences in P300 amplitudes were found while in our study we found significantly lower P300 amplitudes in the auditory condition compared to the visual one. This was probably caused by our decision to use less sequences than Furdea and colleagues (2009) as we wanted to avoid a ceiling effect in performances especially in the visual P300 speller (Furdea et al., 2009; Kleih et al., 2010). In conclusion, we found that motivation can be manipulated by monetary reward in a BCI task, that motivation does influence brain activation when the task is demanding and that this brain activation increase does not affect BCI performance in terms of accuracy, i.e. correctly selected letters.

## Acknowledgements

This work is supported by the European ICT Program Project FP7- 224631 (TOBI: Tools for Brain-Computer Interaction). This manuscript only reflects the authors' views and funding agencies are not liable for any use that may be made of the information contained herein.

## References

- Furdea, A., Halder, S., Krusienski, D. J., Bross, D., Nijboer, F., Birbaumer, N., & Kübler, A. (2009). An auditory oddball (P300) spelling system for brain-computer interfaces. *Psychophysiology*, *46*(3), 617-625. doi: PSYP783 [pii] 10.1111/j.1469-8986.2008.00783.x
- Hammer, E. M., Halder, S., Blankertz, B., Sannelli, C., Dickhaus, T., Kleih, S., . . . Kübler, A. (2012). Psychological predictors of SMR-BCI performance. *Biol Psychol*, *89*(1), 80-86. doi: 10.1016/j.biopsycho.2011.09.006
- Kleih, S. C., Nijboer, F., Halder, S., & Kübler, A. (2010). Motivation modulates the P300 amplitude during brain-computer interface use. *Clin Neurophysiol*, *121*(7), 1023-1031. doi: 10.1016/j.clinph.2010.01.034
- Nijboer, F., Furdea, A., Gunst, I., Mellinger, J., McFarland, D. J., Birbaumer, N., & Kübler, A. (2008). An auditory brain-computer interface (BCI). *J Neurosci Methods*, *167*(1), 43-50. doi: 10.1016/j.jneumeth.2007.02.009
- Schalk, G., McFarland, D. J., Hinterberger, T., Birbaumer, N., & Wolpaw, J. R. (2004). BCI2000: a general-purpose brain-computer interface (BCI) system. *IEEE Trans Biomed Eng*, *51*(6), 1034-1043. doi: 10.1109/TBME.2004.827072.

# The Comparison of Cortical Neural Activity Between Spatial and Non-spatial Attention in ECoG Study

Taekyung Kim<sup>1</sup>, Gayoung Park<sup>2</sup>, Hojong Lee<sup>1</sup>, Jinsick Park<sup>1</sup>,  
InYoung Kim<sup>1</sup>, Dongpyo Jang<sup>1</sup> and JoongKoo Kang<sup>2</sup>

<sup>1</sup> Department of Biomedical Engineering, Hanyang University, Seoul, Korea

<sup>2</sup> Department of Neurology, Asan Medical Center, Seoul, Korea  
dongpjang@gmail.com

## Abstract

In this study, we investigated the spatiotemporal dynamics during spatial and non-spatial attention tasks employing human electrocorticographic (ECoG) signals. Ten epileptic patients who underwent subdural electrodes insertion for epileptic surgery performed spatial and non-spatial attention tasks. Time-frequency and statistical analysis for ECoG signals resulted that both spatial and non-spatial attention commonly had event related desynchronization (ERD) over the low frequency (theta, alpha, beta) and event related synchronization (ERS) over the high frequency (gamma band). The difference between two paradigms have been found in the right parietal area showing ERD in superior parietal area during spatial task and ERD in inferior parietal area during non-spatial task.

## 1 Introduction

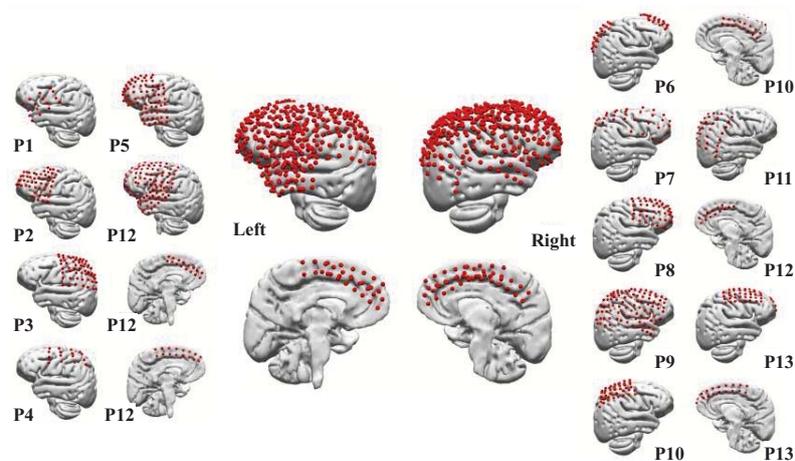
The posterior parietal cortex (PPC) has known to be a main part for cognitive processes including attention, spatial representation, working memory, and visuomotor control [1]. However, details about the neural mechanism underlying spatial and non-spatial attention are unclear because of the limitations of poor spatial and temporal resolution of the method used in previous studies [2]. The electrocorticographic (ECoG) recording may be the one of solutions because of its high temporal and spatial resolution and high signal fidelity.

The aim of this study was to investigate the neural mechanisms associated with two types of attention processing using ECoG signals from epileptic patients.

## 2 Materials and Methods

### 2.1 Participants

Thirteen epileptic patients (6 females, mean age  $32.6 \pm 10.1$  years) who had underwent an invasive study with intracranial electrodes for epileptic surgery participated in this experiment (Figure 1). After finishing a video-monitoring study to localize seizure foci with intracranial electrodes, the patients' seizures were controlled by anticonvulsant medications. The studies were conducted approximately 5-7 days after the electrode implantation therefore all patients were healthy and able to perform this



**Figure 1:** Location of electrodes. Left two columns and right two columns illustrate the location of electrodes for each participants. Four figures in the center show the electrode maps of whole participants overlaid on the standard brain surface.

experiment. This study was approved by Asan Medical Center, Seoul, Korea. Informed consent was obtained in accordance with the regulations of the Research Ethics Board of our institution.

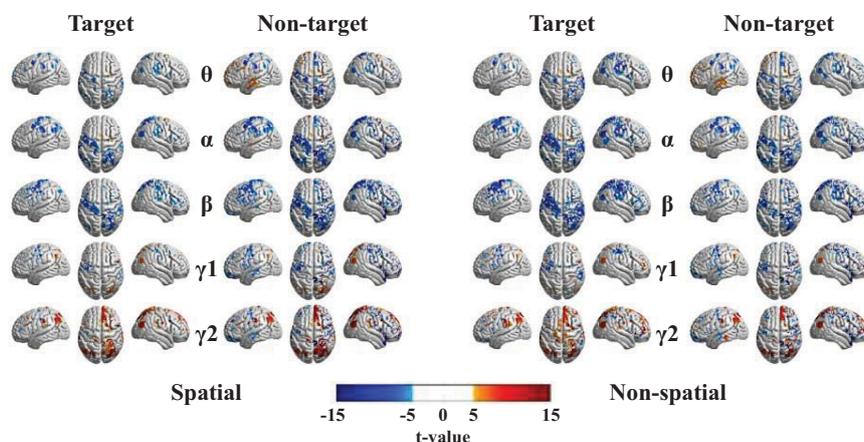
## 2.2 Experimental Task

The tasks were developed using VIZARD software (World Viz Inc.) and adopted from a Malhotra's paradigm [3] which combined spatial (location-based) and non-spatial (feature-based) visual attention. The patients were asked to respond as quickly as possible by pressing the space bar on the keyboard with their right hand when they saw predefined target stimuli according to attention type (spatial: two locations are target stimuli, non-spatial: two patterns are target stimuli). The visual stimuli consist of five different patterns that were presented sequentially in a random order at one of five positions along the vertical midline of the screen. Each stimulus was presented every 2 seconds, and remaining on the screen for 1 second. 500 stimuli were presented in total over a period of ~15min, with 200 target stimuli and 300 non-target stimuli shown during that time period.

## 2.3 ECoG Recordings and Processing

During the Experiment, ECoG signals were recorded continuously at a sampling frequency of 1000Hz and referenced by the Pz on the scalp with a Stellate Harmonie System (Stellate, Montreal, Canada) for 7 participants and with a Nihon Koden EEG system 1200K (Nihon Koden, Tokyo, Japan) for 3 participants. The electrodes were identified by co-registered image between pre-operative T1-MRI data and post-operative CT data using the FMRIB software library (Oxford, UK). The locations of each electrode were transformed into the Talairach coordinate system using Curry software (Compumedics, Charlotte, NC, USA).

The Matlab (MathWorks, Natick, MA, USA) and EEGLAB toolbox (Swartz Center, CA, USA) were used for processing ECoG data. ECoG signals were processed with a band-pass filter from 1 to 200Hz and were re-referenced with a common average reference [4]. An independent component analysis (ICA) was performed to remove artifact components such as eye movement and muscle movements. ECoG signals were then segmented into 2000msec epochs from 500msec before stimulus to 1500msec after stimulus. Noisy epochs were then removed.



**Figure 2:** *T*-values calculated from the paired *t*-test during 400-600ms period after cue onset are mapped on the MNI brain surface. If the power of an electrode is significantly decreased, the *t*-value becomes negative and the region is colored blue. If the power of an electrode is significantly increased, the *t*-value becomes positive and the region is colored red.

Spectral analysis was performed using the short-time Fourier transform (STFT), which is one of the most widely used signal analyzing methods. Briefly, STFT divides the ECoG signal into small overlapping time segments, analyzes each time segment and provides a time-frequency distribution. In each 2000msec epoch, we used a 500msec sliding Hanning window with overlapping 7.5msec and obtained 200 time bins. After that the time-frequency features are reduced into 10 time bins (200ms window with 100ms overlap) and 5 frequency bands (theta: 4-7 Hz, alpha: 8-13 Hz, beta: 13-30 Hz, low-gamma: 30-50 Hz, high-gamma: 50-150Hz).

To evaluate the spatiotemporal change during spatial and non-spatial attention tasks, we computed paired *t*-test using the features of trials for each channel. *P*-values of less than 0.01% were considered statistically significant and the color plot (Figure 2) shows that the *p*-values following the *t*-values are much smaller than this criterion. We compared the spatiotemporal course during each spatial and non-spatial attention task with respect to the baseline interval (from 500ms before the cue to cue onset). In this analysis, the *p*-values were used for defining the significance of the feature and *t*-values were for identifying whether the power of the feature was increasing or decreasing in comparison with the baseline.

Electrodes of all participants were projected on the template brain provided by the Montreal Neurological Institute (MNI). The fractional changes in 5 frequency oscillatory powers were assigned to each electrode and the signal power was displayed with color. A nearest-neighbor method was used where a cortical triangular mesh was colored strongest when the closest electrode and a linear faded to zero as the distance increased [5].

### 3 Results

Significant activations are displayed over cortices following the presentation of the cue onset during both spatial and non-spatial attention tasks (Figure 2). Event related desynchronization (ERD) is common across low frequency bands (theta, alpha, beta), and in gamma bands show event related synchronization (ERS). Oscillatory power changes were salient at 400 ~ 600msec post stimulus onset and maintained to 600 ~ 800msec. Especially spatial attention shows focused ERD in right superior

parietal lobe, whereas non-spatial attention shows stronger ERD in inferior parietal lobe in low frequency bands. Also, in high gamma band, the stronger ERS pattern is observed at the parietal area from the spatial attention and at the anterior frontal lobe from the non-spatial attention. According to the result from the non-target stimulus which the participants don't respond, these activation after 400 ~ 600msec post stimulus onset may not be affected by the button press or motor preparation.

## 4 Discussion

In this study, we analyzed ECoG signals to investigate oscillatory power changes in the processing of spatial and non-spatial attention in the bilateral cortex. The results demonstrated that both spatial and non-spatial attention seem like sharing common areas over the cortices for all frequency bands. In the right parietal area, we found some difference including spatial attention shows decreasing low frequency power in the superior parietal lobe while non-spatial attention appears to be desynchronized in the inferior parietal lobe. Although many studies about attention paradigm shows BOLD change using fMRI and most of them demonstrated that the BOLD signal has increased in the parietal region during the attention period [1], they were not able to show how the neural activity had been changed. To the best of our knowledge, this is the first study demonstrating the neural variation with spatiotemporal oscillatory power changes during spatial and non-spatial attention.

The current study drives a novel result that how the neural activity changes and which areas within cortices are the main role in both spatial and non-spatial attention tasks. A further study will let us know much more about the neural mechanisms of attention and it would be applicable to the passive mapping technique.

## Acknowledgement

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MISP) (No. 2013R1A2A2A04015925)

## References

- [1] J. Culham and N. Kanwisher, "Neuroimaging of cognitive functions in human parietal cortex," *Curr Opin Neurobiol*, vol. 11, pp. 157-163, 2001.
- [2] K. Jerbi, "Task-related gamma-band dynamics from an intracerebral perspective: review and implications for surface EEG and MEG," *Hum Brain Mapp*, vol. 30, pp. 1758-1771, 2009.
- [3] P. Malhotra, "Role of right posterior parietal cortex in maintaining attention to spatial locations over time," *Brain*, vol. 132, pp. 645-660, 2009.
- [4] A. Gunduz, "Neural Correlates of Visual-Spatial Attention in Electroencephalographic Signals in Humans," *Front Hum Neurosci*, vol. 5, no. 89, pp. 1-11, 2011.
- [5] D. Englot, "Impaired consciousness in temporal lobe seizures: role of cortical slow activity," *Brain*, vol. 133, pp. 3764-3777, 2010.

# Decoding of Picture Category and Presentation Duration - Preliminary Results of a Combined ECoG and MEG Study

T. Pfeiffer<sup>1,\*</sup>, C. N. Heinze<sup>1,\*</sup>, E. Gerber<sup>2</sup>, L. Y. Deouell<sup>2</sup>, J. Parvizi<sup>3</sup>,  
R. T. Knight<sup>4</sup> and G. Rose<sup>1</sup>

<sup>1</sup> Institute for Medical Engineering, Otto-von-Guericke-University Magdeburg, Magdeburg, Germany,

<sup>2</sup> Department of Psychology, Edmond and Lily Safra Center for Brain Sciences, The Hebrew University of Jerusalem, Jerusalem, Israel

<sup>3</sup> Department of Neurology and Neurological Sciences, Stanford University, Stanford, USA

<sup>4</sup> Helen Wills Neuroscience Institute, University of California, Berkeley, USA

\* These authors contributed equally to this work

tim.pfeiffer@ovgu.de

## Abstract

Limited recording times and different placement of electrocorticography (ECoG) grids over patients can make it necessary to adjust certain paradigms designated for ECoG data acquisition. The present study compares the accuracies of decoding picture categories and the corresponding presentation durations using data acquired by means of ECoG and magnetoencephalography (MEG). The results - decoding accuracies of up to 95 % in ECoG data and up to 85 % in MEG data (chance 20 %) as well as similarity in selected channels for both modalities - indicate that assumptions can be made from MEG data about the outcome of the same paradigm run on ECoG patients for any kind of grid position. Therefore, MEG provides a way to improve paradigms designated for ECoG studies and thus use the limited recording times more effectively.

## 1 Introduction

Over the last decade, studies more and more suggest ECoG to be the preferred method for acquiring high spatial and temporal resolution data to be used in brain-computer-interfaces (BCI) [1]. Besides its many advantages, acquisition of ECoG data turns out to be a difficult and demanding procedure. It is invasive and provides only limited recording times. Also, because purely based on medical necessity, the placement of the implanted electrode grids rarely covers all brain areas essential for BCI studies. Making best use of the limited patient collective and recording time should therefore be of strong priority. Former studies combined the high spatial resolution of magnetic resonance imaging (MRI) and high temporal resolution of electroencephalography (EEG) [2] to achieve similar signal characteristics to ECoG while being non-invasive. Here we investigate MEG data, which to a certain degree combine the advantages of both imaging modalities [3], and compare it to ECoG. The comparison is based on results of a picture category/duration decoding paradigm using state-of-the-art features and classification routines. Because of its immobility, MEG is not a solution for acquiring data in real-life BCIs. However, with the limited access to ECoG data it is helpful to test paradigms with other modalities and make assumptions about the possible outcomes in ECoG. By showing similarities and differences between MEG and ECoG findings, we conclude that MEG can provide just that.

## 2 Material and Methods

**Data** ECoG data were recorded from two volunteering patients (right handed males). Both received subdural electrode implants for pre-surgical planning of epilepsy treatment at Stanford, CA, USA. Electrode grids were solely placed based on clinical criteria and covered a variety of cortical areas including lateral occipital and medial temporal areas. All patients gave their informed consent before recordings started. The ECoG was recorded with a sampling frequency of 3051.7 Hz. Pre-processing steps included high pass filtering (cut-off: 0.5 Hz) as well as notch filtering around the power line frequency (60 Hz). The electric potentials of all electrodes were re-referenced to Common-Average-Reference. Afterwards, the time series were visually inspected for artifacts (e.g. epileptic activity) and epoched into trials representing the interval between -100 ms and 2000 ms with respect to picture onset times. MEG data were recorded with a bandwidth of 100 Hz (sampling frequency: 1017.25 Hz), using a whole-head 248-sensor BTi Magnes system (4D-Neuroimaging, San Diego, CA, USA). Data have been acquired from four healthy volunteers (age 25-28, one female, one left-handed). Like the ECoG, MEG data were epoched into trials containing the interval [-100 ms, 2000 ms].

**Task** Pictures from four different categories, namely objects, faces, watches and clothing, were presented to the subjects. Presentation durations were randomly chosen from five different time spans (300, 600, ..., 1500 ms). Intervals of varying duration (600, 750, ..., 1200 ms), in which only a focus is visible, intersect the pictures. Stimuli were shown either on a projection screen (MEG data) positioned 1 m away from the subject or on a notebook screen (ECoG data) within the patients reach. Subjects were requested to respond to the presentation of a piece of clothing with a button press. Target stimuli (i.e. clothing) accounted for approximately 10 % of the total trial count.

**Feature Extraction and Selection** Two different feature types have been extracted for this study, namely low-frequency time domain (LFTD) and high gamma (HG) features. LFTD features were obtained by low-pass filtering in Fourier domain with a cutoff frequency of 30 Hz (MEG: 10 Hz) and subsequent down sampling of the time series. For the computation of HG features, spectral power was measured using a sliding Hann-window approach (window length = 250 ms). For each window the square root of the power spectrum was computed by Fast Fourier Transform (FFT). The resulting coefficients were then averaged in the frequency band of 70-200 Hz. Feature selection is performed on training data only using an algorithm based on the Davies-Bouldin index. Full details on feature selection (as well as extraction) routines can be found in our previous study [4]. The algorithm was employed to select a set of the twenty (MEG: ten) most informative channels corresponding to the actual class separation problem (i.e. either discrimination of picture categories or presentation durations).

**Classification** Linear Support-Vector-Machines (SVM) were used in one-vs-one mode for both classification problems, picture category and duration. The influence of the constant  $C$  was analysed on a single data set for each modality. We found the influence of  $C$  to be minimal, as long as not chosen too small ( $C < 2^{-15}$ ). Consequently, we chose  $C = 2^{-5}$  (MEG:  $C = 2^{-10}$ ) across all datasets.

### 3 Results

In order to restrict information to any kind of visual processing within the brain, target stimuli (i.e. clothing) have been omitted for the analysis of decoding accuracies. Otherwise, activation of motor areas during subjects' button presses might influence the decoding.

Signal-to-noise ratio in the high gamma band of the MEG data was too low to be used for single trial analysis. Therefore, only LFTD feature accuracy is plotted for the MEG datasets.

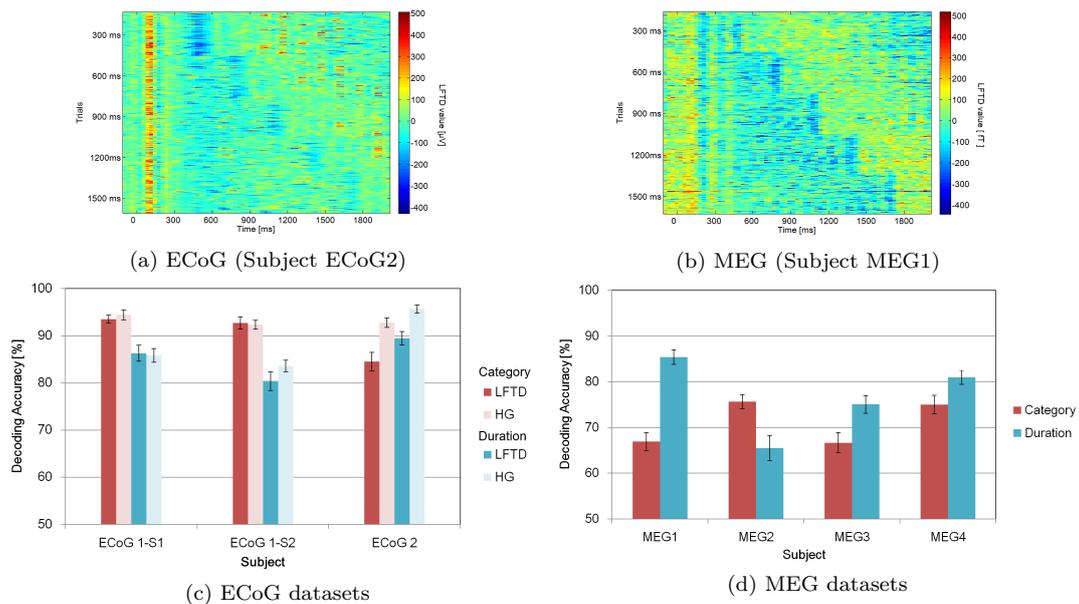


Figure 1: a), b) LFTD features for a dominantly selected channel (trials sorted by duration) c), d) Decoding accuracies (in %) for both classification tasks (i.e. category and duration) and both feature types (i.e. LFTD and HG features; ECoG only).

Decoding accuracies have been computed for all datasets by means of a 50-times-5-fold cross-validation procedure. The results for all ECoG and MEG datasets can be found in figure 1 c) and d). Figure 2 shows a comparison of the averaged decoding accuracies of both acquisition modalities. Chance levels for the two classification tasks are 20 % (duration, five classes) and 33 % (category, three classes) respectively.

We also found that the dominantly selected MEG channels were located in the same areas as the dominantly selected ECoG channels. This holds true for duration detection as well as picture category decoding.

### 4 Discussion and Future Work

The results show that MEG can classify picture category and presentation duration with high accuracy on a single trial basis, though not as stable as ECoG. More importantly, spatial resolution of the MEG seems to be high enough that only channels from the area expected to have highest activity for the given task are selected by our routines. By manually restricting MEG data to certain channels, decoding accuracies of different ECoG grid placements might

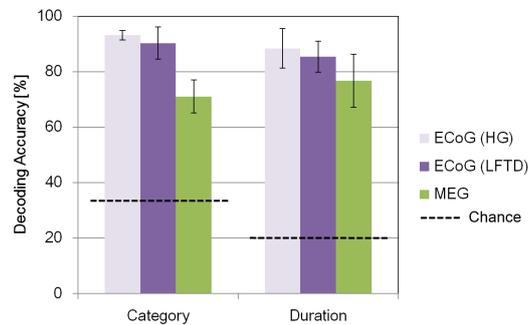


Figure 2: Comparison of the decoding accuracies (in %) of both recording modalities

be roughly assessed. This is helpful to test and improve paradigms designated for ECoG studies. The results also imply a superiority of high frequency information for neural decoding. Because of that, future studies will address gamma activity in MEG in order to get a better comparison. Also, we will investigate simultaneous decoding of both, stimulus category and duration. Findings are expected to be of high interest for BCI due to possible analogies between visual and motoric scenarios. Decoding stimulus quality (here: picture category) and quantity (here: stimulus duration) could translate into the possibility to detect an intended movement direction and duration at the same time, when it comes to realizations of robotic BCI systems.

## Acknowledgements

The work of this paper is partly funded by the German Ministry of Education and Research (BMBF) within the Forschungscampus STIMULATE under grant number '03FO16102A'.

## References

- [1] J. A. Wilson, E. A. Felton, P. C. Garell, G. Schalk, and J.C. Williams. Ecog factors underlying multimodal control of a brain-computer interface. *IEEE Trans. Neural Syst. Rehabil. Eng.*, 14(2):246–250, Jun 2006.
- [2] L. Ding, C. Wilke, B. Xu, X. Xu, W. van Drongelen, M. Kohrman, and B. He. EEG source imaging: Correlating source locations and extents with electrocorticography and surgical resections in epilepsy patients. *Journal of Clinical Neurophysiology*, 24(2):130–136, Apr 2007.
- [3] J. Mellinger, G. Schalk, C. Braun, H. Preissl, W. Rosenstiel, N. Birbaumer, and A. Kübler. An meg-based brain-computer interface (bci). *NeuroImage*, 36(3):581 – 593, 2007.
- [4] T. Wissel, T. Pfeiffer, R. Frysich, R. T. Knight, E. F. Chang, H. Hinrichs, J. W. Rieger, and G. Rose. Hidden markov model and support vector machine based decoding of finger movements using electrocorticography. *Journal of Neural Engineering*, 10(5):056020, 2013.

# N100-P300 Speller BCI with detection of user's input intention

Hikaru Sato and Yoshikazu Washizawa

The University of Electro-Communications, Chofu, Tokyo, Japan  
hikarusato@uec.ac.jp, washizawa@uec.ac.jp

## Abstract

P300 which is evoked by a subject's mental task, has been used for an operation principle of brain computer interfaces (BCIs) such as P300-speller. In this paper, we propose a novel EEG based spelling BCI, N100-P300 speller which uses N100 response in addition to P300. Thanks to using both N100 and P300, the proposed method achieves higher information transfer rate (ITR) than P300-speller. Furthermore, the proposed framework enables us to detect user's input intention with a high degree of accuracy. Our experiment by ten subjects showed that ITR of the proposed system was an average of 0.25bit/sec higher than P300-speller, and the detection accuracy of user's input intention was 90.5 %.

## 1 Introduction

Many types of brain computer interfaces (BCIs) have been developed in the last decade, employing various measuring devices and operation principles [6]. P300-speller which is one of the non-invasive BCIs, is used widely because of its simpleness, high information transfer rate (ITR), and reliability [1]. Several methods to improve the classification accuracy and ITR of P300-speller have been investigated [4, 3, 2].

In practical BCI, detection of user's input intention is a significant function. When we use some input interface, we do not always input information. We sometimes think, wait for a response, or do something else. Although P300-speller can detect user's input intention using P300 response, its detection is not accurate.

In this paper, we investigate new BCI spelling system using N100 as well as P300 in order to improve ITR and the detection accuracy of the user's input intention compared to P300-speller. In the proposed method, nine kinds of stimulus images including a  $2 \times 3$  matrix containing four commands are used (Fig. 1 (a)). P300 is used to detect the target stimulus image containing the desired character. N100 is used to detect the target position that the user gazes on. By using both N100 and P300 features, ITR of the proposed method was 0.25bit/sec higher than that of P300-speller in our experiment by ten subjects. In addition, the detection accuracy of user's input intention was an averaged of 32.7% higher than that of P300-speller.

## 2 Proposed method

### 2.1 Detection of target character

26 alphabets (A, B, ..., Z) and 10 numerals (0,1, ..., 9) are used as spelling commands in the proposed system as well as P300-speller. The characters are arranged in nine stimulus images having four characters and two blanks in the  $2 \times 3$  matrix (Fig. 1 (a)). The desired character is detected by the following three steps: (i) the target image containing the desired character is detected by using P300; (ii) the target position that the user gazes on is detected by using

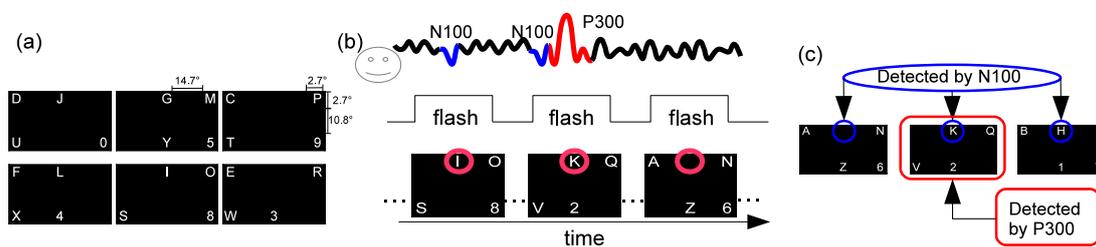


Figure 1: Stimulus images and how to type. (a) Stimulus images. (b) Stimulation and ERP responses for the target character ‘K’. Circles means the user’s gazing position. (c) How to detect the desired character.

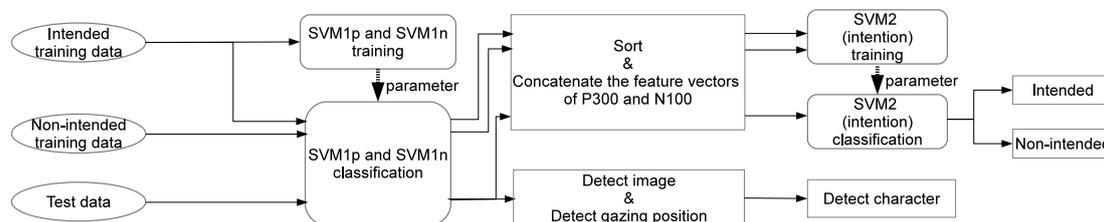


Figure 2: Outline of the proposed method.

N100; (iii) the desired character is detected by the combination of the target image and the target position.

P300 is elicited by mental tasks such as counting how many times the desired character is flashed as well as P300-speller. N100 is elicited by any visual stimuli. In order to detect the user’s gazing position by using N100, two blank positions are arranged for each stimulus image. Since N100 is elicited by any visual stimulus without a mental task [5], we can detect the user’s gazing position by comparing N100 responses for the blank and non-blank positions. Each position has six characters and three blanks for the nine images. The position of characters is informed to the users beforehand by displaying the positioning table to them near the screen. An example of the detection scheme of the desired character ‘K’ is shown in Fig. 1-(b),(c).

Two linear support vector machines (SVMs) are independently used to detect N100 and P300. We respectively denote them by SVM1n and SVM1p.

## 2.2 Detection of the user’s input intention

We categorize the user’s condition into two classes, *intended* and *non-intended*. N100 and P300 responses are obtained during the intended condition, whereas both of them are not elicited during the non-intended condition. This difference is adopted to detect the user’s intention.

Feature vectors to detect the intention are made by concatenating the output values of SVM1n and SVM1p. Output values of SVM1n corresponds to the six positions in the stimulus images. Output values of SVM1p corresponds to the nine stimulus images. These output values are sorted in ascending order before the concatenation. A linear SVM (denoted by SVM2) is trained by the training data of SVM1n and SVM1p, and pre-recorded training data that the user does not intend to input. The procedure is depicted in Fig. 2.

### 3 Experiment and Result

Ten healthy subjects (22-24 years old males) conducted experiments of P300-speller and the proposed system 60 times alternately. These experiments include 20 trials that the user does not input the command without gazing on the display every three times. 40 samples of intended condition and 20 samples of non-intended condition were recorded for each subject. P300-speller flashes 24 times (= 12 flashes  $\times$  two loops), and the proposed system flashes 18 times (= nine flashes  $\times$  two loops) in a trial. Each row and column is presented two times in P300-speller, whereas nine kinds of stimulus images are presented two times in the proposed method per one trial. The flash has 125 ms duration with 62.5 ms inter stimulus interval (ISI). The proposed system takes 3.375 s in a trial, whereas P300-speller takes 4.5 s. The subjects were asked to gaze on only one position that the target character is presented and silently count the number of times the target character flashes. If the maximum instantaneous amplitude of EEG is greater than 100  $\mu$ V, we discarded the trial, and repeated the trial again.

EEG was recorded by using an active EEG (Guger technologies) with a sampling frequency of 512 Hz and a bio-signal amplifier (Digitec) with 0.5Hz analogue high-pass filter and 100 Hz analogue low-pass filter. 16 electrodes (FCz, FC2, FC1, Cz, CP1, CP2, Pz, POz, P3, P4, TP8, TP7, C3, C4, C5, and C6 of the extended international 10-20 system) were used [5]. AFz was used as the ground and A2 was used as the reference. The second-order Butterworth band pass filter (1-13 Hz) and the third-order Butterworth band stop filter (49-51 Hz) were used to extract feature and remove the hum noise, respectively. Signal was downsampled from 512 Hz to 64 Hz.

P300 and N100 components were extracted from 125 ms to 625 ms and from 100 ms to 250 ms after the stimulus, respectively. P300 signals were averaged for each stimulus image in both methods. N100 signals were averaged for each position. The soft margin parameter  $C$  for SVM, such that shows highest averaged accuracy over the five validations for each subject, was selected from  $\{0.1, 1, 10, 100, 1000\}$ . ITR [bit/s] is used as the performance index since the length of one trial is different [6]. ITR is calculated by  $B = \frac{1}{T} \{ \log_2 N + P \log_2 P + (1-P) \log_2 (\frac{1-P}{N-1}) \}$  [bit/s], where  $T$  [s] is the time of one trial,  $P$  is the classification accuracy, and  $N$  is the number of commands. The classification accuracy and ITR were evaluated by 5-fold cross-validation. Hence, the number of training data was 32, and that of test data was eight for SVM1n and SVM1p. For SVM2, the number of training data was 48, (32 samples was from intended trial and 16 samples was from non-intended trial.) The number of test data was 12, (eight samples was from intended trial and four samples was from non-intended trial.)

For P300-speller, since only P300 is used as feature, we obtained the classifier SVM1p and SVM2 to detect the target character and the user's input intention, respectively. SVM1p outputs 12 values corresponding to flashes of six rows and six columns. We sorted these values and obtained a 12-dimensional feature vector to detect the user's input intention.

Table 1 shows the averaged classification accuracies, standard deviations, and ITR. ITR of the proposed method is an average of 0.25bit/sec higher than that of P300-speller because i) the trial length of the proposed system is shorter than that of P300-speller; and ii) the classification accuracy of N100 is higher than that of P300 because N100 signal is averaged over six times, on the other hand, P300 signal is averaged over two times. Moreover, the detection accuracy of the user's input intention of the proposed method is an average of 32.7% higher than that of P300-speller because both N100 and P300 components are used to discriminate the user's intention. The improvement is statistically significant (t-test,  $p < 0.05$ ).

Table 1: Averaged classification accuracy and ITR; “P300” is accuracy of detecting P300 response, and “N100” is accuracy of detecting N100 response. “Intention” is accuracy of detecting the user’s input intention. “Character” is accuracy of detecting the desired character.

Method	Accuracy of feature		Without detection of the user’s input intention		With detection of the user’s input intention		
	P300 [%]	N100 [%]	Character [%]	ITR [bit/sec]	Intention [%]	Character [%]	ITR [bit/sec]
P300-speller	80.9 ± 7.4	–	67.8 ± 15.6	0.60 ± 0.22	57.8 ± 8.5	56.0 ± 21.9	0.46 ± 0.27
Proposed	74.0 ± 18.4	88.5 ± 10.4	70.3 ± 17.1	0.85 ± 0.34	90.5 ± 10.6	66.0 ± 18.6	0.78 ± 0.35

## 4 Conclusion

We have proposed a novel design for a spelling system using both N100 and P300 to reduce the number of flashes per trial, increase ITR, and detect the user’s input intention. The advantages of proposed system are i) it takes shorter time to input one command since the number of flashes can be reduced to nine flashes; ii) the performance of detecting the input intention by using N100 and P300 is 32.7% higher compared to P300-speller in our experiments; iii) each character never flashes in a row whereas at least one character flashes twice in a row in P300-speller; iv) the number of response times can be reduced. That is because the user only has to respond the desired character only two times per trial in the proposed system, whereas in P300-speller the user has to respond it four times, that are the row and column corresponding to the desired character. Hence, the user’s fatigue caused by mental task can be reduced in the proposed system.

In future work, we extend the proposed method to improve ITR by extending the size of matrix. The selection of electrode positions should also be investigated. The comparison between the proposed BCI and other P300-BCIs needs to be done.

## References

- [1] L. A. Farwell and E. Donchin. Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials. *Electroenceph. Clin. Neurophysiol.*, 70:510–523, 1988.
- [2] 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. *NIPS 2008 online*, 2008.
- [3] D. J. Krüsienski, E. W. Sellers, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw. Toward enhanced P300 speller performance. *Journal of Neuroscience Methods*, 167:15–21, 2008.
- [4] W. L. Lee, T. Tan, and Y. H. Leung. An improved P300 extraction using ICA-R for P300-BCI speller. *35th Annual International Conference of the IEEE EMBS*, pages 7064–7067, 2013.
- [5] E. K. Vogel and S. J. Luck. The visual N1 component as an index of a discrimination process. *Psychophysiology*, 37:190–203, 2000.
- [6] J. R. Wolpaw, N. Birbaumer, W. J. Heetderks, D. J. McFarland, P. H. Peckham, G. Schalk, E. Donchin and L. A. Quatrano, C. J. Robinson, and T. M. Vaughan. Brain-computer interface technology: a review of the first international meeting. *IEEE Trans. Rehabil.*, 8:162–173, 2000.

# Training Free Error-Potential Detection

Matthew Dyson<sup>1</sup>, Emmanuel Maby<sup>2</sup>, Laurence Casini<sup>1</sup>, Margaux Perrin<sup>2</sup>,  
J eremie Mattout<sup>2</sup> and Boris Burle<sup>1</sup>

<sup>1</sup> LNC, UMR7291, Aix-Marseille Universit , CNRS, Marseille, France  
matthew.dyson@univ-amu.fr, laurence.casini@univ-amu.fr, boris.burle@univ-amu.fr  
<sup>2</sup> INSERM U1028, CNRS UMR5292, Lyon Neuroscience Research Center, Bron, France  
manu.maby@inserm.fr, margaux.perrin@inserm.fr, jeremie.mattout@inserm.fr

## Abstract

Error potential detection systems typically have to be trained on each individual, often necessitating that users experience sub-optimal active system performance. In this work, by introducing neurophysiological knowledge to preprocess EEG, we present a training-free error potential detection system. Results demonstrate that a subset of error potentials may be discriminated with a high level of precision in previously unseen subjects.

## 1 Introduction

Passive error potential (ErrP) detectors may be utilised to correct or negate the output of active BCI systems, providing output verification. Coupling active and passive BCI systems can increase training requirements. In this analysis we demonstrate an approach which generalises information across subjects, allowing poor BCI performance by some subjects to aid other users.

The underlying EEG component of ErrP detection systems is feedback related negativity (FRN). Key features of FRN make it a primary candidate for the reinforcement learning (RL) theory of error feedback, since it is thought to originate from the anterior cingulate cortex (ACC) [6], and to be triggered by phasic changes in dopaminergic signals from the basal ganglia [3]. In our method neurophysiological knowledge is used to inform projection filters, which perform joint dimensionality and noise reduction [1]. In RL theory, error trials will not necessarily exhibit FRN, therefore we only target a subset of trials using precision based loss functions [4].

## 2 Method

### 2.1 Data

Results from two BCI datasets are presented - detailed descriptions are available for both [2, 5].

The development data set, acquired in Marseille, consisted of 64 channel EEG recorded using a Biosemi Active2. Eleven subjects performed online motor imagery sessions using a Graz protocol modified to display discrete feedback. Subjects performed four runs of 40 trials. Three feedback periods were embedded in each trial. Feedback was manipulated to ensure all subjects experienced minimum error rates of 20%. Data from two subjects was rejected. After artefact removal the average error rate was  $23\pm 2\%$ . We refer to this data as the MI set. A second data set, acquired in Lyon was recorded with 56 EEG sensors using a VSM-CTF compatible system. This analysis is based on 32 electrodes. Sixteen subjects used a pseudorandom stimulation procedure P300-speller. Flash duration and stimulus onset asynchrony parameters were selected to increase the difficulty of spelling. Each subject performed four 12 minute spelling sessions, a session consisted of 12 five letter words with feedback for each letter. After artefact rejection the average error rate was  $30\pm 19\%$ . Four subjects with a error rate below 10% after artefact rejection were not included in this analysis. This data will be referred to as the P300 set.

## 2.2 EEG Preprocessing

EEG data was bandpass filtered between 3 and 12 Hz using 4th order Chebyshev filters. Source time series were obtained using transformation matrices designed for real time use [1]. Applied to EEG, these matrices output  $x$ ,  $y$  and  $z$  current triplet for each of the voxels in the LORETA head model. Voxels belonging to Brodmann areas 24, 32, 33, 8 or 6 were included in a fixed region of interest (ROI), encompassing all FRN activation foci described in meta analysis [6]. The radial,  $z$ , value of all voxels in the ROI were averaged to create a single virtual channel, of sample rate equal to surface channels, on which all further analysis / processing was performed. Using this approach, the virtual channel retains a proxy to polarity observed in central channels.

## 2.3 Feature Space

Training data was generated for each subject using data from others in the same set. Each subject's virtual channel amplitude was normalised by the standard deviation. Subject's virtual channel ERPs were aligned to the median FRN latency of the group, the *normalised latency*, by circular shifting of trial values and peak alignment. A template error FRN was derived from the average of the normalised error trials. Three features were then extracted from a window around the normalised latency. *Amplitude* measured peak negativity within the window. *Jitter* measured the distance between peak negativity and the normalised latency. *Similarity* measured correlation between single trial activity and the template error FRN when accounting for *Jitter*.

## 2.4 Classifier Calibration

SVM<sup>perf</sup>, a classification method for multivariate performance measures [4], was trained on the feature space. A radial basis function kernel and a precision/recall breakeven point loss function were used. SVM parameters cost,  $C$ , and gamma,  $g$ , were obtained through grid search. At each search step nested  $2 \times 10$  cross validation was performed and an F-score obtained from rates of precision and sensitivity. The  $C$  and  $g$  pair which maximised  $F_\beta$  was selected. For P300 calibration five of 12 subjects were excluded from training data (the *hold-back* group) after visual inspection of virtual channel data suggested they had few discriminable error trials.

## 2.5 Online Steps

As subject independent classifiers were calibrated on amplitude and latency normalised feature spaces, online calculations required equivalent distributions. Amplitude normalisation in online steps was based on estimates of statistical moments, updated on a trial by trial basis, using parameterless one pass methods described by Terriberry [7]. Temporal normalisation of the feature space required an estimate be made of each new subject's FRN latency without use of class labels. Latency was tracked via estimation of how a temporal window of activity influenced the *Similarity* feature. A moving average,  $\widehat{Similarity}_t$ , was maintained on a trial by trial basis:

$$\widehat{Similarity}_t = \widehat{Similarity}.UC + \widehat{Similarity}_{t-1} \cdot (1 - UC) \quad (1)$$

The update coefficient,  $UC$ , was bounded between 0 and 0.1. The seed,  $\widehat{Similarity}_0$ , was set to perfect correlation with mean training error FRN, scaled by a mid-range UC value of 0.05. As *Similarity* was invariant to magnitude,  $UC$  was weighted by trial amplitude data in order to only update *Similarity* on trials exhibiting negative polarity in the period of interest:

$$UC = 0.1/1 + e^{13(Amp+0.5)} \quad (2)$$

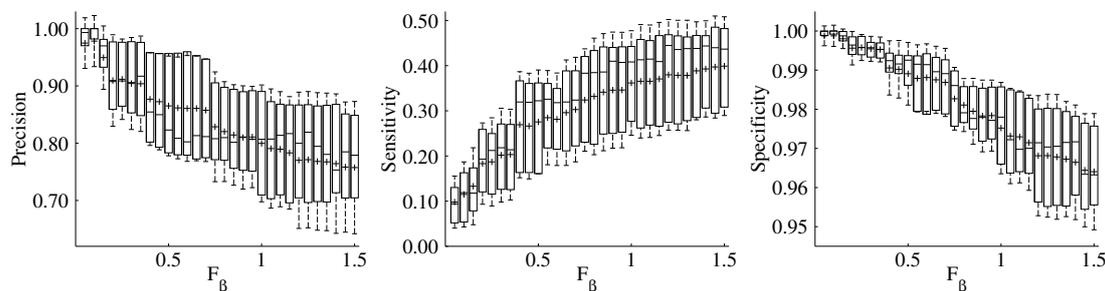


Figure 1: Error detection rates for MI dataset. Influence of  $F_\beta$  selection on overall group rates of precision, sensitivity and specificity. Results calculated from 50 simulation runs for each subject. Whiskers indicate standard deviation. Plus symbol indicates the mean.

$Amp$  being the mean normalised amplitude in a window around the *normalised latency* for the trial. After each trial the estimate of FRN latency,  $\widehat{Latency}$ , was set to the trial period estimated to be most similar to the template error FRN:

$$\widehat{Latency} = t \mid_{\max(\widehat{Similarity}_t)} \quad (3)$$

After amplitude normalisation, *Amplitude*, *Jitter* and *Similarity* features were extracted from a window centred on  $\widehat{Latency}$  and trials classified by SVM. An additional step detected ERP polarity inversion. Polarity inversion may occur due to differences in sulci orientation and the relative position of dipoles. This step was based on sequential calculation of the third central moment and the standard error of skewness. If significant positive skew was detected at  $\widehat{Latency}$ , data was deemed inverted and the polarity of all succeeding trials reversed.

### 3 Results

Figure 1 shows the influence of the F-score  $\beta$  on ErrP detection rates for the MI dataset; demonstrating that overall precision can be increased at the cost of sensitivity and specificity. Further results are based on a  $\beta$  value of 0.5, corresponding to the typical  $F_{0.5}$  measure weighting precision over sensitivity. Post-hoc comparisons were performed on the MI development dataset to verify contributions of the method. Paired t-tests showed source based preprocessing provided better sensitivity ( $M=27.1$ ,  $SD=11.5$ ) than equivalent use of monopolar ( $M=19.7$ ,  $SD=15.0$ ,  $t=-3.23$ ,  $p=0.012$ ); laplacian ( $M=16.9$ ,  $SD=16.1$ ,  $t=-3.75$ ,  $p=0.006$ ); and surface spline ( $M=21.9$ ,  $SD=15.5$ ,  $t=-2.47$ ,  $p=0.039$ ) time series from sensor channel FCz. Use of FRN latency tracking produced significant ( $t=-2.78$ ,  $p=0.024$ ) improvement in sensitivity rates ( $M=27.3$ ,  $SD=11.4$ ) in comparison to use of a fixed value set to the normalised latency ( $M=23.6$ ,  $SD=11.2$ ).

Individual ErrP detection rates are detailed for both the MI and P300 datasets in Table 1. P300 subjects in the *hold-back* group obtain poor overall results, with mean precision, sensitivity and specificity of  $38.6 \pm 27.4$ ,  $5.6 \pm 30$  and  $96.1 \pm 3.6$ . Mean rates for precision, sensitivity and specificity for the seven remaining subjects in the P300 dataset were  $86.9 \pm 15.3$ ,  $21.4 \pm 17$  and  $98.2 \pm 1.9$ . Post-hoc analysis of the properties of these two groups showed the *hold-back* group had a reduced overall difference in surface EEG amplitude between feedback classes, as measured by subtracting correct trial average from error trial average at channel Cz, however this difference was only marginally significant following a one tailed hypothesis test (Welch's t,  $t=-1.8785$ ,  $p=0.049$ ).

	MI Data				P300 Data			
	Prec.	Prec. $\delta$	Sens.	Spec.	Prec.	Prec. $\delta$	Sens.	Spec.
S1	80.0±5.4	+59.3	10.8±1.2	99.3±0.2	<i>9.3±12.3</i>	<i>-28.7</i>	<i>2.2±3.0</i>	<i>89.9±0.8</i>
S2	95.6±1.4	+68.7	42.2±1.8	99.3±0.4	85.1±1.9	+49.4	42.5±3.5	95.9±1.1
S3	83.8±6.2	+61.3	17.6±1.8	99.0±0.3	92.7±3.0	+29.6	10.5±1.1	98.5±0.3
S4	93.2±1.6	+69.1	33.4±2.4	99.2±0.5	99.9±0.8	+44.9	16.7±1.0	100.0±0.3
S5	97.0±2.8	+74.1	32.1±2.2	99.7±0.5	<i>11.5±23.7</i>	<i>-16.3</i>	<i>0.9±1.9</i>	<i>98.5±0.4</i>
S6	77.4±5.1	+55.9	12.3±1.5	99.0±0.3	99.0±6.9	+39.6	0.8±0.7	100.0±0.2
S7	95.9±2.6	+75.5	35.9±3.1	99.6±0.6	<i>57.5±5.7</i>	<i>+33.3</i>	<i>18.5±1.8</i>	<i>95.5±0.4</i>
S8	79.3±6.2	+53.2	23.7±5.9	97.9±1.3	62.3±3.6	+49.4	47.1±3.3	95.8±0.4
S9	77.4±4.7	+57.7	36.8±2.6	97.4±0.5	69.8±6.3	+41.9	14.3±1.6	97.5±0.4
S10					99.5±2.7	+56.8	18.1±3.1	99.9±0.9
S11					<i>44.0±18.3</i>	<i>-2.1</i>	<i>1.7±0.7</i>	<i>98.2±0.3</i>
S12					<i>70.5±5.3</i>	<i>+39.1</i>	<i>9.9±2.0</i>	<i>98.1±0.5</i>
$\mu$	86.6±8.6	+63.9	27.2±11.5	98.9±0.8	66.8±31.9	+28.1	15.3±15.4	97.3±2.8

Table 1: Error-potential detection rates for MI and P300 data sets. Precision  $\delta$  denotes difference between precision obtained and equivalent randomisation of true label distributions. *Italics* denote members of the *hold-back* group. Results obtained using  $\beta$  value of 0.5. Simulation runs were repeated 100 times.

## 4 Conclusion

Discrimination rates demonstrate that strong neurophysiological a priori, with inverse solution methods to apply them, can produce viable training free approaches for ErrP detection. Signal enhancement provided by inverse methods can enable use of simple descriptive interpretable features.

## 5 Acknowledgments

This work was carried out as part of the CoAdapt project, funded via ANR-09-EMER-002-01.

## References

- [1] M. Congedo. OpenViBE tools. <http://sites.google.com/site/marcocongedo>, 2011.
- [2] M. Dyson, L. Casini, and B. Burle. Discrimination of discrete feedback during performance of motor imagery. In *Proc. Int. Workshop on Pattern Recognition in NeuroImaging*, pages 93–96, 2012.
- [3] C. B. Holroyd and M. G. H. Coles. The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, 109(4):679–709, 2002.
- [4] T. Joachims. A support vector method for multivariate performance measures. In *Proceedings of the 22nd International Conference on Machine Learning*, pages 377–384. ACM Press, 2005.
- [5] M. Perrin, E. Maby, S. Daligault, O. Bertrand, and J. Mattout. Objective and subjective evaluation of online error correction during P300-based spelling. *Adv. in Human-Computer Interaction*, 2012.
- [6] K. R. Ridderinkhof, M. Ullsperger, E.A. Crone, and S. Nieuwenhuis. The roles of the medial frontal cortex in cognitive control. *Science*, 306:443–447, 2004.
- [7] T. B. Terriberry. Computing Higher-Order Moments Online. <http://people.xiph.org/~tteribe/notes/homs.html>, 2007.

# 3D visualization modalities can have effects on motor cortex activation

Teresa Sollfrank<sup>1\*</sup>, Daniel Hart<sup>2</sup>, Rachel Goodsell<sup>3</sup>, Jonathan Foster<sup>3</sup>,  
Andrea Kübler<sup>1</sup> and Tele Tan<sup>2†</sup>

<sup>1</sup>University of Würzburg, Würzburg, Germany

<sup>2</sup>Department of Mechanical Engineering, Curtin University, Perth, Western Australia

<sup>3</sup>School of Psychology and Speech Pathology, Curtin University, Perth, Western Australia &  
Neurosciences Unit, Health Department of WA

Teresa.sollfrank@uni-wuerzburg.de, T.tan@curtin.edu.au

## Abstract

When pathways for normal motor function are interrupted (e.g. after stroke), brain-computer interfaces (BCI) can be used i) as an alternative channel for communication by translating brain signals measured with electroencephalography (EEG) into a computer output (Kübler et al., 2005) or ii) for rehabilitation by influencing brain plasticity processes to induce recovery of motor control (Pichiorri et al., 2011). This study investigated if a realistic visualization of an upper limb movement can amplify motor related potentials during motor imagery (MI). We hypothesized that a 3D sensory richer visualization might be more effective during instrumental conditioning, resulting in more pronounced event related desynchronisation (ERD) of the mu band (10-12Hz) over the sensorimotor cortices and can therefore improve sensorimotor rhythm-based based BCI protocols for motor rehabilitation.

## 1 Introduction

Over the past years advances in the analysis of EEG signals and improved computing capabilities have enabled people with severe motor disabilities to use their brain activity for communication and control of objects in their environment, thereby bypassing their impaired neuromuscular system (Kübler et al., 2001; Allison et al., 2007). A new potential BCI therapeutic approach is generating substantial interest in the use of EEG-based BCI protocols to improve volitional motor control, that has been impaired by trauma or disease: A repetitive movement practice e.g., with non-invasive sensorimotor rhythm-based (SMR) BCIs should lead to an increase in the motor cortical excitability (Pichiorri et al., 2011). By inducing a better engagement of motor areas with respect to motor imagery, BCI protocols might be able to guide the neuroplasticity to promote recovery in the affected brain regions to restore motor function (Cincotti et al., 2012).

Since BCI systems use immediate, typically visual feedback of performance, the influence of the visual feedback presentation should be considered. For instance, there seems to be no difference in the classification results between groups provided with ‘realistic feedback’ (moving hand performing an object-related grasp) and an “abstract feedback” (moving bar; Neuper et al., 2009) in a SMR based-BCI task. However there is some evidence that instead a rich visual representation of the feedback

---

\* Masterminded EasyChair and created the first stable version of this document

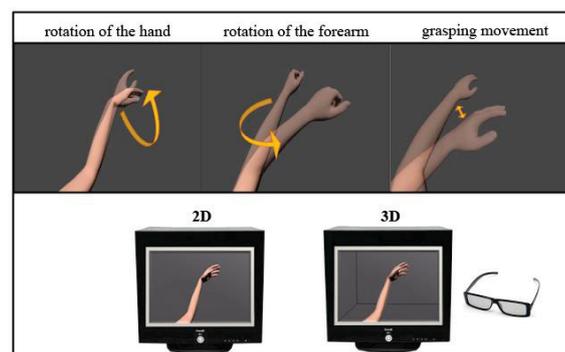
† Created the first draft of this document

signal, e.g., in the form of a three dimensional video game or virtual reality environment, may enhance the user's control of a SMR based-BCI (Pineda et al., 2003; Pfurtscheller et al., 2006).

This study investigated if a 3-dimensional visualization of upper limb movement can amplify motor cortex activation during motor imagery (MI) and thereby support the use of a sensorimotor rhythm based BCI. We hypothesize that this "realistic" sensory richer visualization might be more effective during instrumental conditioning, resulting in more pronounced event related desynchronisation of the mu band (10-12Hz) over the sensorimotor cortices.

## 2 Methods

Fifteen healthy SMR BCI naïve participants were recruited for the study. In a within subject design all participants were instructed to watch attentively randomized videos of three different left and right upper limb movements (Fig. 1) on a True3Di monitor in 2D and in 3D visualization, by using stereoscopic glasses. Every session consisted of 3 runs with 12 trials each (3 different movements for the left and right upper limb in randomized order consecutively in 2D and 3D). After every video, participants were instructed to replicate subsequently the observed movements by motor imagery, while the EEG signals were recorded from a grid of 40 Ag/AgCl scalp electrodes. For statistical analyses, we used the ERD values obtained from the right (recording position CP4) versus left sensorimotor cortex (recording position CP3), temporally aggregated over imagery period (1-5 s). The probability of a Type I error was maintained at 0.05

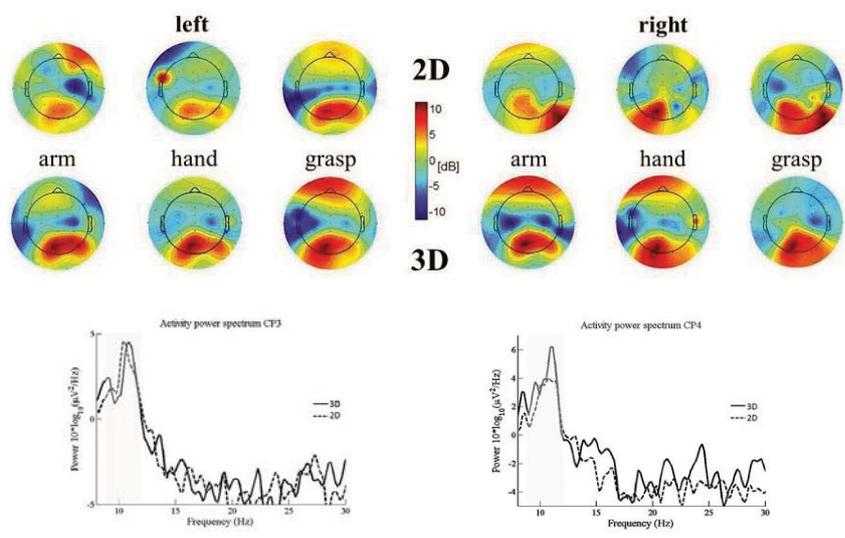


**Figure 1:** Visualization of three different upper limb movements: rotation of the hand around wrist, rotation of the whole forearm around elbow and a grasping movement of the whole arm. All movements were shown for the left and right upper limb randomized in 2D and 3D.

## 3 Results

Fig. 2 presents the topographical analyses of the mean ERD values for the two visualization modality groups, separately for the respective task (right and left hand, arm and grasp motor imagery) and electrode position (CP3 and CP4). Characteristic patterns of ERD of mu band components for left and right upper limb motor imagery were present mainly contralateral at electrode positions CP3 and CP4 over the sensorimotor areas. In order to analyze the potential influence of the visualization modality on the ERD patterns during task performance, we performed a repeated measures ANOVA on the ERD data using the Visualization Modality (VM, 2 levels: 2D vs. 3D), Electrode Position (2

levels: CP3 vs. CP4), Task (3 levels: arm rotation, hand rotation and grasping motor imagery) and Task Side (2 levels: left vs. right) as within-subjects variables. Overall, significant differences were observed as a function of Visualization Modality. This main effect indicates that largest ERD was obtained with a 3-dimensional visualization modality. ANOVA revealed a significant interaction of ‘VM’ x ‘Task Side’ x ‘Electrode Position’ ( $F_{3,01} = 8.77, p < 0.001$ ), ‘VM’ x ‘Task’ ( $F_{2,22} = 5.45, p < 0.005$ ) and a significant interaction between electrode position and side of the motor imagery (Electrode Position x Task Side;  $F_{(1,56)} = 18.25, p = 0.001$ ). Largest enhancement in the mu band event related desynchronisation was obtained in the tasks: ‘3D arm mi left’ on the right hemisphere at electrode position CP4 ( $t = 2.75, p < 0.005$ ), ‘3D arm mi right’ on the left hemisphere at CP3 ( $t = 2.62, p < 0.005$ ) and ‘3D hand mi left’ on the right hemisphere at CP4 ( $t = 2.64, p < 0.005$ ), compared to the 2-dimensional task.



**Figure 2:** Maps displaying the topographical distribution of mu band power (10-12 Hz; ERD) during motor imagery tasks (left and right upper limb) after presenting 2D and 3D limb movement visualizations. The two graphs show the modulated power spectral density during motor imagery (averaged 15 participants, baseline corrected).

## 4 Discussion

Our findings support the assumption that stimulus-rich and realistic feedback conditions can lead to stronger motor cortex activation during motor imagery (Pineda et al., 2003; Pfurtscheller et al., 2007). Depending on the visualization modality different activations of the hemispheres could be detected corresponding to the two different electrode positions CP4 and CP3. Furthermore, this motor cortex activation was also related to the three different motor imagery tasks depending on the type of visualization. Participants were able to enhance the desired electrophysiological signals significantly in three of the 3-dimensional tasks and according to findings of Neuper et al. (2009) a contralateral dominance of the ERD was detectable. In the tasks where the 2-dimensional visualization lead to comparable strong ERDs the visualization seemed to become closely associated to the MI task and no additional benefit could be gained from the sensory richer realistic 3D visualization.

## 5 Conclusion

Motor imagery offers a promising technique for motor rehabilitation (Cincotti et al., 2012). With a realistic visualization of the limb movements, we tried to increase motor cortex activation during motor imagery and thereby gradually improve performance in SMR based BCI tasks. These preliminary results suggest that the patient's involvement in the motor imagery task and the functional outcome in the motor related potentials may be improved by the use of the 3-dimensional visualization modality.

## 6 Acknowledgement

This work was funded by the Deutscher Akademischer Austauschdienst (DAAD), the Department of Mechanical Engineering at the Curtin University (Perth/ Western Australia) and partly supported by the GK Emotion and the EU grant FP7-ICT-2011 "CONTRAST" project.

## 7 References

- Allison, B.Z., Wolpaw, E.W. and Wolpaw, J.R. (2007) Brain-computer interface systems: progress and prospects. *Expert Review of Medical Devices*, Vol. 4, No. 4, 463-474.
- Cincotti, F., Pichiorri, F., Arico, P., Leotta, F., de Vico Fallani, F., del R Millan, J., Molinari, M. and Mattia, D. (2012) EEG-based Brain-Computer Interface to support post-stroke motor rehabilitation of the upper limb. *Proceedings of the 34<sup>th</sup> Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Diego, CA*, pp. 4112-4115.
- Kübler, A., Kotchoubey, B., Kaiser, J., Wolpaw, J.R. and Bierbaumer, N. (2001) Brain-computer communication: Unlocking the locked in. *Psychological Bulletin*, Vol. 127, No. 3.
- Kübler, A., Nijober, M.S., Mellinger, J., Vaughan, T.M., Pawlzik, H., Schalk, G., McFarland, D.J., Birbaumer, N. and Wolpaw, J.R. (2005) Patients with ALS can use sensorimotor rhythms to operate a brain-computer interface. *Neurology*, Vol. 64, No. 10, pp. 1775-1777.
- Neuper, C., Scherer, R., Wriessneggerand, S. and Pfurtscheller, G. (2009) Motor imagery and action observation: Modulation of sensorimotor brain rhythms during mental control of a brain-computer interface. *Clinical Neurophysiology*, Vol. 120, Issue 2, pp 239-247.
- Pfurtscheller, G., Scherer, R., Leeb, R., Keinrath, C., Neuper, C., Lee, F. & Bischof, H. (2007) Viewing Moving Objects in Virtual Reality Can Change the Dynamics of Sensorimotor EEG Rhythms. *Presence*, Vol. 16, No.1, pp 111-118.
- Pichorri, F., De Vico Fallani, F., Cincotti, F., Babiloni, F., Molinari, M., Kleih, S.C., Neuper, C., Kübler A. and Mattia, D. (2011) Sensorimotor rhythm-based brain-computer interface training: the impact on motor cortical responsiveness. *J. Neural Eng.*, Vol.8, No.2, 025020.
- Pineda, J.A., Silverman, D.S., Vankov, A. and Hestenes, J. (2003) Learning to control brain rhythms: making a brain-computer interface possible. *IEEE Trans Neural Syst Rehabil Eng*, 11, pp. 181-184.

# The MOPS-Speller: An Intuitive Auditory Brain-Computer Interface Spelling Application

Sonja C. Kleih<sup>1</sup>, Andreas Herweg<sup>1</sup>, Tobias Kaufmann<sup>2</sup> and Andrea Kübler<sup>1</sup>

<sup>1</sup>University of Würzburg, Würzburg, Germany

<sup>2</sup>University of Oslo, Oslo, Norway

sonja.kleih@uni-wuerzburg.de

## Abstract

In auditory Brain-Computer Interface (BCI) systems for communication, usually the letters of the alphabet are grouped to be spelled since presenting them all one after another would be too time consuming. Each group is usually assigned a number or a sound and letters are selected by making several grouping choices in a row. Herein we introduce the MOPS-Speller, which uses natural language words for grouping. The desired letters can be selected intuitively as they appear in the grouping word, thereby presenting an intuitive auditory spelling paradigm for communication.

## 1 Introduction

To allow for BCI based communication even in patients who are unable to control their eye movement and therefore, cannot benefit from visual input has been one crucial research goal in Brain-Computer Interfacing (BCI). In the P300-based auditory paradigm presented by Schreuder and colleagues (2011) letters of the alphabet were divided into six groups and each group was represented by one sound (consisting of a base tone ranging between 440 and 1099 Hz and a noise ranging between 3 and 10.5 KHz) presented from spatially distributed loudspeakers. The six loudspeakers were arranged in a circle around the person (in a distance of 60° from each other) and one could choose a letter group by focusing attention to the sound as well as the direction from which the sound was presented (spatial information). After the group was selected, single letters were presented to the participant and the desired letter could be chosen. Another promising P300-based approach was presented by Höhne and colleagues who combined sound frequency (high=708Hz, medium=524Hz and low=480Hz) and spatial information (right ear, left ear, both ears) via headphones in a paradigm that was based on the T9 system that is used for mobile phones. By focusing attention on the desired group of letters, possible letter combinations were narrowed down to one single desired word and the user could spell fast and easy.

However, in both systems, the letters or groups of letters were represented by sounds in combination with spatial information. Therefore, one has to keep in mind which group is represented by which tone and from which direction. Also the presentation of a visual support matrix decreases external validity as in case visual input cannot be perceived, the presentation of a support matrix would not represent a realizable solution for a patient. Thus, in this paper, we present a speller which uses natural words as stimuli and thereby allows intuitive communication for which no visual support matrix is needed. Spatial information was not implemented in the here presented speller as we aimed at an intuitive and easy-to-set-up solution for auditory BCI use in people with motor impairment.

## 2 Methods

### 2.1 Participants

We tested N=2 severely motor impaired patients. Participant A was paralyzed in a wheelchair because of a traumatic accident, participant B was diagnosed with muscle dystrophy (Duchenne) and also used a wheelchair. Both participants were male and 43 and 49 years old. Both participants used assistive devices for communication but could communicate single words by voice. Participants gave informed consent for the study, which had been reviewed and approved by the Ethical Review Board of the University of Würzburg.

### 2.2 Material and Procedure

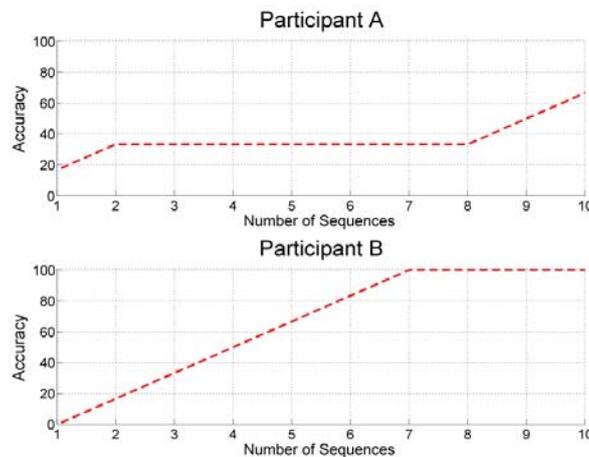
In the MOPS speller the German words “MOPS”, “BUCH”, “KLANG”, “FEDER”, “WITZ” and the non-word “JQVXY” represented the letter groups. Additionally the word “RÜCK” (engl: “undo”) which allows for correction of a false choice was added. Participants underwent a calibration in which all word stimuli were presented 10 times. While participant A needed 10 sequences for spelling 80% correct which was judged to be sufficient, participant B needed 6 sequences according to calibration results. For the online spelling, stimuli were presented to the participants in random order while they focused on the word containing the letter they intended to select. After correct choice of this word, the single letters of the word were presented to the user in random order. Thus, for every letter to spell one word and one letter had to be correctly selected. The lengths of the auditory stimuli varied between 401 and 1162ms, the Inter-Stimulus-Interval (ISI) was set to 800ms. During the copy-spelling task, participants were told to spell the words “BOJE”, “SYLT” and “HARZ”. They were also informed which word to focus on first and after that choice, which letter to spell. Their spelling progress was fed back by presenting to them the single letters already spelled. All stimuli were recorded with the Cubase LE5 Software (female speaker).

### 2.3 Data Acquisition and Analysis

Stimulus presentation was controlled via Python© (version 2.5, Python Software Foundation) and was linked via UDP to BCI 2000 (version 3, Schalk et al., 2004) in which data were recorded and stored. For EEG acquisition we used 12 passive Ag/AgCl electrodes on the locations Fz, PC1, PC2, C3, Cz, C4, P3, Pz, P4, O1, Oz, O2 according to the international 10-20 system. Each electrode was referenced to the right and grounded to the left mastoid. The EEG was amplified using a 16-channel g.USB amplifier (Guger Technologies, Austria), sampled at 256 Hz, and bandpass filtered between 0.01 – 30 Hz. Fifty Hz noise was filtered using a notch filter implemented in the BCI2000 software. Data processing, storage and stimulus presentation was controlled with an LG computer (Intel Core 2 Duo, 4 GHz, Windows 7), loudspeakers were Hama AL-140 Stereo Speaker (Monheim, Germany). The EEG data were corrected for artifacts and baseline (-100 to 0 ms) using MATLAB© (v2011b). The P300 was defined as the maximum positive peak occurring between 200 and 800 ms after stimulus onset. As the here presented sample exists of two participants only, we did not perform statistical analysis on the data but present the results descriptively. To generate the feature weights for the classifier in the online spelling as well as for offline classification accuracy calculation after artifact removal, stepwise linear discriminant analysis (SWLDA) was used.

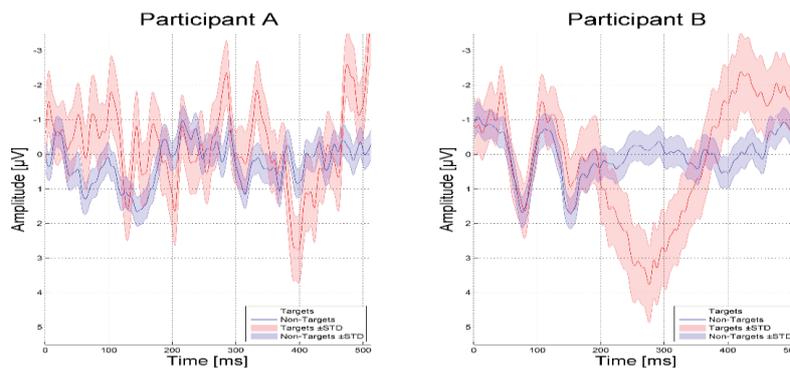
### 3 Results

Participant A spelled very successfully with the MOPS Speller with an average online accuracy of 90% correctly selected letters. The three times he had to use the undo function, he corrected his previous false choice properly and spelled correctly all three words “BOJE”, “SYLT” and “HARZ”. His offline classification accuracy was 100% after 7 sequences (see figure1). Participant B was less successful with the MOPS-speller and reached an average accuracy of 53% correctly selected letters online. His offline classification accuracy was 65% with 10 sequences of stimulus presentation (see figure1).



**Figure 1:** Classification accuracy as a function of number of presented sequences.

In line with the presented spelling result we found main differences for signal distinction between targets and non-targets between the two patients. While in participant B, the P300 is very clearly detectable (see figure 2), targets and non-targets were less distinguishable in participant A (see figure 2). The Information Transfer Rate (ITR) was 0.30 for participant A and 1.04 for participant B.



**Figure 2:** P300 evoked by targets (red) compared to non-targets (blue). Red and blue shades indicate the standard deviation (SD) of EEG signals.

## 4 Discussion

With the here presented auditory speller, we showed that intuitive auditory communication independent from a visual support matrix can be used successfully by people with severe motor impairment. However, we cannot judge general usability of the MOPS-speller as we only included two patients, and one did not reach sufficient accuracy for meaningful communication (Kübler et al., 2001), but we demonstrated the proof-of-principle.

Our major goal of using intuitive stimuli in an easy-to-setup-paradigm was very well accepted by the participants. They reported the auditory stimulation with natural words to be pleasant and simple to understand, even though after spelling the second word, they started to feel exhausted. Also the stimuli themselves were long and together with the here implemented ISI, ITR is low. Nonetheless, we believe that for a patient who is unable to use the visual input channel, the time that is needed for spelling would still be acceptable.

## Acknowledgements

This work is supported by the European ICT Program Project FP7-287320 (CONTRAST). This manuscript only reflects the authors' views and funding agencies are not liable for any use that may be made of the information contained herein.

## References

- Höhne, J., Schreuder, M., Blankertz, B., & Tangermann, M. (2011). A Novel 9-Class Auditory ERP Paradigm Driving a Predictive Text Entry System. *Front Neurosci*, 5, 99. doi: 10.3389/fnins.2011.00099.
- Kübler, A., Kotchoubey, B., Kaiser, J., Wolpaw, J. R., & Birbaumer, N. (2001). Brain-computer communication: unlocking the locked in. *Psychol Bull*, 127(3), 358-375.
- Schalk, G., McFarland, D. J., Hinterberger, T., Birbaumer, N., & Wolpaw, J. R. (2004). BCI2000: a general-purpose brain-computer interface (BCI) system. *IEEE Trans Biomed Eng*, 51(6), 1034-1043. doi: 10.1109/TBME.2004.827072
- Schreuder, M., Rost, T., & Tangermann, M. (2011). Listen, You are Writing! Speeding up Online Spelling with a Dynamic Auditory BCI. *Front Neurosci*, 5, 112. doi: 10.3389/fnins.2011.00112

# An asynchronous BMI for autonomous robotic grasping based on SSVEF detection

C. Reichert<sup>1,2,3</sup>, M. Kennel<sup>4</sup>, R. Kruse<sup>2</sup> and H. Hinrichs<sup>1,3</sup>

<sup>1</sup> Dept. of Neurology, University Medical Center A.ö.R., Magdeburg, Germany  
christoph.reichert@med.ovgu.de, hermann.hinrichs@med.ovgu.de

<sup>2</sup> Dept. of Knowledge and Language Processing, Otto-von-Guericke University, Magdeburg, Germany

<sup>3</sup> Forschungscampus STIMULATE, Magdeburg, Germany

<sup>4</sup> Fraunhofer Institute for Factory Operation and Automation IFF, Magdeburg, Germany

## Abstract

Severely impaired persons could greatly benefit from assistive devices controlled by brain activity. However, the low information transfer rate of noninvasive neuroimaging techniques complicates complex and asynchronous control of robotic devices enormously. In this paper we present an asynchronous brain-machine interface (BMI) relying on autonomous grasp planning. The system enables a user to grasp and manipulate objects with a minimal set of commands. We successfully tested the system in a virtual environment with eight subjects. Our results suggest that the system represents a promising approach for real-world application of brain-controlled intelligent robotic devices.

## 1 Introduction

One of the most convenient applications of brain-machine interfaces is the control of prosthetic devices. Many approaches have been developed so far, including non-invasive [1] and invasive [2] control of an artificial upper limb. However, for continuous control, the commonly performed decoding of motor sensory rhythms, originated in motor imagery, has two drawbacks: due to only a few distinguishable classes the degrees of freedom to control the effector are limited and the training phase for users is long. An alternative control signal is the steady state visual evoked potential (SSVEP) which provides a robust signal for detection and an easy usage. In a previous study [3] we showed that the magnetic fields induced by SSVEPs (steady state visual evoked field, SSVEF) can be reliably decoded to select virtual reality objects. The control of a robotic hand with SSVEP detection was presented by [4] but still the degrees of freedom which can be controlled are very limited. Here we present an SSVEF based prototype system, aimed at autonomously grasping and manipulating objects, which is controlled asynchronously via visual overt attention. The decoding scheme we applied is designed to minimize erroneous actions of the manipulator which facilitates successful and secure grasping when combined with an intelligent actuator.

## 2 Methods

### 2.1 Robotic Grasping System

We developed a virtual reality scenario designed to simulate the control of a robotic device. Virtual objects, placed in the working area of the robot, can be grasped autonomously by the robot [5, 6]. In the current setup we placed four spherical objects on the table serving as potential grasp targets. Furthermore, we placed a green start button and a red stop button on the table to send control commands to the robot. All objects were tagged with stimulation

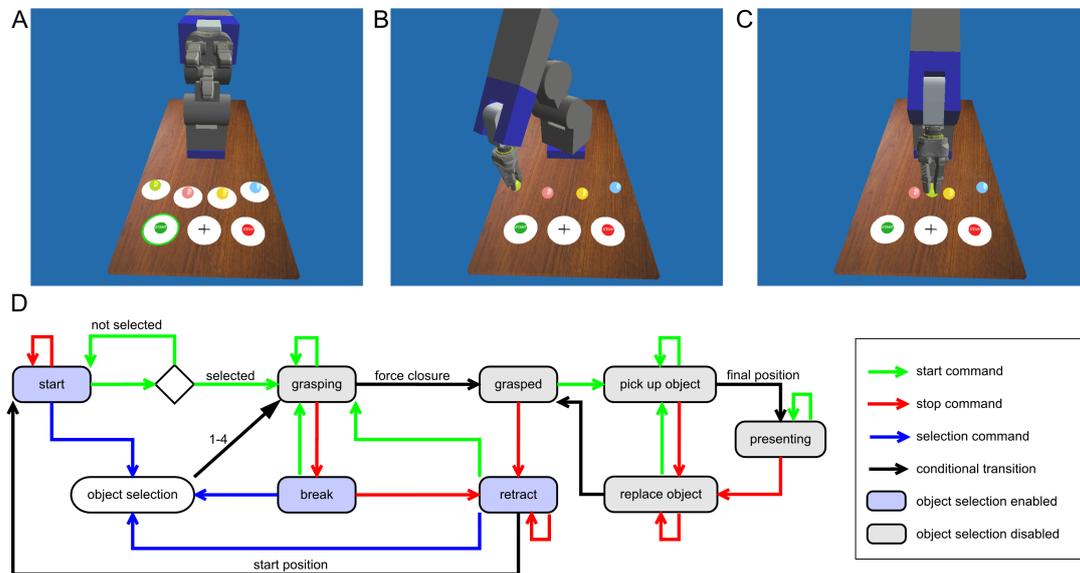


Figure 1: Robotic gripper in the states *start* (A), *grasped* (B) and *presenting* (C) and the state diagram (D) of the system which is controllable by only three commands (start, stop and object#).

surfaces flickering at different frequencies (object 1–4: 6.818 Hz, 7.5 Hz, 8.333 Hz, 9.375 Hz; start button: 10.714 Hz; stop button: 12.5 Hz). Whenever only the selection of a command button was reasonable, e.g. during a running grasping process, the stimulation of objects 1–4 was stopped. Thus, selection of a new object to grasp was only possible when the gripper was in the states *retract*, *break* and *start*. This reduces the probability of unwanted selections during the grasping and manipulation process. As a representative manipulation the object could be picked up and presented. For resting periods, a non-flickering area, e.g. the fixation cross must have been focused. For the selection of the object to be grasped next, the corresponding flickering surface has to be attended. The selection of the start button initiates a grasp to the lastly selected object, the presentation of the object and the resuming of a stopped grasp movement. In contrast, a selection of the stop button causes the reverse operations. Figure 1 (D) visualizes the state transitions of the robot dependent on activation of the start button (green transitions) and the stop button (red transitions). States marked in grey indicate stimulation of start and stop button only, states marked in blue indicate stimulation of all six surfaces.

## 2.2 Decoding Algorithm

We decoded the SSVEF by means of canonical correlation analysis (CCA) [7]. This method finds the optimal weighting  $W_{x,f}$  and  $W_{y,f}$  of input channels  $X$  (here the signals of 59 occipital MEG sensors) and model channels  $Y_f$  (sine and cosine waves of the stimulation frequency  $f$  and optionally harmonics) to maximize the correlation  $\rho_{CCA}(f)$  between linear combinations  $x_f = X^T W_{x,f}$  and  $y_f = Y_f^T W_{y,f}$ . The CCA was applied to MEG data asynchronously every 1000 ms from the preceding 2000 ms interval. To distinguish resting intervals from selection intervals, we determined a confidence threshold  $th_f$  for each frequency  $f$  from the empirical

distribution of coefficients  $\rho_{CCA}$  which we revealed in calibration runs. Finally, the classification was determined according to the frequency  $f_{max} = \operatorname{argmax}(\rho_{CCA}(f))$  that revealed the maximum correlation coefficient and exceeded the threshold  $th_f$ .

Feedback of the instantaneous target detection of the decoder was presented at each decoding step by coloured rings around the decoded target. Only after three successive predictions of the same target it was finally selected. The stage of prediction was indicated by stepwise colour changes from light blue to green for the final selection. After a selection, the next decoding was performed three seconds after feedback presentation. The detection of a resting period was displayed at the fixation cross identical to the selectable targets and did not affect the prediction stage of targets.

### 2.3 Experimental Setup

Eight healthy subjects (age 21–34 years) participated in this study. MEG data were acquired with a 248 magnetometer 4DNeuroimaging system, at a sampling rate of 678.17 Hz.

The task of the subjects was to control the autonomously grasping robot depicted in Figure 1 (A–C) by overtly attending the flickering surfaces. The experiment started with calibration runs in which one of six targets or the fixation cross was cued by a black ring during each trial. If the final selection was different from the cued target a red ring indicated the erroneous decoding. A trial was counted as false after 30s of selection failure. The subjects performed 3–4 calibration runs to obtain a reliable performance estimate of the proposed decoding scheme. Additional 3–4 runs were performed to control the robot. In these runs, a high amount of resting periods was required due to the slow movement of the robot. Consequently, relatively few selection trials were performed. Subjects were instructed to grasp, pick up and replace the objects in a predefined order, which was announced by the experimenter directly before the run. In case of an unwanted selection due to erroneous SSVEF detection, the subject had to perform the consequential selection to continue the desired workflow. The sequence of control commands was as follows: Select the object to be grasped, select the start button to grasp the object, select the start button to present the object, select the stop button to replace the object, select the stop button to retract the gripper. In intervals of robot movement, subjects had to prevent selections.

## 3 Results

Although feedback was provided from the first run the decoding of resting intervals started with the second run, when initial calibration data were available. On average, 74% of target trials ( $T_{corr}/trials$ ) and 73% of resting trials ( $R_{corr}/trials$ ) were decoded correctly while the guessing level was 14%. Average selection time  $t_{select}$  was 11.7s while fastest possible selection was 5s. In robot control runs, the speed of a complete grasp sequence ( $G_{compl}/time$ ) was  $0.52 \text{ min}^{-1}$  on average. In order to safely perform gripper control, the movement of the robot was relatively slow (grasping: 28s, picking up: 8s) such that an optimal complete grasp sequence, without complete retracting but including selections, was possible in at least 69s (maximum  $G_{compl}/time = 0.87 \text{ min}^{-1}$ ). As an important result, in our setup false object selections ( $S_{err}$ ) occurred reasonably (26%) but accidentally initiated robot commands ( $C_{err}$ ) occurred rarely (4.7%) although long intervals of preventing erroneous selections were required. Table 1 summarizes the results. The number of trials performed depends on the latency of a final selection and on the number of unwanted selections. One of the subjects was able to perfectly control the system without any erroneous selection.

Subject #	Calibration			Robot control		
	$T_{corr}$ /trials	$R_{corr}$ /trials	$t_{select}$ [s]	$G_{compl}$ /time	$C_{err}$ /trials	$S_{err}$ /trials
1	56 / 75	18 / 25	12.1	8.2 / 18'04"	3 / 38	3 / 14
2	11 / 27	4 / 10	16.5	7.4 / 24'04"	8 / 44	2 / 13
3	62 / 62	23 / 23	10.9	16.6 / 24'06"	0 / 67	0 / 20
4	40 / 63	15 / 23	10.7	10.2 / 24'23"	2 / 44	6 / 26
5	46 / 64	21 / 23	10.5	12.4 / 24'22"	1 / 50	13 / 33
6	46 / 54	17 / 21	12.8	12.6 / 24'48"	1 / 60	4 / 24
7	60 / 69	16 / 25	9.7	15.0 / 21'39"	1 / 61	14 / 35
8	45 / 67	16 / 23	10.3	10.3 / 23'39"	1 / 53	21 / 38

Table 1: Subject performance

## 4 Conclusions

As an alternative modality to EEG we probed the MEG for detecting SSVEFs. Provided that the EEG allows for comparable detection success, our results indicate that an intelligent actuator in combination with noninvasive detection of SSVEPs could help severely paralyzed patients to initiate complex grasp movements asynchronously. The presented system allows for voluntary robot control with a low error rate even in subjects with moderate selection performance, indicating that noninvasive BMIs may be suitable for grasping despite its low information transfer rate. It is important to note that the transfer of the system into an EEG-driven real-world setting requires further development towards more elaborated stimulus presentation techniques and intelligent prostheses. In addition, a combination with other modalities, for instance tracking of eye movements, could serve to make the system even more robust.

## References

- [1] G. R. Müller-Putz, R. Scherer, G. Pfurtscheller, and R. Rupp. EEG-based neuroprosthesis control: a step towards clinical practice. *Neuroscience Letters*, 382(1-2):169–174, 2005.
- [2] L. R. Hochberg, D. Bacher, B. Jarosiewicz, N. Y. Masse, J. D. Simeral, J. Vogel, S. Haddadin, J. Liu, S. S. Cash, P. van der Smagt, and J. P. Donoghue. Reach and grasp by people with tetraplegia using a neurally controlled robotic arm. *Nature*, 485(7398):372–375, 2012.
- [3] C. Reichert, M. Kennel, R. Kruse, H. Hinrichs, and J. W. Rieger. Efficiency of SSVEF recognition from the magnetoencephalogram - A comparison of spectral feature classification and CCA-based prediction. In *Proceedings of the International Congress on Neurotechnology, Electronics and Informatics*, pages 233–237. SciTePress Digital Library, Sep 2013.
- [4] G. R. Müller-Putz and G. Pfurtscheller. Control of an electrical prosthesis with an SSVEP-based BCI. *IEEE Trans Biomed Eng*, 55(1):361–364, 2008.
- [5] C. Reichert, M. Kennel, R. Kruse, H.-J. Heinze, U. Schmucker, H. Hinrichs, and J. W. Rieger. Robotic grasp initiation by gaze independent brain-controlled selection of virtual reality objects. In *Proceedings of the International Congress on Neurotechnology, Electronics and Informatics*, pages 5–12. SciTePress Digital Library, Sep 2013.
- [6] M. Kennel, C. Reichert, U. Schmucker, H. Hinrichs, and J. W. Rieger. A robot for brain-controlled grasping. In *Proceedings of the Human-Robot Interaction Workshop*, March 2014.
- [7] Z. Lin, C. Zhang, W. Wu, and X. Gao. Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs. *IEEE Trans Biomed Eng*, 54(12):1172–1176, 2007.

# Towards a user-friendly BCI for elderly people

Felix Gemblar, Piotr Stawicki and Ivan Volosyak

Faculty of Technology and Bionics, Rhine-Waal University of Applied Sciences, Kleve, Germany  
[ivan.volosyak@hochschule-rhein-waal.de](mailto:ivan.volosyak@hochschule-rhein-waal.de)

## Abstract

In this paper we compare the ease of use, comfort, efficiency and accuracy of two different graphical user interface (GUI) designs for a SSVEP-based BCI application. Main focus is on a user-friendly and intuitive layout design having in mind users of advanced age. The new 3-steps GUI with fewer selection options addresses those factors. Our results show that the majority of subjects favored the 3-steps GUI despite slightly deteriorated information transfer rate.

## 1 Introduction

The effects of ageing present physical handicaps that all-too-often prevent older people from maintaining relationships and actively participating in social life, even when they can live safely and independently in their own homes. Modern communication technologies could help elderly people with physical impairments if they provided special interfaces that worked independently of the person's limitations. For those people there is a way to express their needs, wishes and feelings by using a Brain-Computer Interface (BCI). This technology translates brain signals, usually acquired non-invasively using electroencephalogram (EEG), in computer commands without using the brain's normal output pathways of peripheral nerves and muscles [3].

The use of BCI as a spelling interface has been one of the main focuses in many BCI studies. One of the BCI paradigms used for realization of a spelling interface is the SSVEP BCI which measures the brain responses to a visual stimulation at specific constant frequencies [3]. High system speed is an essential goal in BCI studies. However, especially regarding elderly users, the layout of the graphical user interface plays a key role in ensuring user friendliness and effective control, therefore, the arrangement of the visual stimuli and the desired targets are important. For older people readability and simplicity are crucial, therefore we decreased the number of simultaneously displayed targets to reduce cognitive load and misselections.

## 2 Methods and Materials

### 2.1 Subjects

Seven subjects participated in the study, with a mean age of  $23.6 \pm 3.06$  years (range 21 – 30). Two subjects were female. All subjects were students or employees of the Rhine-Waal University of Applied Sciences. The EEG recording took place in a standard laboratory room with low background noise and luminance. None of the subjects had neurological or visual disorders. Spectacles were worn when appropriate. All persons who volunteered to participate in the study became research subjects after reading the subject information sheet and signing a consent form. Subjects did not receive any financial reward for participating in this study.

## 2.2 Hardware

The subjects were seated in front of a LCD screen (BenQ XL2420T, resolution:  $1920 \times 1080$  pixels, vertical refresh rate: 120 Hz) at a distance of about 60 cm. The used computer system operated on Microsoft Windows 7 Enterprise and was based on an Intel processor (Intel Core i7, 3.40 GHz). Standard Ag/AgCl electrodes were used to acquire the signals from the surface of the scalp. The ground electrode was placed over  $AF_Z$ , the reference electrode over  $C_Z$ , and the eight signal electrodes were placed at predefined locations on the EEG-cap marked with  $P_Z, PO_3, PO_4, O_1, O_2, O_Z, O_9$  and  $O_{10}$  according to the international system of EEG electrode placement. Standard abrasive electrolytic electrode gel was applied between the electrodes and the scalp to bring impedances below  $5 k\Omega$ . An EEG amplifier g.USBamp (Guger Technologies, Graz, Austria) was utilized. The sampling frequency was set to 128 Hz. During the EEG signal acquisition, an analog band pass filter between 2 and 30 Hz and a notch filter around 50 Hz were applied directly in the amplifier.

## 2.3 Software, Experimental setup

Each subject used two different user interface layouts subsequently. The order of the presented GUI's as well as the spelling tasks were randomized for each subject in order to minimize the risk of adaption. The *2-steps GUI* resembles to an earlier developed GUI layout [2]. In the *2-steps GUI* seven commands were represented on the computer screen by flickering boxes of default sizes (125 x 125 pixels). The size of the boxes varied during the experiment as described in [1]. Initially the subject faced seven boxes containing "A-F", "G-L", "M-R", "S-X", "Y-\_" and "Del" in a first window (see Figure 1(a)). After initial selection, the GUI changed to the second window with six individual letters of the initially selected group, so this layout allowed the user to select every letter with only two commands.

In order to improve user friendliness and to decrease the number of frequencies, in the here presented *3-steps GUI* the number of boxes was reduced to four. This also allows us to choose the frequency set more freely. The box for the written word and the word to spell was moved to the centre of the screen. The sizes and appearances of the boxes were the same as in the *2-steps GUI*, however, the number of steps needed to select a letter had increased to three and the position of the boxes changed in every step to increase the user-friendliness. In the first window there were three boxes in the upper part of the screen, arranged horizontally with all the letters divided into three groups "A-I", "J-R" and "S-\_", respectively. The additional 4th box, containing the command "Del" (delete the last spelled character) was located on the right side of the screen. In the second window the position of three boxes changed from upper horizontal to left-hand vertical according to the first selection made in the initial window, e.g. "A B C", "D E F", "G H I" or "J K L", "M N O", "P Q R" or "S T U", "V W X", "Y Z \_". After selecting a desired box in the second window the user faced a third window. Here each box contained a single letter and their positions changed once more from left vertical to horizontal bottom position. In both the second and the third window the far right box ("Del" in window 1) would contain the command "back", giving the user the option to switch to the previous window. The first window as well as an overview of the three steps necessary to choose a single letter are shown in Figure 1(b) and 1(c).

In both GUIs every command classification was also followed by an audio feedback with the name of the selected command or the letter spelled in order to reduce the information load of the visual channel. For SSVEP signal classification we used a minimum energy combination method. The SSVEP classification was performed on-line every 13 samples (ca. 100ms) on the basis of the adaptive time segment length of the acquired EEG data. More details about the

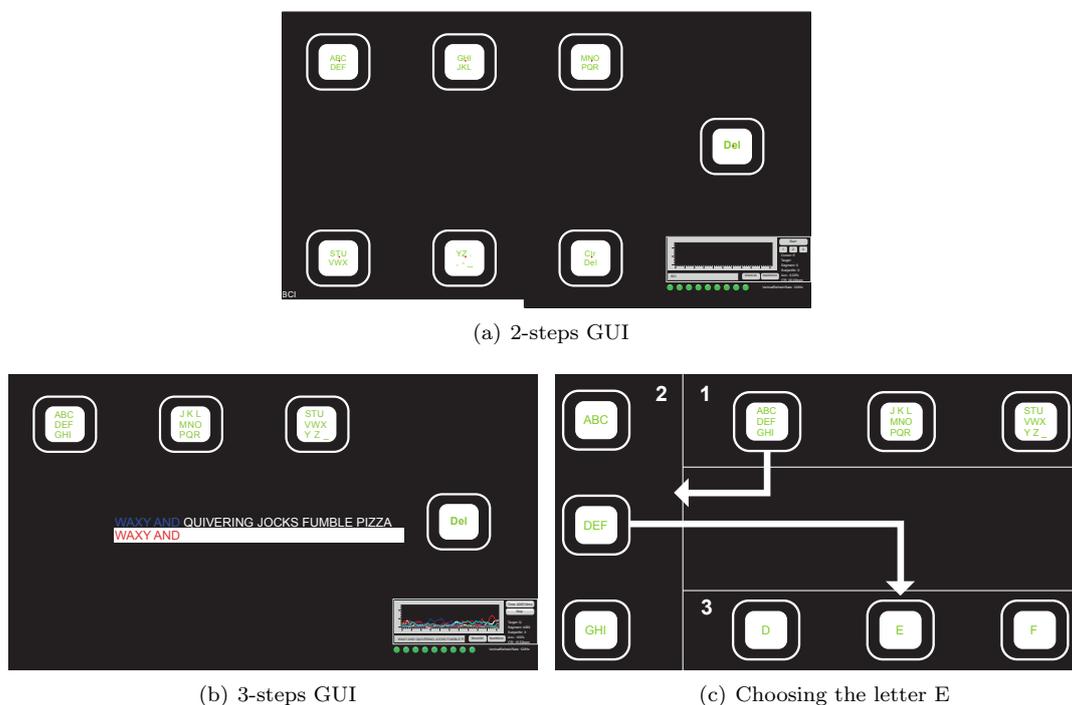


Figure 1: GUI's used in study

used SSVEP detection method can be found in [1].

## 2.4 Procedure

After completing the consent form, each subject completed a brief questionnaire and was prepared for the EEG recording. Subjects participated in a familiarization run spelling words “BCI”, “BRAIN”, and the word of choice (e.g. the own first name) and the experimenters used the collected data to manually adjust the individual parameters. Next, each subject used both GUI's to spell the phrase “WAXY AND QUIVERING JOCKS FUMBLE PIZZA” (commonly used pangram, a sentence using every letter of the alphabet at least once). Each spelling phase ended automatically when the presented word was spelled correctly. The experiment was stopped manually in case a subject could not execute a desired classification within a certain time frame, the subject wished to end the experiment or unintentional repeated misclassifications occurred. Spelling errors were corrected via the implemented Del button. Information needed for the analysis of the test was stored anonymously during the experiment. After finishing the spelling, subjects were asked to rate how well the BCI-System worked for them.

## 3 Results

BCI performance for each subject was evaluated by calculating the commonly used information transfer rate in bit/min, employing the formula as discussed e.g. in [4]. In the here presented GUI's, the overall number of possible choices was equal to seven and four for the *2-steps GUI*

and the *3-steps GUI*, respectively. The accuracy was calculated based on the number of correct command classifications divided by the total number of classified commands. The overall BCI performance is given in Table 1. All subjects were able to complete the pangram. Since for one subject the data were not stored due to a computer error, this table presents the data from only six subjects.

Subject	2-steps GUI			3-steps GUI			Subjective impression	
	Time sec	Acc %	ITR bpm	Time sec	Acc %	ITR bpm	2-steps GUI	3-steps GUI
1	<b>894.66</b>	<b>95.06</b>	<b>13.02</b>	<b>1686.55</b>	92.09	<b>7.28</b>	2	4
2	430.42	81.51	27.19	845.61	85.89	13.71	3	4
3	702.71	72.00	23.59	599.93	85.55	20.27	4	4
4	557.38	70.00	23.53	<b>455.61</b>	<b>76.23</b>	24.34	3	3
5	<b>325.51</b>	92.39	<b>37.69</b>	516.85	88.74	22.97	2	4
6	623.09	<b>69.48</b>	23.20	460.08	<b>95.94</b>	<b>27.08</b>	2	4
<b>Mean</b>	<b>588.96</b>	<b>80.07</b>	<b>24.70</b>	<b>760.77</b>	<b>87.41</b>	<b>19.27</b>	1- worked not at all	
<b>SD</b>	<b>201.54</b>	<b>11.47</b>	<b>7.95</b>	<b>476.06</b>	<b>6.74</b>	<b>7.43</b>	5- worked very well	

Table 1: Results for each subject and layout. Mean values are at the bottom of the table.

## 4 Discussion

Four out of six subjects reached higher ITR with the *2-steps GUI*, also the spelling time was faster for three subjects with this layout. However, the majority of subjects (five out of six) stated that they prefer the *3-steps GUI* over the *2-steps GUI* (right sight of Table 1). Also, some subjects stated that they found the overall design of the *3-steps GUI* more intuitive. All in all, the here presented *3-steps GUI* did not surpass the *2-steps GUI* in terms of spelling time and ITR. But because of other factors, such as user comfort, user-friendly design and better accuracies the *3-steps GUI* could be a valuable option especially for elderly users. Main advantage of the *3-steps GUI* is that only 4 boxes are displayed, which decreases the misselections and seems to be less stressful for the users.

This research was supported by the German Federal Ministry of Education and Research (BMBF) under Grant 16SV6364. We also thank all the participants of this research study.

## References

- [1] Ivan Volosyak. SSVEP-based Bremen-BCI interface – boosting information transfer rates. *J. Neural Eng.*, 8(3):036020, 2011.
- [2] Ivan Volosyak, Anton Moor, and Axel Gräser. A dictionary-driven SSVEP speller with a modified graphical user interface. In *Advances in Computational Intelligence*, pages 353–361. Springer, 2011.
- [3] Jonathan Wolpaw and Elizabeth Winter Wolpaw, editors. *Brain-computer interfaces: principles and practice*. Oxford University Press, 2011.
- [4] J.R. Wolpaw, N. Birbaumer, D.J. McFarland, G. Pfurtscheller, and T.M. Vaughan. Brain-computer interfaces for communication and control. *Clin. Neurophysiol.*, 113:767–791, 2002.

# Prediction of the Saccadic Eye Movement: Using Epidural ECoG in Non-Human Primate

Hoseok Choi<sup>1</sup>, Jeyeon Lee<sup>1</sup>, In Young Kim<sup>1</sup>, Kyung-Ha Ahn<sup>2</sup>,  
Kyoung-Min Lee<sup>2</sup>, and Dong Pyo Jang<sup>1</sup>,

<sup>1</sup>Department of Biomedical Engineering, Hanyang University, Seoul, Korea

<sup>2</sup>Department of Neurology, Seoul National University, Seoul, Korea  
[dongpjang@gmail.com](mailto:dongpjang@gmail.com)

## Abstract

Recently, several studies have reported use of epidural electrocorticogram (eECoG) for brain computer interface (BCI). However, the feasibility and performance of eECoG on BCI were not fully evaluated yet. In this study, we verified the usability of implanted eECoG in non-human primate by predicting saccadic movement using eECoG signals. Two micro electrode patches (32 channels) were inserted over duramater on rhesus monkey. The monkey performed four directional eye movement tasks responding to target's color change. As results, we classify the eye movement directions using eECoG and showed significant and stable decoding performance over two months. This could support the efficacy of BCI using eECoG.

## 1 Introduction

On brain computer interface (BCI), decoding and predicting intention of the subject is one of main issues. In primates, saccadic eye movements are often used to fixate objects of interest. The frontal eye field (FEF), supplementary eye field (SEF), and superior parietal lobule (SPL) including Intraparietal sulcus (IPS) are known as principal neocortical regions involved in the execution of both saccadic and pursuit eye movements. However it is not fully understand because of the limited measurement equipment such as fMRI, which has low time-resolution (LunaB, ThulbornK, StrojwasM, SweeneyJ, 1998) or depth electrode, which has low spatial-resolution (LeeK, AhnK, KellerE, 2012).

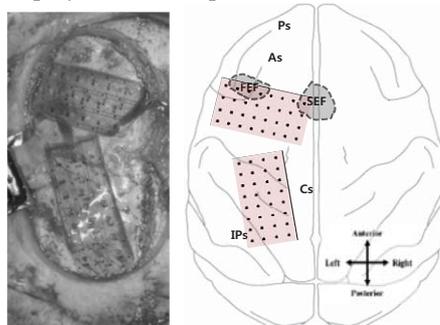
Recently, epidural electrocorticogram (eECoG) is widely known to give a great SNR, high spatial-resolution, and broad frequency bandwidth. In this study, we verified the usability of implanted eECoG in non-human primate by predicting saccadic movement using eECoG signals. We used implanted multi channels micro electrode for non-human primate and eye movement behavioral task for verification.

## 2 Materials and Methods

### 2.1 Surgical Procedure

All surgeries were carried out in the animal surgical suite at the Primate Center of Seoul National University Hospital. Throughout the surgery, body temperature, heart rate, blood pressure, oxygen saturation and respiratory rate were continuously monitored. The adult male rhesus monkey (*Macaca mulatta*) was then placed in a stereotaxic frame before the incision of the scalp. After skin incision, a craniotomy of 2cm radius was conducted. Then, the monkey was implanted epidurally with two micro

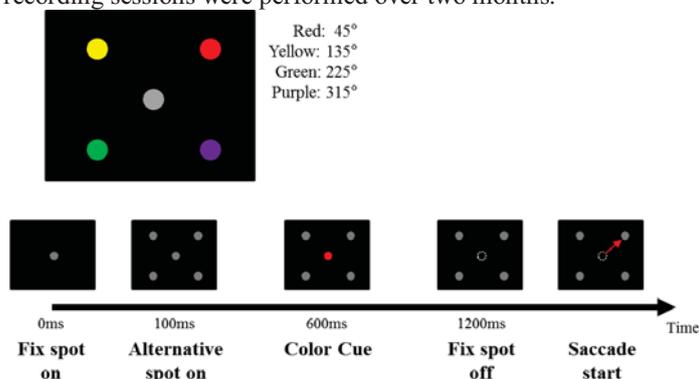
electrode patches (32 channels, 4 by 8), positioned in the left hemisphere including dorsal parietal cortex and posterior part of frontal cortex (Figure 1). It covered the SEF, FEF and SPL including IPS, which were well-known as related to saccadic behavior. The implanted electrode grids consisted of gold electrodes that were embedded in poly-imide and spaced at an inter-electrode distance of 3 mm.



**Figure 1:** Position of the implanted electrode patches

## 2.2 Behavioral Task

Monkey was trained to perform a choice saccade task (LeeK, AhnK, KellerE, 2012). In Figure 2, upper panel shows the pre-trained location and color association, e.g. red is associated with the upper-right visual field. Lower panel depicts a task procedure. A trial began when the monkey fixated at a central gray disc. After the fixation, four gray targets appeared in the peripheral visual field, and after 500ms the central disc changed to one of the four colors which associated with a particular target location. After 600ms, the central disc disappeared, which was the cue for the monkey to make a saccade response. The mapping between color and location was held constant throughout the training and experiments. Electrical recordings were started a week after the surgery, to allow sufficient time for recovery and three recording sessions were performed over two months.



**Figure 2:** "Choice saccade" task paradigm

The performance in the task was monitored by infrared video-oculography with a sampling rate of 500Hz (Eyelink2, SR Research Ltd, Kanata, Ontario, Canada). Saccade behavior was measured off-line using programs written in MATLAB (The Mathworks, Natick, MA, USA). The onset and offset of saccades were determined by velocity criteria (30°/s radial velocity for onset and 10°/s for offset).

eECoG signals were recorded with a sampling rate of 512Hz per channel, using data acquisition system (Brainbox EEG-1164 amplifier, Braintronics B. V., Almere, Netherlands). 60 Hz analog notch filter was used during the data acquisition.

## 2.3 Data Analysis

All data were band-pass filtered from 1 Hz to 170 Hz for processing. For pre-processing, trials with artifacts were excluded by the visual inspection, baseline remove, and the ICA decomposition was conducted to remove artifacts. The analysis epoch was -500 to 0ms relative to saccade execution time. For each channel and each 100ms time period (stepping by 10ms), normalized average power spectral densities were computed in 10 Hz frequency bins using Wavelet. 4 frequency bands were used,  $\alpha$ -band (8~13Hz),  $\beta$ -band (18~26Hz), low  $\gamma$ -band (30~50Hz), and high  $\gamma$ -band (70~170Hz). The powers of each frequency, time window and channel were used as input features (a feature vector) for classification analysis (35 time bins  $\times$  4 frequency bands  $\times$  64 channels; total of 8960 features). The feature vector obtained from each trial was labeled by the target direction (left: 135 degree and 225 degree / right: 45 degree and 315 degree). We used a linear support vector machine (SVM) classifier (BoserB, GuyonI, VapnikV, 1992; VapnikV, 1999) to decode target direction and dimensionality of feature vectors was reduced by adopting SVM-based recursive feature elimination (SVM-RFE) (GuyonI, WestonJ, BarnhillS, VapnikV, 2002). We ranked the features by the weights value and selected top features above +2 standard deviation. The classification accuracy was calculated with a 10 fold cross-validation process.

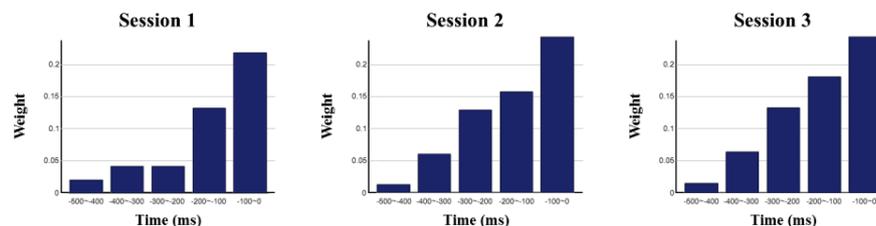
We also quantified the spatio-spectro-temporal contribution of brain activity for predicting each target direction, by calculating weight contribution ratio value from the features' weight magnitude which was derived from adopting the SVM-RFE algorithm (GuyonI, WestonJ, BarnhillS, VapnikV, 2002).

$W(ch)$  quantifies the spatial contribution ratio of each channel for predicting across all frequency bins and time bins,  $W(freq)$  quantifies the spectral contribution of each frequency bin across all recording channels and time lags, and  $W(time)$  quantifies the temporal contribution of each time bin across all channels and frequency bins. (ChaoZ, NagasakaY, FujiiN, 2010)

## 3 Results

We measured the eECoG's impedance to observe the impedance change after implantation. As a result, *in-vivo* impedance test showed stable values as time passes. Electrodes had impedance of  $5.492 \pm 0.046 k\Omega$ ,  $5.470 \pm 0.048 k\Omega$ ,  $5.392 \pm 0.050 k\Omega$  and  $5.358 \pm 0.047 k\Omega$  (mean  $\pm$  SEM for 64 electrodes at 1 kHz), from 1st to 4th week, respectively.

The eECoG decoding processing was carried on three sessions to predict the monkey's saccadic movement. The classification accuracies were 87%, 86% and 89% for predicting saccade eye movement's direction, from 1st session to 3rd session, respectively. Calculating the weight contribution ratio in decoding models revealed that the greater spatial contributions were found in SEF, FEF and SPL, which are consistent with previous reports and support the feasibility of eECoG based on neuroscientific background. In addition, Contribution of temporal information was significantly higher as closer to saccade onset time. The spectral contributions were greater in  $\beta$  (18~26Hz) band than other frequency bands.



**Figure 3:** Time-Weight graph of the each session's  $\beta$  (18~26Hz) band

## 4 Discussion

The 4-direction classification accuracies, which were 53%, 66%, and 56%, were higher than a chance level (25%). The left/right classification accuracies were significantly higher than a chance level (50%). It shows that 2-direction classification is much more efficient than 4-direction classification at BCI. Moreover, the prediction accuracy of eECOG-based decoding showed no significant monotonic decrease for 2 months. This result indicates that eECOG has durability and that the signal quality of the eECOG was maintained for several months at least.

As seen in the Figure 3, the weight of  $\beta$ -band was increased gradually over time and the highest at just before the physical eye movement. It seems that the  $\beta$ -band played a large role in saccadic plan and execution.

In the present study, we predicted saccadic eye movement using eECOG and investigated the feasibility of long term implanted eECOG on BCI. This work demonstrated that it was possible to detect the eye movement's direction before physical eye movement using eECOG, which reflects the possibility of eECOG for brain signal decoding and BCI study.

## Acknowledgments

This research was supported by the National Research Foundation of Korea (NRF) grant funded by the Ministry of Education, Science and Technology (No. 2010-0020807, No. 2013R1A2A2A04014987).

## References

- Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers. *ACM Proceedings of the fifth annual workshop on Computational learning theory*, 144-152.
- Chao, Z. C., Nagasaka, Y., & Fujii, N. (2010). Long-term asynchronous decoding of arm motion using electrocorticographic signals in monkeys. *Frontiers in neuroengineering*, 3.
- Guyon, I., Weston, J., Barnhill, S., & Vapnik, V. (2002). Gene Selection for Cancer Classification using Support Vector Machines. *Machine Learning*, 46(1-3), 389-422.
- Lee, K. M., Ahn, K. H., & Keller, E. L. (2012). Saccade generation by the frontal eye fields in rhesus monkeys is separable from visual detection and bottom-up attention shift. *PLoS One*, 7(6).
- Luna, B., Thulborn, K. R., Strojwas, M. H., & Sweeney, J. A. (1998). Dorsal cortical regions subserving visually guided saccades in humans: an fMRI study. *Cereb Cortex*, 8(1), 40-47.
- Vapnik, V. N. (1999). An overview of statistical learning theory. *IEEE Trans Neural Netw.*, 10(5), 988-999.

# Shared Decision-Making between Clinicians and Patients for Cognitive Rehabilitation after Stroke

Sonja C. Kleih<sup>1</sup>, Ruben G.L. Real<sup>1</sup>, Helena Erlbeck<sup>1</sup>,  
Guilherme Wood<sup>3</sup>, Silvia E. Kober<sup>3</sup>, Claus Vögele<sup>4</sup>, Simone Witzmann<sup>4</sup>,  
Monica Risetti<sup>5</sup>, Donatella Mattia<sup>5</sup> and Andrea Kübler<sup>1</sup>

<sup>1</sup>University of Würzburg, Germany

<sup>2</sup>Center for Clinical Neuropsychology, Würzburg, Germany

<sup>3</sup>University of Graz, Austria

<sup>4</sup>University of Luxembourg, Luxembourg

<sup>5</sup>Fondazione Santa Lucia, IRCCS, Rome, Italy

sonja.kleih@uni-wuerzburg.de

## Abstract

Stroke, one of the leading causes for acquired disability in adults affects motor and cognitive abilities, and greatly increases the risk for depression. In search for new therapies, traditional biofeedback as well as BCI controlled neurofeedback have been suggested to improve emotional well-being and cognitive functions. To provide the best possible therapy for an individual patient, we developed an easy-to-use online interface for shared decision-making between clinicians and patients. After data entry, the system presents a visually attractive review of aggregated and individual test results, relates performance to normative samples, and suggests possible treatment options. This approach to rehabilitation follows the Chronic Care Model that considers patients as partner and fosters their empowerment. Patients, health care providers, and developers of technology work closely together to define problems, set priorities, establish goals and create a training plan. Such a user-centered approach is the pre-requisite for the implementation of bio- and neurofeedback supported rehabilitation in clinical practice.

## 1 Introduction

Stroke is the second leading cause (6.3%) for acquired disability in adults in Europe and affects motor and cognitive abilities. Most prevalent are cognitive deficits in attention and memory, which contribute to reduced quality of life, depression and impairments in activities of daily living (Sheldon & Winocur, 2014). In search for new therapies after stroke we suggest to include bio- and neurofeedback based interventions, which can support patients' convalescence and, thereby, reduce long-term effects after stroke.

Within the EU funded project CONTRAST, we, therefore, developed specific bio- and neurofeedback-based modules for rehabilitation training of patients with sub-acute and chronic stroke. The training modules comprise five different domains: (1) depression, (2) attention, (3) declarative memory, (4) working memory, and (5) inhibitory control. To provide every patient with the appropriate training module and to support the shared decision-making between clinician and patient, an algorithm was developed. In doing so, we followed two major goals: To facilitate shared decision-making by creating a standardized, comfortable and efficient tool for presenting a structured overview of patient

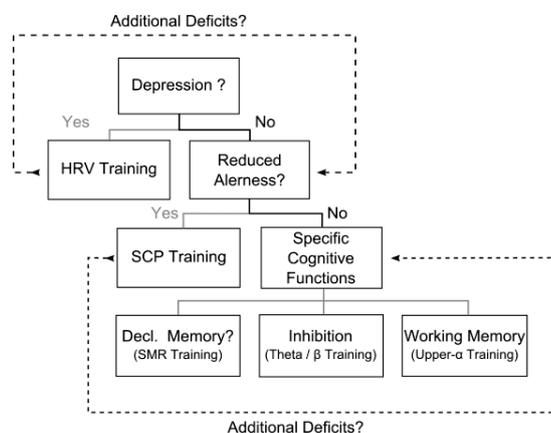
test results. This enables the clinician to explain further steps concerning rehabilitation treatments while patients can follow, comment, discuss and finally co-decide on appropriate interventions. In addition, we aimed to create an individualized hierarchy of treatments on the basis of patients' psychometric test scores.

## 2 Methods

In a first step, a comprehensive profile of possible impairments is obtained from a battery of standardized and well-normed psychometric tests (see Table 1). All included tests fulfilled the quality criteria of being objective, reliable, and valid and are available in German, English, French and Italian. Based on the patient's demographics, such as age, gender, or education, test scores are automatically related to the appropriate test norms, and visualized. In a second step, our system proposes a set of recommendations for rehabilitation training (see Figure 1).

Domain	Measure
Depression	Center for Epidemiologic Studies Depression Scale, Perseverative Thinking Questionnaire
Attention	Test of Attentional Performance (TAP) Alertness, TAP Divided Attention
Declarative memory	California Verbal Learning Test
Working memory	TAP Working Memory, Digit Span Task, Visuo-Spatial Memory
Inhibitory control	TAP Go/No-Go

**Table 1:** Dimensions and respective tests serving as a basis for the algorithm



**Figure 1:** The algorithm including the bio-and neurofeedback modules for cognitive deficits and if necessary, treatment for depression.

We suggest to address depression first, since mood disorders have a ubiquitous negative impact on cognitive training and need to be treated with psychological therapy additionally to HRV biofeedback, then to focus on attentional problems, which are likely to affect other cognitive functions (Lezak, 1987), and finally to address specific cognitive functions such as declarative memory, inhibitory control, and working memory (see Figure 1).

## 2.1 User interface

The evaluation of psychometric test scores is labour intensive and complex. Thus, we provide a user-friendly interface (see Figure 2), which allows for intuitive data entry and retrieval, relieving the therapist from the complexities of test score analysis. After data entry, a comprehensive presentation of the patient’s results in relation to appropriate norms is available with the click of a button.

**Patient: WUE\_P\_001**

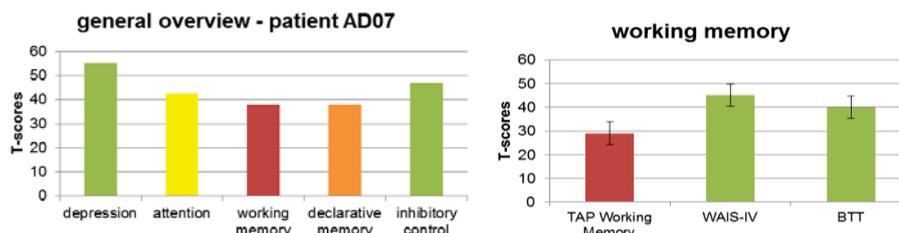
Back

All tests Tests of patient General notes

Name	Cognitiv function	Specification	Date pre	Date post
CVLT (California Verbal Learning Test - Version 1)	Memory	Declarative	9/2/2013	--
CVLT (California Verbal Learning Test - Version 2)	Memory	Declarative	--	9/18/2013
RAVLT (Rey Auditory Verbal Learning Test)	Memory	Declarative	--	--
TAP Alertness	Attention	Alertness	9/2/2013	9/18/2013
TAP Divided Attention (I. auditory)	Attention	Selective/Divided Attention	--	--
TAP Divided Attention (I. visual)	Attention	Selective/Divided Attention	--	--
TAP Divided Attention	Attention	Selective/Divided Attention	9/2/2013	9/18/2013
TAP Visual Field (Neglect)	Attention	Spatial Attention	--	--
TAP Go/NoGo	Attention	Inhibition	9/2/2013	9/18/2013

**Figure 2:** Online overview of test performance scores including date, cognitive function and possibility to request test report.

Graphical presentation of test scores follows a hierarchical approach, where the most critical results and recommendations of the most appropriate training module are summarized on a one-page report. In addition, a consistent colour coding scheme is used to highlight deviant test scores, allowing the rapid identification of critical test results. Figure 3 illustrates the presentation of results from a (fictitious) patient. In this example, “working memory” is easily identified as the most critical domain, with the below-average performance in the TAP working memory test being responsible for this result. To address potential colour vision deficiencies after stroke we plan to provide an additional black and white high-contrast display of results.



**Figure 3:** Colour coded overview of patient AD07 test results. Red=highly critical, orange=critical, yellow=less critical, green=most likely normal.

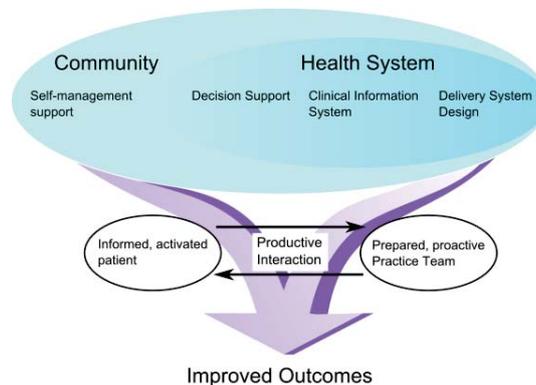
## 2.2 Shared decision-making

Shared decision-making has been suggested as a valuable instrument to increase patients’ autonomy, adherence and commitment to treatment interventions (Godolphin, 2009). In this view, the patient is considered the expert for her or his life while the clinician is considered the medical expert. Thus, shared decision-making is essentially a dialogue between experts with different backgrounds, requiring the sharing of knowledge. By providing accurate yet easy to understand information, our system helps

clinicians and patients to agree on a treatment choice (a training module) as we provide a common basis for this decision with the presentation of tests scores indicating deficits in cognitive functions.

### 3 Conclusion and Outlook

To help clinicians and patients decide on possibly beneficial (neuro-) feedback training for rehabilitation after stroke, we have developed an easy-to-use web-based software suite. Our goal to increase usability for clinicians in a stroke rehabilitation context could be achieved as the algorithm output facilitates shared decision-making between clinicians and patients. Furthermore, it represents a first guideline for integration of biofeedback and BCI based neurofeedback interventions in stroke rehabilitation. This approach to rehabilitation follows the Chronic Care Model (see Figure 4, Wagner et al., 1996) that considers the patient as partner and fosters empowerment of the patients. Patients, health care providers, and developers of technology work closely together to define problems, set priorities, establish goals and create a training plan. Such a user-centred approach is the pre-requisite for the implementation of bio-and neurofeedback supported rehabilitation.



**Figure 4:** The Chronic Care Model (modified from Wagner et al., 1996) requires that Community and Health Systems contribute to well-prepared practice teams and informed active patients who are in a productive interaction allowing for shared decision-making. Improved outcome is demonstrated, for example, by increased patients' self-efficacy, which is associated with low levels of depression and feelings of helplessness.

### References

- Godolphin, W. (2009). Shared decision-making. *Healthcare Quarterly*, 12, (e186-e190).
- Lezak, M. D. (1987). Assessment for rehabilitation planning. In M. J. Meier, A. L. Benton, & L. Diller (Eds.), *Neuropsychological rehabilitation*. New York: Guilford Press.
- Sheldon, S., & Winocur, G. (2014). Memory Loss After Stroke. In T. A. Schweizer & R. L. Macdonald (Eds.), *The Behavioral Consequences of Stroke* (pp. 151–176). New York, NY: Springer New York.
- Wagner EH, Austin BT, Von Korff M. Organizing care for patients with chronic illness. *Milbank Q* 1996;74(4):511-44.

# CoAdapt P300 speller: optimized flashing sequences and online learning

E. Thomas<sup>1</sup>, E. Daucé<sup>2</sup>, D. Devlaminck<sup>1</sup>, L. Mahé<sup>1</sup>, A. Carpentier<sup>3</sup>, R. Munos<sup>3</sup>,  
M. Perrin<sup>4</sup>, E. Maby<sup>4</sup>, J. Mattout<sup>4</sup>, T. Papadopoulo<sup>1</sup> and M. Clerc<sup>1</sup>

<sup>1</sup> INRIA, Sophia Antipolis, France

Maureen.Clerc@inria.fr

<sup>2</sup> Institut de Neurosciences des Systèmes, Inserm UMR 1106, Marseille, France

<sup>3</sup> INRIA, Lille, France

<sup>4</sup> Centre de Recherche en Neurosciences de Lyon, INSERM U1028, France

## Abstract

This paper presents a series of recent improvements made on the P300 speller paradigm in the context of the CoAdapt project. The flashing sequence is elicited by a new design called RIPRAND, in which the flashing rate of elements can be controlled independently of grid cardinality. Element-based evidence accumulation allows early-stopping of the flashes as soon as the symbol has been detected with confidence. No calibration session is necessary, thanks to a mixture-of-experts method which makes the initial predictions. When sufficient data can be buffered, subject-specific spatial and temporal filters are learned, with which the interface seamlessly makes its predictions, and the classifiers are adapted online. This paper, which presents results of three online sessions totalling 26 subjects, is the first to report online performance of a P300 speller with no calibration.

## 1 Material and Methods

The P300 speller presented in this work was implemented in C++ with OpenViBE [7], and a dedicated stimulating software controlled the keyboard display. The software is opensource and part of OpenViBE release 0.18 We used a single Windows laptop to run all software components. The P300 speller keyboard was displayed on a separate LCD screen. A TMSi Refa8 amplifier, synchronized via hardware to the laptop, was used to record from 12 actively shielded electrodes.

The visual stimulations consisted of briefly flashing “smiley” pictures. The P300 wave was detected via 3 channels of an xDAWN spatial filter [8], combined with a Regularized LDA classifier hereforth called RDA, which incorporates a regularisation of the common covariance matrix. The output of the classifier at each flashing time  $t$  is denoted  $\tilde{y}(t)$ .

To save time, elements are always flashed in groups. Initial design of P300 speller groups involved rows and columns of a square matrix [2] or their randomizations [1]. The target element is then found at the intersection of the groups eliciting a P300 response. But repetitively flashing the same groups causes elements within the target groups to be wrongly selected, because of visual attention effects, and because of the contamination of all group elements by classification errors.

**Element-wise evidence accumulation** avoids these two effects. A different random permutation can then be performed at each repetition of the flashes, effectively changing elements’ group membership across repetitions. At each flash  $t$ , let the binary vector  $\mathbf{a}(t)$  represent the set of  $n$  flashed elements within the grid of cardinality  $N$ . The score  $\alpha(t)$  of each element (initialized to 0 at time 0) is updated with the following scheme, in which both target and non-target flashes contribute to the accumulation:

$$\alpha(t) = \alpha(t-1) + \log \left[ \frac{1}{n} \mathbf{a}(t) \tilde{y}(t) + \frac{1}{N-n} (1 - \mathbf{a}(t))(1 - \tilde{y}(t)) \right] \quad (1)$$

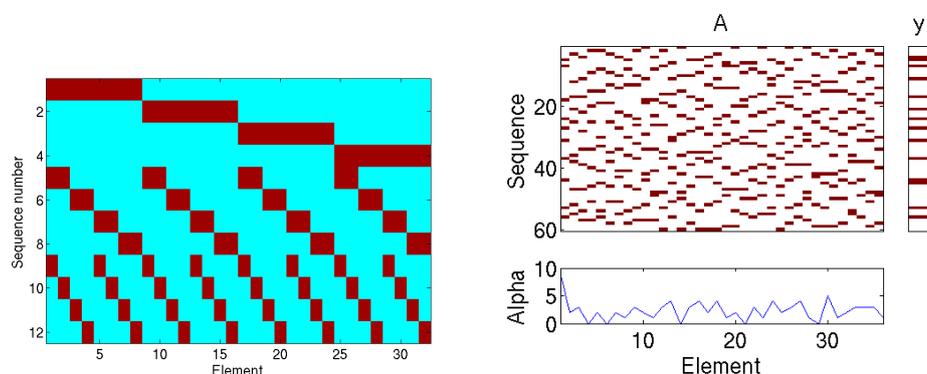


Figure 1: Left: template for a 36-element grid and 1-in-4 flash-rate. Each line represents the flashing elements  $\mathbf{a}(t)$ . Right: example of accumulation based on the randomized grid on the left, over 60 flashes, with element 1 as target.

The groups' definition, encoded in  $\mathbf{a}(t)$ , follows a restricted isometry principle (RIP) [10]. This principle offers an minimal intersection between groups, while allowing the **flash-rate to be adapted** (e.g. 1-in-4, 1-in-6, ...), independently of the number of elements. A random permutation of flashed elements is applied after each repetition (RIPRAND). Figure 1 illustrates the RIP (left) and RIPRAND (right) flashing strategies. Large display grids had already been investigated by [11] and [3], but without optimizing the sequences employed for faster search of the target letter.

Another prominent advantage of element-wise accumulation is that it allows **early stopping** of flashes, as soon as the score of one element clearly outperforms the others. This Early Stopping has been shown to improve user motivation, and in turn, the quality of the P300 signal [6].

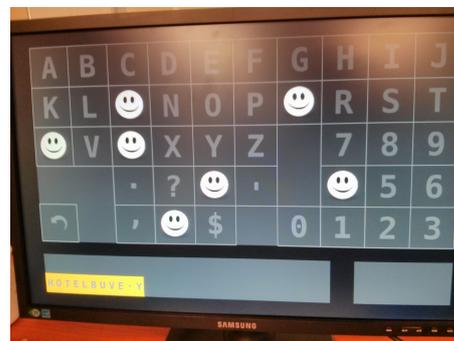
RIPRAND with early stopping was first evaluated offline, using simulated data created by fitting pdfs for P300 detection from a 20-subject database from Inserm Lyon. To test its actual usage, **two online experiments** were performed with 10 subjects each, on a 6x6 grid (**Exp 1**), and on a 9x9 grid (**Exp 2**).

To remove the need for a calibration session, a transfer learning method has been implemented, which uses a “**mixture of experts**” (MOE) classifier learned from the 20-subject data of Experiments 1 and 2. A decision  $\tilde{y}(t)$ , bounded between 0 and 1, is taken by taking the mean of all binary decisions of the experts. This subject-independent decision is then accumulated using (1).

After the P300 speller has been initialized through the above subject-independent approach, and enough data has been collected, a subject-dependent learning takes place, which exploits the notion of target relabelling: decisions made after the accumulation provide labels for supervised learning. xDAWN and RDA parameters are thus learned, and the RDA parameters are subsequently adapted online, by considering a data buffer.

In summary, the MOE classifier, implemented in OpenViBE, thus replaces the classic calibration phase, after which a new individual classifier is calibrated and the experts are ignored. The individual classifier is updated on the data contained in a buffer (a sliding window, which for the moment has a fixed size).

An **online experiment (Exp 3)** was performed with 6 naive subjects to test the zero-calibration and online learning. Experiment 3 used the keyboard layout displayed on the right. The concepts of zero-calibration and online adaptation have been suggested earlier, and validated offline [5], but this is the first time an online study is presented.



## 2 Results and discussion

Table 1 (simulated data) shows the advantages of RIPRAND over row-column flashing in reducing the number of flashes required for an accurate selection. Accuracy refers to character selection accuracy (not target detection accuracy). Table 2 presents results from Online Experiments 1 (6x6 grid) and 2 (9x9 grid), with two types of bit-rate: theoretical bit-rate, and actual bit-rate with a 5 second break between characters. In Experiment 1, one subject was disregarded, as he was unable to stay focussed during the session, and his performance degraded, making comparisons delicate. Finally, Table 3 shows results of the session with no calibration: the session starts with a Mixture-of-experts prediction, which quite high accuracy (86.7%) but requires many flashes (88.7 target + non-target per character) to reach a decision. Bit-rate for this initial period is thus quite low (13.8). In the next part of the session, adaptive learning on the users' own data allows the bit-rate to improve to about 30 bits/minute (real bit-rate).

**Table 1:** Comparison of flashing strategies on simulated data (average over 10000 trials).

Type	Accuracy	Flashes	Theo. Bit-rate
6x6 RIPRAND 1/6	<b>94.5%</b>	<b>34.2</b>	<b>62.8</b>
6x6 Row-Column	93.7%	43.7	47.9
9x9 RIPRAND 1/6	<b>93.9%</b>	<b>42.9</b>	<b>63.8</b>
9x9 Row-Column	87.2%	66.9	37

**Table 2:** Results of Exp 1 (6x6 grid, averaged over 9 subjects) and Exp 2 (9x9 grid, averaged over 10 subjects). Character selection accuracy, average number of flashes per character selection before early stopping, bit-rate accounting for the 5 s pause between characters, and theoretical bit-rate.

Type	Accuracy	Flashes	Bit-rate	Theo. Bit-rate
<b>6x6 Row-Column</b>	92.2%	21.6	30.1	71.3
<b>6x6 RIPRAND 1/6</b>	<b>93.3%</b>	21.6	<b>31.2</b>	72.7
<b>9x9 Row-Column</b>	74.5	53.8	18	30.8
<b>9x9 RIPRAND 1/6</b>	<b>88</b>	42.1	<b>26.5</b>	<b>47.9</b>

**Table 3:** Results of Exp 3 (averaged over 6 subjects). Character selection accuracy, average number of flashes and bit-rate per minute including a 5s pause between characters.

Type	Accuracy	Flashes	Bit-rate
Mixture-of-experts	86.7%	88.7	13.8
<b>Adaptive learning, buffer length 10</b>	<b>95.7%</b>	37.7	27.2
<b>Adaptive learning, buffer length 30</b>	94.6%	<b>27.2</b>	<b>30.8</b>

We have thus developed a new P300 speller paradigm which does not need any calibration session and boosts user motivation thanks to early stopping. A clinical feasibility study is currently taking place at Nice University Hospital on 20 ALS patients. Further improvements considered include the use of Natural Language Models to improve the prediction stage [4, 9] and further improve the bit-rate thanks to word completion.

**Acknowledgment:** This work was carried out as part of the CoAdapt project, funded through the French National Research Agency grant ANR-09-EMER-002-01.

## References

- [1] Hubert Cecotti and Bertrand Rivet. One step beyond rows and columns flashes in the p300 speller: a theoretical description. *International Journal of Bioelectromagnetism*, 2010.
- [2] 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, 1988.
- [3] J. Jin, B.Z. Allison, E.W. Sellers, C. Brunner, P. Horki, X. Wang, and C. Neuper. An adaptive P300-based control system. *Journal of Neural Engineering*, 8(3), June 2011.
- [4] Pieter-Jan Kindermans, Hannes Verschore, David Verstraeten, and Benjamin Schrauwen. A P300 BCI for the masses: Prior information enables instant unsupervised spelling. In *NIPS*, pages 719–727, 2012.
- [5] Pieter-Jan Kindermans, David Verstraeten, and Benjamin Schrauwen. A Bayesian Model for Exploiting Application Constraints to Enable Unsupervised Training of a P300-based BCI. *PLOS One*, 2012.
- [6] Margaux Perrin. *Coadaptation cerveau machine pour une interaction optimale: application au P300-Speller*. PhD thesis, Université Claude Bernard Lyon 1, 2012.
- [7] Yann Renard, Fabien Lotte, Guillaume Gibert, Marco Congedo, Emmanuel Maby, Vincent Delannoy, Olivier Bertrand, and Anatole Lécuyer. Openvibe: An open-source software platform to design, test, and use brain-computer interfaces in real and virtual environments. *Presence: Teleoperators and Virtual Environments*, 19(1):35–53, 2010.
- [8] B. Rivet, A. Souloumiac, V. Attina, and G. Gibert. xDAWN Algorithm to Enhance Evoked Potentials: Application to Brain-Computer Interface. *Biomedical Engineering, IEEE Transactions on*, 56(8):2035 –2043, 2009.
- [9] William Speier, Corey Arnold, Jessica Lu, Ricky K Taira, and Nader Pouratian. Natural language processing with dynamic classification improves p300 speller accuracy and bit rate. *Journal of neural engineering*, 9(1):016004, 2012.
- [10] Eoin M. Thomas, Maureen Clerc, Alexandra Carpentier, Emmanuel Daucé, Dieter Devlaminck, and Rémi Munos. Optimizing P300-speller sequences by RIP-ping groups apart. In *IEEE/EMBS 6th international conference on neural engineering (2013)*, San Diego, United States, November 2013. IEEE/EMBS.
- [11] G. Townsend, B.K. LaPallo, et al. A novel P300-based brain-computer interface stimulus presentation paradigm: Moving beyond rows and columns. *Clinical Neurophysiology*, 121(7):1109 – 1120, 2010.

# A Fast “Single-Stimulus” Brain Switch

Anastasia A. Fedorova<sup>1</sup>, Sergei L. Shishkin<sup>1</sup>, Yuri O. Nuzhdin,<sup>1</sup>  
Marat N. Faskhiev<sup>1,2</sup>, Albina M. Vasilyevskaya<sup>1,3</sup>, Alexei E. Ossadtchi<sup>4,5</sup>,  
Alexander Y. Kaplan<sup>3</sup> and Boris M. Velichkovsky<sup>1</sup>

<sup>1</sup> NBICS Centre, National Research Centre “Kurchatov Institute”, Moscow, Russia  
anastasya.teo@gmail.com, sergshishkin@mail.ru, nuzhdin.urii@gmail.com,  
velich@applied-cognition.org

<sup>2</sup> Faculty of Physics, Lomonosov Moscow State University, Moscow, Russia  
liquidmole@mail.ru

<sup>3</sup> Faculty of Biology, Lomonosov Moscow State University, Moscow, Russia  
albina-vasil@mail.ru, akaplan@mail.ru

<sup>4</sup> Centre for Cognition and Decision Making, National Research University Higher School of  
Economics, Moscow, Russia

<sup>5</sup> St. Petersburg State University, St. Petersburg, Russia  
ossadtchi@gmail.com

## Abstract

We present a novel brain switch that enables sending an urgent command in asynchronous mode at a relatively high speed and accuracy. Five able-bodied participants issued a command to a moving robotic arm in 4 s with only 0.1 false activation per minute. Even higher speed was observed in less demanding condition of an image inspection task, as well as in an offline test with a hybrid BCI design including a saccade detector.

## 1 Introduction

The need for a fast brain switch emerges in the development of brain-computer interface (BCI) control for assistive robotic devices. While low speed and impreciseness of noninvasive BCIs can be, to a large extent, compensated by assigning the low level control to robotic intelligence, the user must be given an opportunity to rapidly stop an action if something goes wrong [3]. A BCI for this function (a fast “brain switch”) should possess the following features: (1) be as fast as possible, (2) be asynchronous, i.e. able to detect the user’s intention at any time moment, (3) have a low false activation rate (FA). An ability to detect a command irrespective of the preceding activity in which the user was involved is also desirable.

Rebsamen et al. [3] proposed a fast asynchronous switch based on a modified P300 BCI. In their design, stimuli were presented in a 3x3 matrix, but a command could only be issued using attention to the central stimulus. Robotic wheelchair was stopped with mean response time (RT) of 6.0 s and FA of 1.2 per min. In [1], the P300 BCI was combined with steady-state visual evoked potential (SSVEP) based BCI, providing RT of 5.3 s and FA of 0.5 per min.

One may ask if the non-target stimuli (8 of 9 stimuli in [3] design) can be removed from the P300 BCI stimulus sequence, to improve the conditions for perceiving the target stimuli and make possible its fast presentation. Although the absence of non-targets contradicts the traditional view on the P300 BCI, we demonstrated that a simplified and more ergonomic variant of the P300 BCI without non-targets can be successfully used for calibrating a classifier for a normal P300 BCI [6]. Using this approach, reminiscent of the “single-stimulus paradigm” known in psychophysiology [2], we designed a new “single-stimulus switch”. In offline simulation using data from four participants, the BCI classifier exceeded the threshold already about 2 s after a saccade to the stimulus, with FA rate of 1.2 to 3.4 per minute [5].



Figure 1: Left: A screenshot from the auditory feedback phase. Right: a view from behind the participant in the robot control phase.

In the current study, the “single-stimulus switch” was tested online for the first time. To demonstrate the effectiveness of the switch in diverse conditions, we tested it in two tasks. In these tasks, the participants were involved into different types of visual activity (still image inspection task and passive pursuing a moving robot arm) from which they had to switch urgently to BCI control at unpredictable time moments.

## 2 Methods

Five healthy volunteers participated in the study. EEG was acquired from Cz, Pz, POz, P2, Oz, O1, O2 locations. We used BCI2000 system [4] with our module for “single-stimulus” presentation and online processing. Stimuli were stylized animal faces presented for 150 ms with interstimulus interval varied from 300 to 550 ms.

For calibration, the participants silently counted stimuli (8 “control” runs, each consisted of 8 blocks of 8 stimuli with 2 s pauses between blocks) and silently read a text (a 10 min “non-control” run). Rebsamen et al. [3] found no difference between reading and other conditions without stimuli counting when they represented the “non-control” condition in estimating classification performance of their BCI in similar settings. We chose the reading task because it is simple and can be well controlled. 512 epochs of 700 ms length were extracted from the “control” EEG starting from the stimulus. From “non-control” EEG, 598 epochs of the same length were extracted with fixed 300 ms interval between them. The EEG was downsampled to 50 Hz and the channel data were concatenated, resulting in feature vector length of 245. Fisher Discriminant Analysis with Tikhonov regularization (ridge regression) was applied to “control” and “non-control” epoch sets to compute the classifier weights.

In the subsequent two phases of BCI online testing, the stimuli (one of the animal faces used in calibration) were presented with the same interstimulus interval, but without 2 s pauses. Each time a stimulus-related EEG epoch was obtained after a new stimulus, the classifier was applied to an average of the last five stimulus-related epochs and its output was compared to a predefined threshold. Once the threshold was exceeded, the BCI was “activated”, i.e. caused actions specific for each testing phase (see below). Data from the first 3 min run (not used in the analysis of the BCI performance) were used to adjust the threshold so that RT and FA were minimized.

In the auditory feedback phase, the background task was to guess what cities were shown on the screen (Figure 1, left). A photo on the screen was changed each 60 s. A sound signal was

presented at random moments. After this “signal event”, the participant had to start counting stimuli until he or she heard a recorded voice saying “Yes” and the stimulus was replaced by a green circle. In the robot control phase, the participants watched a robotic arm (AREXX Mini Robot Arm in the first two experiments and ST Robotics R12-six in the last three experiments) at 1.5 m distance (Figure 1, right). The robot made fast movements (Mini Robot Arm was waving its hand and R12-six was drawing a horizontal line) until it suddenly changed movement direction (a “signal event” in this phase). When this occurred, the participant had to “stop” the robot as fast as possible by counting the stimuli until the green circle appeared and the robot returned to its “normal” behavior.

The BCI activation was considered as correct if it occurred between 750 ms and 8 s after the “signal event”, and as a false activation otherwise. RT was computed for correct activations as their time relative to the preceding “signal events”. FA was computed as the false activation number divided by duration of testing, and miss rate was computed as the proportion of “signal events” not followed by a correct activation. In an additional analysis, a hybrid EEG and EOG based command detection was simulated, with number of averaged EEG epochs decreased from five to three. Large saccades in the direction of the stimulus were detected in EOG using a combined threshold-based and template-based detector and used to confirm activations: BCI threshold exceeding produced a command only if preceded by such a saccade.

### 3 Results and Discussion

In the auditory feedback phase, mean RT was 3.7 s, FA was  $0.3 \text{ min}^{-1}$  and miss rate was 0.1. In the robot control phase, mean RT was 4.0 s, FA rate was  $0.1 \text{ min}^{-1}$  and miss rate was 0.1.

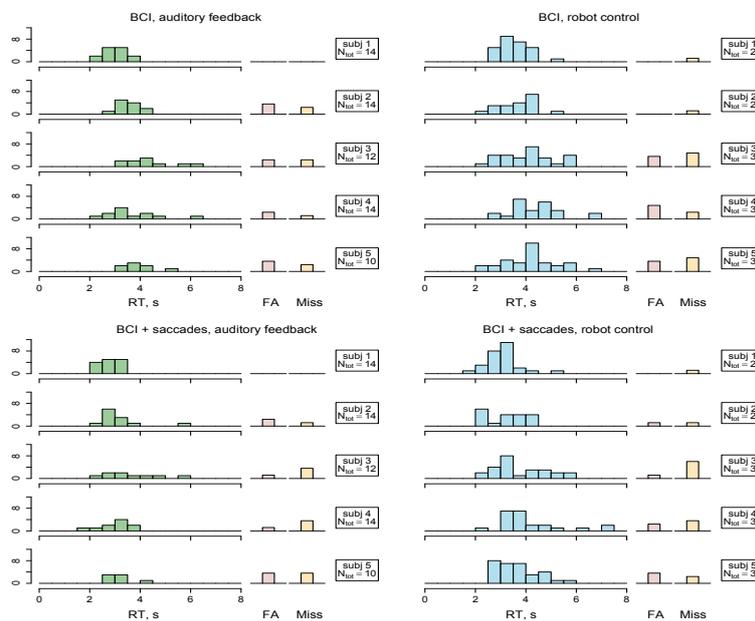


Figure 2: Distribution of Reaction Times (RT), False Alarms (FA) and Misses in online (BCI) and offline (BCI + saccades) tests. Here, FA and Misses data are not normalized. The y-axis shows the number of observations, while  $N_{tot}$  is the total number of “signal events”.

In offline simulation of hybrid EEG-EOG-based command detection, mean RT was reduced to 3.1 s and 3.6 s for the auditory feedback and robot control phases, respectively. FA was  $0.2 \text{ min}^{-1}$  and  $0.1 \text{ min}^{-1}$  in the auditory feedback and robot control phases, respectively. Miss rate was 0.2 and 0.1, respectively (Figure 2, lower row).

This study demonstrated improvement of the BCI switch online performance in issuing an urgent command to a robotic device comparing to the previous online results [3][1]. FA was low even in the image inspection task (the auditory feedback phase), where gaze and attention could be easily attracted by the stimulus positioned near the inspected picture (Figure 1, left).

Further improvement of the initial results obtained in this study can be expected, in our view, by improving the classifier (e.g., by automatic adjustment of the regularization parameter and the threshold) and optimizing parameters of stimulation. Additional possible way for performance improvement can be the use of a committee of classifiers trained on “control” epochs with different overlap between the responses to stimuli and/or several non-control conditions related to different scenarios (e.g., fixating a position close to the stimuli but without attending them). Hybridization of the “single-stimulus” BCI with saccade detection also seems promising, particularly with implementation of advanced saccade detectors and a Bayesian approach to BCI and saccade data fusion.

## Acknowledgements

Part of this study related to the hybrid EEG and EOG based command detection was supported by Russian Science Foundation Grant 14-28-00234.

## References

- [1] Y. Li, J. Pan, F. Wang, and Z. Yu. A hybrid bci system combining p300 and ssvp and its application to wheelchair control. *IEEE Transactions on Biomedical Engineering*, 60(11):3156–3166, 2013.
- [2] J. Polich and M. R. Heine. P300 topography and modality effects from a single-stimulus paradigm. *Psychophysiology*, 33(6):747–752, 1996.
- [3] B. Rebsamen, C. Guan, H. Zhang, C. Wang, C. Teo, V. M. H. Ang, and E. Burdet. A brain controlled wheelchair to navigate in familiar environments. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 18(6):590–598, 2010.
- [4] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw. Bci2000: a general-purpose brain-computer interface (bci) system. *IEEE Transactions on Biomedical Engineering*, 51(6):1034–1043, 2004.
- [5] S. L. Shishkin, A. A. Fedorova, Y. O. Nuzhdin, I. P. Ganin, A. E. Ossadtchi, B. B. Velichkovsky, A. Y. Kaplan, and B. M. Velichkovsky. Towards high-speed eye-brain-computer interfaces: combining the single-stimulus paradigm and gaze translation (in russian). *Herald of Moscow University, Ser. 14: Psychology*, (4):4–19, 2013.
- [6] S. L. Shishkin, A. A. Nikolaev, Y. O. Nuzhdin, A. Y. Zhigalov, I. P. Ganin, and A. Y. Kaplan. Calibration of the p300 bci with the single-stimulus protocol. In *Proceedings of the 5th International BCI Conference*, pages 256–259, 2011.

# Towards BCI Cognitive Stimulation: From Bottlenecks to Opportunities

Begonya Otal<sup>1\*</sup>, Eloisa Vargiu<sup>1†</sup> and Felip Miralles<sup>1</sup>

<sup>1</sup>Barcelona Digital Technology Centre (BDigital), Barcelona, Spain  
botal@bdigital.org, evargiu@bdigital.org, fmiralles@bdigital.org

## Abstract

Emerging BCI based cognitive stimulation applications, like neurofeedback and gaming, are probably among the most growing market opportunities today. Besides, the rising aging population and its associated cognitive decline create a pressing demand for innovative technologies offering a real chance to preserve general wellbeing. Related research breakthroughs that use BCI as main technology to enhance user's performance are examined, as also the common bottlenecks that still hinder BCI end user adoption.

## 1 BCI Evolution

Brain-Computer Interaction (BCI) based technology is thriving and has the potential of spreading into society by addressing the needs of various user groups under different application scenarios. Wolpaw and Wolpaw (2012) identified the five principal BCI scenarios or application types that may have evolved so far, namely, *replace*, *restore*, *improve*, *supplement* and *enhance*. Communication is perhaps the most fulfilling immediate use of BCI systems for a patient, her family and caregivers when no intelligible interaction can otherwise take place (Birbaumer et al., 1999). In this case, the BCI output clearly *replaces* the natural patient's communication function lost as a result of injury or disease. Even simple interactions to make needs known, answer questions with a simple yes or no, and select among a small matrix of choices may reintegrate the isolated patient with others. Similarly, a person may wish to *replace* lost limb function using BCIs as wheelchair controllers in real (Philips et al., 2007) or virtual environments (Leeb et al., 2007), or as appliance adjusters, altering body position in an electric bed for comfort as well as to decrease the chance for developing a bed sore. Additionally, BCIs can be used to operate prosthetic or functional electrical stimulation devices in invasive (Hochberg et al., 2006) or non-invasive (Pfurtscheller et al., 2003) *restoration* of lost natural outputs, such as motor or bladder function in paralyzed humans.

Interestingly, BCIs can now be considered as neurorehabilitation tools to *improve* muscular activation and limb movements in impaired post-stroke patients in clinical settings, for example. Generally speaking, BCIs could help stimulate cortical plasticity leading to the recovery of some lost functions (Carabalona et al., 2009). Following this approach, BCI based cognitive rehabilitation may be among the most outstanding applications that could benefit a large number of patients ranging from completely locked-in patients (Kübler and Birbaumer, 2008) to patients with cognitive impairment, to increasingly *improve* their cognitive deficits. Evidence to support the use of computer based cognitive rehabilitation programs has been growing for the last decade with examples extending to memory (Tam and Man, 2004); working memory (Johansson and Tornmalm, 2012); attention (Zickefoose et al., 2013); and, visual perception (Kang et al., 2009). A more futuristic scenario might be using a BCI

---

\* B. Otal is the corresponding author: botal@bdigital.org

† B. Otal and E. Vargiu contributed equally to this document

to *supplement* a natural neuromuscular output with an additional, artificial (i.e., robotic) output (Wolpaw and Wolpaw, 2012). Last but not least, another area of increasing recent research interest is in the recognition of the user's mental states (e.g., stress or attention levels, "cognitive load" or "mental fatigue") and cognitive processes (e.g., learning or awareness of errors) that will facilitate interaction and stimulate user's interest. In these cases, by preventing stress or attentional lapses, the BCI *enhances* the natural output. Basically, we can deal with neurofeedback or gaming applications to *enhance* user's performance. In line with the *enhance* scenario, and overlapping the *improve* scenario, the rest of the paper concentrates on exploring potential opportunities and research breakthroughs that may use BCI based cognitive stimulation applications to *enhance* people overall performance in clinical and non-clinical settings to maintain general wellbeing and quality of life. Prior to that, we aim at identifying common bottlenecks that prevent current BCI systems from end user adoption.

## 2 Solving BCI Bottlenecks & Shortcomings

The usability of current BCI devices is far from perfect. Medical BCI applications are still very limited, and many critical issues need to be addressed before they can be effectively adopted in clinical settings or in users' homes. Carabalona et al. (2009) have discussed several points that need to be considered when designing, selecting and using a BCI system for neurorehabilitation purposes. In particular, they have emphasized the importance of technology acceptance and usability, as well as, issues related to the impact on the patient's emotional and motivational states. Although some users and assistive technology experts may be quite satisfied with some BCI devices (Zickler et al., 2011), others could not imagine using most of the devices in daily life without further improvements. User-centered design is critical, and testing medical BCI applications with healthy users may be inadequate.

As a whole, medical and other non-medical emerging applications especially addressed to healthy users may encounter similar shortcomings. Main obstacles vary from the long preparation and setup times, to the ergonomics of the electrode caps, and the low speed and lack of reliability of the BCI system. The learning curve is also a major drawback. That is, the user must learn completely new skills to operate a BCI system. Therefore, there is still an important need to overcome the following identified "key factors" (Allison, 2010) for better BCI adoption:

1. *Cost* (financial, help, expertise, training, invasiveness, time, attention, fatigue)
2. *Throughput* (accuracy, speed, latency, effective throughput)
3. *Utility* (support, flexibility, reliability, illiteracy)
4. *Integration* (functional, distraction quotient, hybrid/combined BCIs, usability)
5. *Appearance* (cosmetics, style, media, advertising)

Following a user-centered approach and increasing the engagement with the appropriate end users while designing, selecting and using a BCI system will offer the opportunity to increase technology acceptance and usability. The possibility of using dry electrodes for EEG acquisition, and the monitoring of psychological effects during BCI tasks may give rise to a broader range of cognitive rehabilitation and stimulation programs (Carabalona et al., 2009). According to such view, we claim that the current problems would be overcome when new technologies provide non-conventional sensors for less obtrusive brain signal recording, and affective interfaces able to adapt the BCI according to emotional status changes in the patient.

## 3 Potential Opportunities & Research Breakthroughs

Cognitive decline is one of the most pervasive consequences of aging, and it can be associated with loss of autonomy, functional impairment and deterioration in quality of life in modern societies

(Lee et al., 2013). The increasing elderly population is a potential new target for BCI cognitive stimulation applications, such as neurofeedback and gaming, not only for enhancing motor and cognitive rehabilitation therapies in clinical settings, but especially as cognitive training programs in patient's homes. Development of BCI devices with such cognitive stimulation capabilities will definitely benefit almost all user groups, including many groups of patients or disabled users, and even healthy people in training programs (brain-gym like applications). In cognitive rehabilitation, the use of event-related potentials is traditionally limited to an assessment of injuries incurred or disorder severity. However, a tentative BCI based neurofeedback therapy was designed to treat attention-deficit patients with brain injury (Neshige et al., 1995): five patients with chronic mental disturbances received a BCI based neurofeedback therapy for a four week period, and all showed remarkable improvement. Nevertheless, this pioneering work has not yet been followed by a larger clinical study. New research breakthroughs in BCI cognitive stimulation could profit from the inclusion of "cognitive neuroprosthetics" in invasive BCI research programs, in which the cognitive state of the subject, rather than signals strictly related to motor execution or sensation is recorded (Andersen et al., 2010). This knowledge could be used to capture user's intentions and assess emotional states to keep motivation alive, enhance cognitive and motor recovery more efficiently, and increase overall quality of life in the long term. That is, by stimulating patients to acquire new skills – activating specific cortical areas – BCIs might be used for innovative and more effective occupational therapies with new promising clinical applications. Likewise, BCIs could lead to accurate cognitive trainings and learning programs for the entire population.

Lee et al. (2013) examine the feasibility of using a BCI based system with a new game in improving memory and attention in a pilot sample of healthy elderly. The study investigated the safety, usability and acceptability of the BCI system to elderly, and obtained an efficacy estimate to warrant a phase III trial. In the EU project BackHome (<http://www.backhome-fp7.eu/>), a framework that provides a backdrop to cognitive stimulation tasks performed using a home based BCI system is currently under study. Two serious games – one on memory, and, another one on attention and concentration – were defined and implemented. This is part of a broader telemonitoring home support system (Vargiu et al., 2013) that allows clinicians to prescribe tasks and scales, and monitor its results on quality of life through the use of a remote therapist station. Further, socialization, education, entertainment and even support groups are feasible using BCI via internet email, chats and games (Karim et al., 2006). Thus, in order to ensure equal access to Information and Communication Technologies (ICTs) for all population groups and assure that everyone can benefit from ICT developments, not only people with disabilities, but specially the elderly should not be left behind. New occupational programs aiming at cognitive stimulation tasks to improve ICT skills will turn to be new potential opportunities for BCI based neurofeedback training and innovative serious games.

## Acknowledgments

This work is supported by the European ICT Programme Projects FP7-ICT-2011-7 (BackHome, No. 288566) and FP7-ICT-2013-10 (BNCI Horizon 2020, No. 609593).

## References

- Allison, B. Z. (2010). Toward Ubiquitous BCIs. In B. Graimann, B. Z. Allison, & G. Pfurtscheller (Eds.), *Brain-Computer Interfaces*. Springer Berlin Heidelberg.
- Andersen, R. A., Hwang, E. J., & Mulliken, G. H. (2010). Cognitive neural prosthetics. *Annual review of psychology*, 61, 169-190.

Birbaumer, N., Ghanayim, N., Hinterberger, T., Iversen, I., Kotchoubey, B. (1999). A spelling device for the paralysed, *Nature* 398, 297-298.

Carabalona, R., Castiglioni, P., & Gramatica, F. (2009). Brain-computer interfaces and neurorehabilitation. *Stud. Health Technol. Inform* 145: 160-176.

Hochberg, L.R., Serruya, M.D., Friehs, G.M., Mukand, J.A., & Saleh, M. (2006). Neuronal ensemble control of prosthetic devices by a human with tetraplegia, *Nature* 242, 164-171.

Johansson, B., & Tornmalm, M. (2012). Working memory training for patients with acquired brain injury: effects in daily life. *Scandinavian Journal of Occupational Therapy*, 19, 176-183.

Kang, A.S.H., Kim, D.K., Kyung, M.S., Choi, K.N., Yoo, J.Y., Sung, S.Y., & Park, H.J. (2009). A computerized visual perception rehabilitation programme with interactive computer interface using motion tracking technology – a randomized controlled, single-blinded, pilot clinical trial study. *Clinical Rehabilitation*, 23, 434-446.

Karim, A.A., Hinterberger, T., Richter, J., Mellinger, J., Neumann, N., Flor, H., Kübler, A., & Birbaumer, N. (2006). Neural internet: Web surfing with brain potentials for the completely paralyzed. *Neurorehabil Neural Repair*. 20(4):508-15.

Kübler, A., & Birbaumer, N. (2008). Brain-computer interfaces and communication in paralysis: extinction of goal directed thinking in completely paralysed patients? *Clin Neurophysiol.*, 119(11):2658-66

Lee, T. S., Goh, S. J. A., Quek, S. Y., Phillips, R., Guan, C., Cheung, Y. B., & Krishnan, K. R. R. (2013). A brain-computer interface based cognitive training system for healthy elderly: A randomized control pilot study for usability and preliminary efficacy. *PloS one*, 8(11), e79419.

Leeb, R., Friedman, D., Mueller-Putz, G.R., Scherer, R., Slater M., & Pfurtscheller, G. (2007). Self-paced (asynchronous) BCI control of a wheelchair in virtual environments: A case study with a tetraplegic, *Computational Intelligence and Neuroscience*, 1-8.

Neshige, R., Endou, T., & Miyamoto, T. (1995). Proposal of P300 biofeedback therapy in patients with mental disturbances as cognitive rehabilitation, *The Japanese Journal of Rehabilitation Medicine* 2, 323-329.

Philips, J., del R. Millán, J., Vanacker, G., Lew, E., Galán, F., Ferrez, P.W., Van Brussel, H., & Nuttin, M. (2007). Adaptive shared control of a brain-actuated simulated wheelchair, *Proceedings of the 2007 IEEE 10th International Conference on Rehabilitation Robotics, Noordwijk, The Netherlands*, 408-414.

Pfurtscheller, G., Mueller-Putz, G.R., Pfurtscheller, J., Gerner, H.J., & Rupp, R. (2003). 'Thought' – control of functional electrical stimulation to restore hand grasp in a patient with tetraplegia, *Neuroscience Letters* 351, 33-36.

Tam, S.F., & Man, W.K. (2004). Evaluating computer-assisted memory retraining programmes for people with post-head injury amnesia. *Brain Injury*, 18, 461-470.

Vargiu, E., Fernández, J.M., Torrellas, S., Dauwalder, S., Solà, M., & Miralles, F. (2013). A sensor-based telemonitoring and home support system to improve quality of life through BNCI. In P. Encarnação, L. Azevedo, G.J. Gelderblom, A. Newell, & N.-E. Mathiassen (Eds.), *Assistive Technology: From Research to Practice, AAATE 2013*.

Wolpaw, J., & Wolpaw, E. W. (2012). Brain computer interfaces: Something new under the sun. In J. Wolpaw & E. W. Wolpaw (Eds.), *Brain-Computer Interfaces: Principles and Practice* (pp. 3-14). Oxford: Oxford Univ Press.

Zickefoose, S., Hux, K., Brown, J., & Wulf, K. (2013). Let the games begin: A preliminary study using Attention Process Training-3 and Lumosity™ brain games to remediate attention deficits following traumatic brain injury. *Brain Injury*, 27, 707-716.

Zickler, C., Riccio, A., Leotta, F., Hillian-Tress, S., Halder, S., Holz, E., Staiger-Sälzer, P., Hoogerwerf, E.J., Desideri, L., Mattia, D., & Kübler, A. (2011). A brain-computer interface as input channel for a standard assistive technology software. *Journal of Clinical EEG & Neuroscience*.

# Brain-Computer Interfacing for Stroke Rehabilitation: a Feasibility Study in Hospitalized Patients

Floriana Pichiorri<sup>1</sup>, Giovanni Morone<sup>1</sup>, Iolanda Pisotta<sup>1</sup>, Sonja Kleih<sup>2</sup>, Febo Cincotti<sup>1,3</sup>, Marco Molinari<sup>1</sup>, Stefano Paolucci<sup>1</sup>, Andrea Kübler<sup>2</sup> and Donatella Mattia<sup>1</sup>

<sup>1</sup> Fondazione Santa Lucia, IRCCS, Rome, Italy  
f.pichiorri@hsantalucia.it

<sup>2</sup> Department of Psychology I, Intervention Psychology, University of Würzburg, Würzburg, Germany

<sup>3</sup> Department of Computer, Control, and Management Engineering, University of Rome Sapienza, Rome, Italy

<sup>4</sup> Department of Computer, Control, and Management Engineering, University of Rome Sapienza, Rome, Italy

## Abstract

The feasibility of a Brain Computer Interface (BCI)-assisted motor imagery (MI) training to support hand/arm motor rehabilitation after stroke was evaluated on eight hospitalized stroke patients. The BCI-based intervention was administered as an “add-on” to usual care and lasted 4 weeks. Under the supervision of a therapist, patients were asked to practice MI of their affected hand and received as a feedback the movements of a “virtual” hand superimposed on their own. Following a user-centered design, we assessed system usability in terms of motivation, satisfaction and workload. Usability was also evaluated by professional end-users who participated to a focus group. All patients successfully accomplished the BCI training; significant positive correlations were found between satisfaction and motivation; BCI performance correlated with interest and motivation. Professionals positively acknowledged the opportunity offered by a BCI-assisted training to measure patients adherence to treatment.

## 1 Introduction

During stroke rehabilitation, patients’ involvement rewarded with performance-dependent feedback has been shown to be crucial in improving compliance-adherence [1]. Also evidence exists that the level of patients’ participation in rehabilitation has an impact on the outcome [2]. Several authors have explored the potentialities of Brain Computer Interfaces (BCI) in stroke rehabilitation [3]. To effectively encourage training and practice the BCI design should incorporate principles of current rehabilitative settings, apt to stimulate patients engagement. In this proof of concept study, we report on an electroencephalographic (EEG)-based BCI system intended to support hand motor imagery (MI) training. Our system was designed with the participation of professional users; it included the presence of a therapist to guide the patient during training sessions and it was introduced into a conventional rehabilitation setting. The BCI system was endowed with a visual feedback mimicking movements of patient own hands to maintain consistency with the MI task and it was intended as an add-on tool to enhance hand motor functional recovery of hospitalized patients affected by stroke. The intervention was tested on a small selected hospitalized patient sample, admitted for their rehabilitation treatment after stroke, in order to describe acceptability and usability. In accordance with a

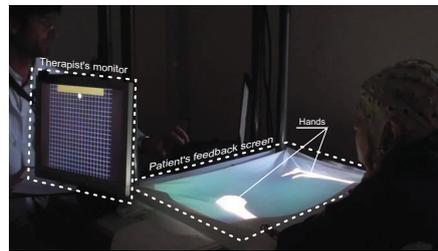


Figure 1: Training session with the BCI tool.

user-centered approach, professional users were requested to participate to the evaluation of the proposed BCI-assisted rehabilitation training.

## 2 Methods

**Participants.** Eight patients with stroke were recruited from a consecutive cohort admitted to Hospital. Outcomes of rehabilitation efficacy were evaluated at baseline and after training: arm-section of the Fugl-Meyer (FM) (minimal clinically important difference - MCID - set to 7 points), National Institute of Health Stroke Scale (NIHSS) and Barthel Index (BI) (MCID set to 14). A descriptive analysis is reported.

**BCI training protocol.** Figure 1 illustrates a training session performed with the proposed EEG-based BCI system. A customized software was implemented to provide patients with a real-time feedback consisting of a visual representation of their own arms and hands. The software allowed the therapists to create an artificial image of a given patient's hand by adjusting a digitally created image in shape, colour and size. The digital image was projected over the patient's real hands covered by a white blanket. To drive the two states of the "virtual hand" (grasping or finger extension) the BCI system exploited the EEG sensorimotor rhythms modulation (SMRs) induced by hand MI of the same movements. The BCI2000 software platform ([www.bci2000.org](http://www.bci2000.org)) was used for real-time estimation and classification of the SMRs state modulation and to drive the instant BCI feedback (i.e. a cursor motion on the therapist's screen) and the corresponding "virtual hand" action (actuated through a UDP connection). The BCI-based device was installed in a rehabilitation hospital ward. Training lasted 4 weeks, with 3 weekly sessions. The EEG features to control the cursor motion during the BCI training were extracted from a screening session during which patients were asked to perform MI of their own affected hand in the absence of feedback. Performances were expressed as mean percentage of correct trials per run; the second and last training sessions were considered for statistical analysis, conducted by means of a t-test for dependent samples. Scalp EEG potentials during the screening session were collected from 61 positions (according to an extension of the 10-20 International System) bandpass filtered between 0.1 and 70Hz, digitized at 200Hz and amplified by a commercial EEG system (BrainAmp, Brainproducts GmbH, Germany). During the BCI-training, EEG was recorded from a subset of 31 electrodes distributed over the scalp centro-parietal regions.

**Acceptability and usability assessment.** Acceptability and usability were explored by means of participants mood, motivation and satisfaction assessment and participants perceived

Pt	Sex	Age (yrs)	Time from event (wks)	Lesion Side	Lesion Type	Fugl Mayer Arm Section		NIHSS		Barthel Index		Control features	
						pre	post	pre	post	pre	post	channel	frequency (Hz)
P1	M	58	32	R	I	11	14	9	8	65	70	Cpz, Cp2	14-16
P2	M	72	20	L	I	13	17	10	7	45	85*	Cp3, Cp1	16-18, 22-24
P3	M	58	44	R	H	13	15	8	8	90	90	C4, Cp4	10-12
P4	M	41	65	R	H	9	17*	5	5	45	70*	C2, C4, Cp4	22-24
P5	F	75	6	L	I	31	58*	12	5	40	55*	C3, C5	16-18
P6	M	52	9	R	I	10	17*	7	5	75	85	C2, Cp2	18-20
P7	M	58	7	L	I	7	11	12	8	50	70*	Cpz, Cp1	22-24
P8	F	66	12	R	I	17	37*	9	6	45	55	Cz, Cp4, Cp6	14-16
<b>Mean</b>						<b>13,9</b>	<b>23,3</b>	<b>9,0</b>	<b>6,5</b>	<b>56,9</b>	<b>70</b>		
<b>SD</b>						<b>7,1</b>	<b>16,1</b>	<b>2,4</b>	<b>1,4</b>	<b>17,9</b>	<b>12,5</b>		

Figure 2: Clinical characteristics and training outcome of patients. Arm section of the Fugl-Mayer scale ranges from 0 (most affected) to 66 (normal); NIHSS ranges from 0 (normal) to 42 (most affected); Barthel Index ranges from 0 (most affected) to 100 (normal). The asterisk indicates achievement of the Minimally Clinical Important Difference.

workload. Usability was also assessed by professional users by means of a focus group setting. Before starting each training session, patients mood and motivation was monitored by means of visual analogue scales (VAS). Mood was also assessed by means of the Center of Epidemiologic Studies Depression Scale (CES-D) scale, administered once a week across the 4 weeks of training. Motivation was also assessed by means of an adapted version of the Questionnaire for Current Motivation (QCM) which was administered at the end of each training session. Satisfaction was reported by users by means of a VAS. The NASA-Task Load Index (NASA-TLX) was administered at the end of the first and last training session as a measure of workload. The evaluation of the proposed BCI-based rehabilitation approach was also addressed with professional users identified as therapists in the context of a focus group. Fifteen therapists attended a training session with the participation of one stroke patient. Professionals users were administered a slightly modified version of the The Quebec User Evaluation of Satisfaction with Assistive Technology 2.0 (QUEST2.0), and an open discussion was held. The Spearmans coefficient was applied to explore separately the correlation between BCI performance and each psychological variable and between psychological variables.

### 3 Results

Figure 2 depicts clinical and BCI-control features for each patient. All patients succeeded in controlling the “virtual hand” by practicing affected hand MI, with mean performances of  $57 \pm 24\%$  (n=8 patients; 73 training sessions; chance level of 5%). No significant change in performance was found between second ( $62,3 \pm 20,4\%$ ) and last ( $49,1 \pm 19,1\%$ ) BCI training sessions ( $p > .05$ ). Patients score at CES-D scale was  $6.86 \pm 4.8$  on average for all patients and determinations. Patients rated their mood as good (VAS  $7.15 \pm 1.91$ , average of all patients/sessions) and were highly motivated during training (VAS  $7.70 \pm 1.90$ , average of all patients/sessions); satisfaction was also high (VAS  $8.36 \pm 1.65$ , average of all patients/sessions).

VAS mood and motivation scores positively correlated ( $p=.00, r=.479$ ) and satisfaction was also positively correlated with motivation ( $p=.001, r=.393$ ). The “Interest” QCM factor and the BCI performance percentage were positively correlated ( $p=.027, r=.257$ ) as well as the BCI performance and VAS motivation scores ( $p=.012, r=.289$ ). Analysis of the NASA-TLX questionnaires revealed no significant differences in perceived workload obtained at the end of the first and last training sessions (NASA-TLX:  $50.22 \pm 21.73$  end of first session;  $54.17 \pm 23$  end of last session). According to the QUEST2.0 results, all therapists (total  $n=15$ ) stated as the most important system features: “Effectiveness”, “Ease to Use”, “Learnability”, “Safety”, and “Reliability”. Several strengths and weaknesses of the BCI-assisted MI training design emerged from the open discussion. Professionals identified as a strength point the potentiality of such BCI-based system to provide them with a quantitative measure of the patients adherence to a cognitive-motor rehabilitation session (i.e. SMR modulation induced by MI). They considered as most relevant weaknesses: i) prototype setup and functioning (hardware and software) which requires technical skills that “a therapist might not have” and they would not feel confident in being able to carry out a session without some technical assistance (cap and electrodes adjusting, EEG signal monitoring, software operating); ii) the lack of a “goal-directed action” feedback (e.g. holding and releasing a glass of water); iii) the need to monitor a possible increase in arm spasticity during MI task practice.

## 4 Discussion

All patients were highly motivated supporting the idea that the specific BCI training was positively accepted with a good compliance/adherence. In particular, motivation was maintained across training sessions and correlated with satisfaction. Encouragingly, we also found no significant self-rated workload differences across training. Our BCI-assisted motor imagery training has proven to be tolerable and acceptable by patients and although usability still requires some improvement, also professional users are prone to accept such technology when added to standard motor rehabilitation during hospitalization. Our preliminary data suggest that BCI-technology might successfully be adopted to support the practice of MI tasks and thus positively influence recovery outcome in stroke patients. A clinical trial with a large cohort of patients is needed to establish the extent to which any clinical improvement might be imputed to the BCI, and to confirm the current positive results on the acceptability of the system by patients.

**Acknowledgements.** This work was partially supported by the Italian Ministry of Healthcare (grant:RF-2010-2319611).

## References

- [1] M C Cirstea and M F Levin. Improvement of arm movement patterns and endpoint control depends on type of feedback during practice in stroke survivors. *Neurorehabilitation and neural repair*, 21(5):398–411, October 2007. PMID: 17369514.
- [2] S Paolucci, A Di Vita, R Massicci, M Traballese, I Bureca, A Matano, M Iosa, and C Guariglia. Impact of participation on rehabilitation results: a multivariate study. *European journal of physical and rehabilitation medicine*, 48(3):455–466, September 2012. PMID: 22522435.
- [3] S Silvoni, A Ramos-Murguialday, M Cavinato, C Volpato, G Cisotto, A Turolla, F Piccione, and N Birbaumer. Brain-computer interface in stroke: a review of progress. *Clinical EEG and neuroscience*, 42(4):245–252, October 2011. PMID: 22208122.

# Steady-State Somatosensory Evoked Potentials in Minimally Conscious Patients – Challenges and Perspectives

C. Pokorny<sup>1</sup>, G. Pichler<sup>2</sup>, D. Lesenfants<sup>3</sup>, Q. Noirhomme<sup>3</sup>, S. Laureys<sup>3</sup> and  
G. R. Müller-Putz<sup>1</sup>

<sup>1</sup> Graz University of Technology, Institute for Knowledge Discovery, 8010 Graz, Austria  
christoph.pokorny@tugraz.at, gernot.mueller@tugraz.at

<sup>2</sup> Albert Schweitzer Clinic, Department of Neurology, 8020 Graz, Austria  
gerald.pichler@stadt.graz.at

<sup>3</sup> Coma Science Group, Cyclotron Research Centre, University of Liège, 4000 Liège, Belgium  
damien.lesenfants@doct.ulg.ac.be, quentin.noirhomme@ulg.ac.be, steven.laureys@ulg.ac.be

## Abstract

In the present study, we aimed to detect the "resonance-like" frequencies of the somatosensory system in patients in a minimally conscious state using a screening paradigm. EEG measurements were conducted in seven patients during tactile stimulation of their left and right wrist. A significant tuning curve could be found in one of the patients. Various reasons that could explain the inconclusive outcome of most measurements, as well as future perspectives are discussed.

## 1 Introduction

A brain-computer interface (BCI) based on electroencephalography (EEG) can provide severely brain-injured people with a new output channel for communication and control [8]. BCIs may also be used as an objective and motor-independent diagnostic tool for patients with disorders of consciousness (see [1] for a review). For patients with impaired hearing or vision, BCIs based on tactile stimuli could be one possible alternative since the somatosensory system is expected to remain functional [4]. By repeatedly applying tactile stimuli with a sufficiently high rate, steady-state somatosensory evoked potentials (SSSEPs) can be evoked and measured using EEG [7]. SSSEPs can intentionally be modulated by attention [2] and, therefore, are one possible way to realize a tactile BCI [4].

As a first step to realize such an SSSEP-based BCI in patients with severe neurological diseases or brain injuries, the "resonance-like" frequencies, i.e. the frequencies with the highest SSSEP response of the somatosensory system [3] need to be identified. Within our work, a well-established screening paradigm was adapted for this purpose to be applied to patients in a minimally conscious state (MCS), i.e. to patients showing non-reflexive behavior but being unable to communicate. Challenges, problems, and results of this attempt are presented. Possible improvements and reasons why the results are not as promising as expected are discussed.

## 2 Materials and Methods

### 2.1 Screening Paradigm

Two C-2 tactors (Engineering Acoustics, Inc., USA) were attached to the left and right volar wrist using elastic wrist bands. The wrists were stimulated with seven frequencies ranging from

14 to 32 Hz (3 Hz steps). A modulated stimulation pattern (200 Hz sine carrier), generated by a self-made, medically approved stimulation device [5], was used.

Each trial started with a 2.5 s reference interval without stimulation, followed by seven 2 s stimulation intervals with frequency and wrist randomly chosen (without using the same frequency and wrist twice in a row). To avoid attentional modulation effects of the SSSEPs, relaxing music was presented via headphones to distract the participants. The whole paradigm lasted around 40 minutes and consisted of 40 repetitions per frequency and wrist.

The EEG was recorded with two g.USBamps (g.tec medical engineering GmbH, Austria) using 32 active electrodes. The reference electrode was connected to the left earlobe, the ground electrode to the right mastoid. Bipolar channels were derived at three frontal, seven central, and four parietal positions (international 10-20 system). Tuning curves showing the percentage band power increase of the stimulation intervals relative to the reference intervals [3] were computed. For statistical validation, 95 % confidence intervals were estimated by bootstrapping using 1000 bootstrap samples.

## 2.2 Participants

Seven patients in an MCS participated in this study (one or two sessions) at the Albert Schweitzer Clinic (Graz, Austria) and the Liège University Hospital (Liège, Belgium). The patients were either sitting in a wheelchair or lying in bed with the upper part of their body slightly elevated. Before or after each EEG measurement, the patients were behaviorally assessed using the Coma Recovery Scale-Revised (CRS-R). Table 1 provides clinical and demographic data together with the CRS-R scores of all patients. Informed consent was obtained from the patients' legal representatives. The study was approved by the Ethics Committees at the participating institutions and was conducted in accordance with the Declaration of Helsinki.

Patient no.	Location	Age (years)	Sex	Etiology	CRS-R	
					s1	s2
PA <sub>01</sub>	Graz	28	male	Traumatic	9	11
PA <sub>02</sub>	Graz	58	female	Anoxia	8	10
PA <sub>03</sub>	Graz	67	male	Traumatic	17	17
PA <sub>04</sub>	Liège	22	male	Traumatic	6	–
PA <sub>05</sub>	Liège	15	male	Hemorrhagic stroke	15	–
PA <sub>06</sub>	Liège	51	female	Hemorrhagic stroke	4	–
PA <sub>07</sub>	Liège	45	female	Traumatic	7	–

Table 1: Clinical and demographic data of the patients, together with the CRS-R scores of the first (s1) and, where applicable, second (s2) session.

## 3 Results

Fig. 1 shows the SSSEP screening results of all patients and sessions from three representative EEG channels contralateral to the stimulated wrist. Only in one patient, PA<sub>05</sub>, a significant tuning curve could be found for right wrist stimulation at the bipolar channel F3-C3. The frequency with the highest relative bandpower increase (140 %) was found to be 20 Hz. In all other patients, no significant tuning curves were found at any of the channels contra- or ipsilateral to the stimulated wrist. To demonstrate that the screening paradigm is suitable

to identify the individual "resonance-like" frequencies, the results of a healthy control were included (same tactor location; reduced channel set only), showing high tuning curve peaks at 23 Hz for left (373 %) and right (363 %) wrist stimulation.

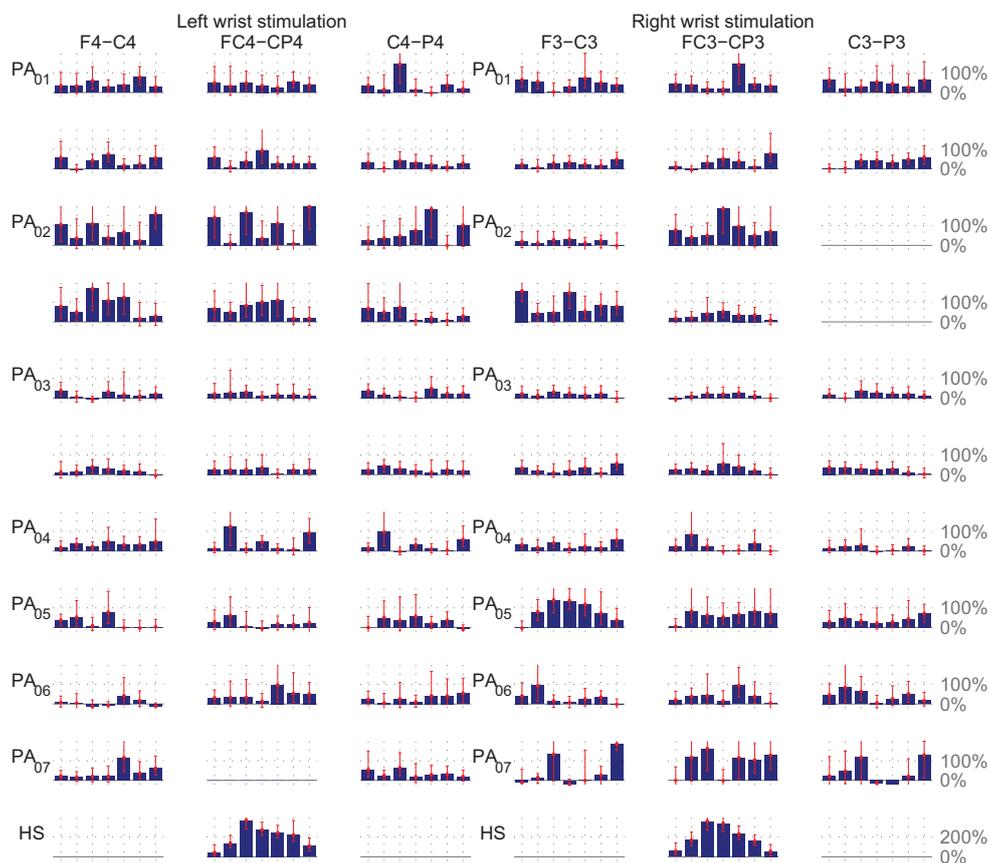


Figure 1: Screening results of all patients and sessions (rows) from three representative (bipolar) EEG channels contralateral to the stimulated wrist (columns). The bars show the relative bandpower increase (in %) with 95 % confidence intervals of all seven stimulation frequencies. The last row shows the results of a healthy subject (HS), using a different y axis scaling.

## 4 Discussion

Within this work, a screening paradigm was developed with regard to the specific needs and capabilities of patients in an MCS. The wrists were selected as target location, since some of the patients suffered from hand spasticities, making it not easily possible to use more sensitive locations like finger tips. Screening results obtained from a healthy control were totally in accordance with literature (e.g. [3]). However, only in one of the seven patients, a significant tuning curve could be found. In all other patients, stable SSSEPs were not present. In some patients, an increase in band power of only certain single frequencies could be found. However,

it is not yet known if perhaps such frequencies could intentionally be modulated and thus be sufficient to realize a BCI. While technical problems seem unlikely (as shown by the control experiment), various other reasons could explain the inconclusive outcome of most patient measurements. First, uncontrolled body movements of the patients resulted in a huge amount of biological (EOG, EMG) and technical (cable movements, electrodes touching the pillow) artifacts. Even though trials containing strong artifacts were manually removed, outliers and huge confidence intervals were still present in the screening results. Second, it was not clear if the position and contact pressure of the tactors allowed the patients to perceive the stimuli strong enough at all, as they could not be simply asked about their perception of the stimuli. Spasticities may have also had a severe influence on the SSSEPs, since the tendons of the finger flexors are located at the volar side of the hand. Third, maybe SSSEPs were not present because of an impaired somatosensory system, or could simply not be measured with EEG due to alterations in the brain topology. Interestingly, the one patient showing significant results was a stroke survivor with a CRS-R score of 15. In comparison to the others, this patient had a high score and no traumatic injury. This could be evidence that the structures in his brain were not that damaged and therefore SSSEPs could be measured.

Similar difficulties regarding a paradigm transition from healthy subjects to patients in an MCS were already reported in [6]. In future, better artifact avoidance or rejection methods, longer stimulation intervals, or other target body locations could be beneficial. Moreover, a thorough neurophysiological examination prior to SSSEP measurements may be helpful.

## 5 Acknowledgments

This work is supported by the European ICT Programme Project FP7-247919. The text reflects solely the views of its authors. The European Commission is not liable for any use that may be made of the information contained therein.

## References

- [1] C. Chatelle, S. Chennu, Q. Noirhomme, D. Cruse, A. M. Owen, and S. Laureys. Brain-computer interfacing in disorders of consciousness. *Brain Inj*, 26:1510–1522, 2012.
- [2] C.-M. Giabbiconi, C. Dancer, R. Zopf, T. Gruber, and M. M. Müller. Selective spatial attention to left or right hand flutter sensation modulates the steady-state somatosensory evoked potential. *Cogn Brain Res*, 20:58–66, 2004.
- [3] G. R. Müller, C. Neuper, and G. Pfurtscheller. Resonance-like frequencies of sensorimotor areas evoked by repetitive tactile stimulation. *Biomed Tech*, 46:186–190, 2001.
- [4] G. R. Müller-Putz, R. Scherer, C. Neuper, and G. Pfurtscheller. Steady-state somatosensory evoked potentials: Suitable brain signals for brain-computer interfaces? *IEEE Trans Neural Syst Rehabil Eng*, 14:30–37, 2006.
- [5] C. Pokorny, C. Breitwieser, and G. R. Müller-Putz. A tactile stimulation device for EEG measurements in clinical use. *IEEE Trans Biomed Circuits Syst*, (in press), 2013.
- [6] C. Pokorny, D. S. Klobassa, G. Pichler, H. Erlbeck, R. G. L. Real, A. Kübler, et al. The auditory P300-based single-switch brain-computer interface: Paradigm transition from healthy subjects to minimally conscious patients. *Artif Intell Med*, 59:81–90, 2013.
- [7] D. Regan. *Human Brain Electrophysiology: Evoked Potentials and Evoked Magnetic Fields in Science and Medicine*. Elsevier, New York, 1989.
- [8] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan. Brain-computer interfaces for communication and control. *Clin Neurophysiol*, 113:767–791, 2002.

# Reconstructing gait cycle patterns from non-invasive recorded low gamma modulations

Martin Seeber<sup>1</sup>, Johanna Wagner<sup>1</sup>, Reinhold Scherer<sup>1,2</sup>, Teodoro Solis-Escalante<sup>3</sup> and Gernot R. Müller-Putz<sup>1</sup>

<sup>1</sup> Graz University of Technology, Institute of Knowledge Discovery, Graz, Austria  
[seeber@tugraz.at](mailto:seeber@tugraz.at)

<sup>2</sup> Rehabilitation Clinic Judendorf-Strassengel, Judendorf-Strassengel, Austria

<sup>3</sup> Delft University of Technology, Department of Biomechanical Engineering, Delft, Netherlands

## Abstract

This work presents a simple to set-up system for reconstructing gait cycle patterns from non-invasive recorded electroencephalographic (EEG) signals. It is based on the prior finding that low gamma amplitudes are modulated locked to the gait cycle in central sensorimotor areas. Therefore, we focused on a Laplacian Cz derivation and low gamma amplitude modulations to reconstruct the gait patterns. Our results show that this method was successful in reconstructing gait cycle patterns in 8/10 subjects during active walking and in every subject during passive walking. The median reconstruction error was  $0.24 \pm 0.13$  s for active and  $0.26 \pm 0.10$  s for passive walking. The presented methods and findings are a further step towards analysing and monitoring ongoing cortical activity during human upright walking.

## 1 Introduction

Reconstructing gait cycle patterns from brain activity could be a useful technique, e.g., first, to enable neuromonitoring for therapists, and second, to provide neurofeedback for patients during rehabilitation after brain injury. Using reconstructed gait related brain patterns for neurofeedback can potentially enhance training in individuals with impaired motor function. These signals could, e.g. be used for supportive control of functional electric stimulation or robotic gait orthosis. Recently, some efforts in gait kinematics reconstruction were made [4] using frequencies in the delta range (0.1-2Hz) for decoding. However, the usage of low frequency components for decoding movements was discussed controversially [1] and their cortical origin were doubted [2]. Previous work from our group [5], [6] showed that low gamma (24-40 Hz) amplitudes are modulated in relation to the gait cycle. These activities were localized in central sensorimotor regions. In this work, we suggest that these low gamma modulations can be recognized on single-trial level and thus enables the reconstruction of gait cycle patterns in real-time.

## 2 Methods

### 2.1 Experiment

Ten able-bodied subjects (S1-S10, 5 female, 5 male,  $25.6 \pm 3.5$  years) walked with a robotic gait orthosis (Lokomat, Hocoma, Switzerland). The experiment involved 4 runs (6 min each) of active and passive walking as well as 3 runs of upright standing (3 min each), which was used as a control condition. The walking speed was adjusted for every subject ranging from 1.8 until 2.2 km/h, and a body weight support of less than 30% was provided.

## 2.2 Recordings and Preprocessing

EEG was recorded from 120 sites with four 32-channel amplifiers (BrainAmp, Brainproducts, Munich, Germany). The sampling frequency was 2.5 kHz, filter settings were 0.1 Hz and 1 kHz for high and low pass respectively. The electrodes were arranged in accordance to the 5% international 10/20 EEG system (EasyCap, Germany). Left and right mastoids were chosen to place reference and ground electrodes. Electrode impedances were lower than 10 k $\Omega$ . The interval between two right leg heel ground contacts defined one gait cycle. Foot contact was measured by mechanical switches. Further details can be found elsewhere [6]. EEG data was filtered and down sampled to 500 Hz. Channels with a variance greater than 5 times the median variance were rejected from further analysis. If one channel exceeded a threshold of 200  $\mu$ V, the according gait cycle epoch was removed from the EEG recordings. In total 94.1 % of the data were used for the analysis.

## 2.3 Reconstruction of the gait cycle

We built a model to reconstruct the gait cycle pattern from EEG recordings, based on physiological evidence; namely that the low gamma amplitudes are modulated with the step frequency in the sensorimotor feet area. To reconstruct the gait cycle pattern from EEG we consequently focused on the temporal modulations of low gamma (20-40Hz) amplitudes. To test a system that is in principle able to work in clinical practice we used five channels (FCz, C1, Cz, C2, CPz) to calculate Laplacian Cz derivation in this work, since we already know the spatial region of interest from previous work [5], [6].

The EEG data were split into two subsets. First, one third of randomly drawn EEG trials (max. 8 min) were used as training set to simulate clinically feasible system training durations. Second, the unseen two thirds of the data were concatenated and used for evaluating the gait cycle reconstruction. The training set was used to identify the parameters carrier frequency, mean step frequency and phase lag between low gamma modulations and actual foot contacts. Spectral analyses were performed using complex Morlet wavelets [3] (center frequency: 1 Hz, full width half maximum: 3). Low gamma amplitudes are a result of spectral analysis. To determine the frequency of an amplitude modulation (AM) these amplitudes were again wavelet transformed (center frequency: 1 Hz, full width half maximum: 6).

The frequency in the low gamma range which amplitude were maximally modulated by the step frequency was used as optimal carrier frequency. The phase lag between low gamma AM and the actual foot contact was determined using the training data. Once we know the optimal carrier frequency for AM and the phase relation between the AM and the actual foot contacts; it is straightforward to reconstruct the foot contacts and thus the gait cycle patterns from EEG recordings.

Data of the evaluation set were analysed on single-trial level to simulate a real-time experiment. The EEG data were wavelet transformed at the optimal carrier frequency and the resulting amplitude were again analysed using a wavelet with the step (modulation) frequency to gain the phase of ongoing low gamma AM. The phase lag gained from the training set was then used to estimate the foot triggers. Reconstructed and mechanical measured foot contacts were then compared and errors were calculated as L1 norm. To test if these errors are significantly smaller than chance, we performed the same analysis using EEG data from the standing condition to evaluate the chance level errors. The reconstruction errors for active and passive walking were then statistically tested to be smaller than these chance level errors using the Wilcoxon rank sum test.

### 3 Results

In 8/10 subjects the reconstruction of the gait cycle patterns during active walking were significantly better than chance ( $p < 0.05$ ). In the best subject, the mean error (L1 norm) was 120 ms. The reconstruction was also successful during passive walking in every subject. No significant difference ( $p = 0.77$ , Wilcoxon signed rank test) was found between the errors of the active and passive condition. Detailed error reports are shown in Table 1 with the corresponding carrier frequency ( $f$ ),  $p$ -values and the relative amount of trials in which the error was better than random (BTR). Note that in 6/10 the BTR percentage was greater than 90%. The error histograms are illustrated in Fig. 1.

subject	Active					Passive				
	f Hz	$err_{recon}$ s	$err_{chance}$ s	p	BTR %	f Hz	$err_{recon}$ s	$err_{chance}$ s	p	BTR %
S1	27	0.12	0.51	<0.001	100	26	0.09	0.51	<0.001	100
S2	34	0.14	0.50	<0.001	96.84	35	0.17	0.50	<0.001	92.91
S3	29	0.24	0.60	<0.001	92.56	29	0.30	0.62	<0.001	88.63
S4	31	0.22	0.59	<0.001	92.33	30	0.28	0.61	<0.001	88.73
S5	30	0.23	0.48	<0.001	91.80	27	0.24	0.53	<0.001	87.56
S6	29	0.21	0.52	<0.001	91.61	27	0.19	0.52	<0.001	93.91
S7	20	0.46	0.57	<0.001	63.79	20	0.39	0.56	<0.001	76.07
S8	27	0.47	0.47	0.54	53.75	34	0.34	0.48	<0.001	84.54
S9	20	0.29	0.54	<0.001	82.49	20	0.40	0.61	<0.001	68.62
S10	20	0.43	0.44	0.16	48.80	35	0.22	0.41	<0.001	88.44

Table 1: Low gamma carrier frequencies, reconstructed and chance level error means with according  $p$ -value and BTR percentage for active (left panel) and passive (right panel) walking.

### 4 Discussion

In this work we presented a system that is able to reconstruct gait cycle patterns from non-invasive EEG recordings. We were able to extract the low gamma modulation pattern, the carrier frequencies, and their phases (time lag relative to the heel strike) using only one third of the data for training. These values are consistent with previous analysis reported in [5], [6]. These findings suggest that the reconstruction method in this work indeed is based on the cortical activity not on artefacts.

The proposed method also works for passive walking in every subject while during active walking it was successful in 8/10. In the analysis we observed that the quality of the reconstruction is reduced when artefacts are present. Our data suggests that head/neck muscular activity is lower during passive walking than during active walking. This effect is notable at lateral and dorsal electrodes. Thus, it is plausible that the lower influence of muscular artefacts during passive walking is the reason why a good reconstruction is possible in all participants. However, EEG recordings are contaminated with muscular artefacts in a broad frequency range during walking [2]. Analysing EEG data on a single-trial level during walking is therefore very challenging. Further work is needed to evaluate to which extent muscular activities influences the reconstruction of gait related parameters in comparison to the overall performance. In this work, the reconstruction performance was found to be driven by physiological meaningful features. In conclusion, we consider this work as a further step towards analysing and monitoring cortical activity in real-time during human upright walking.

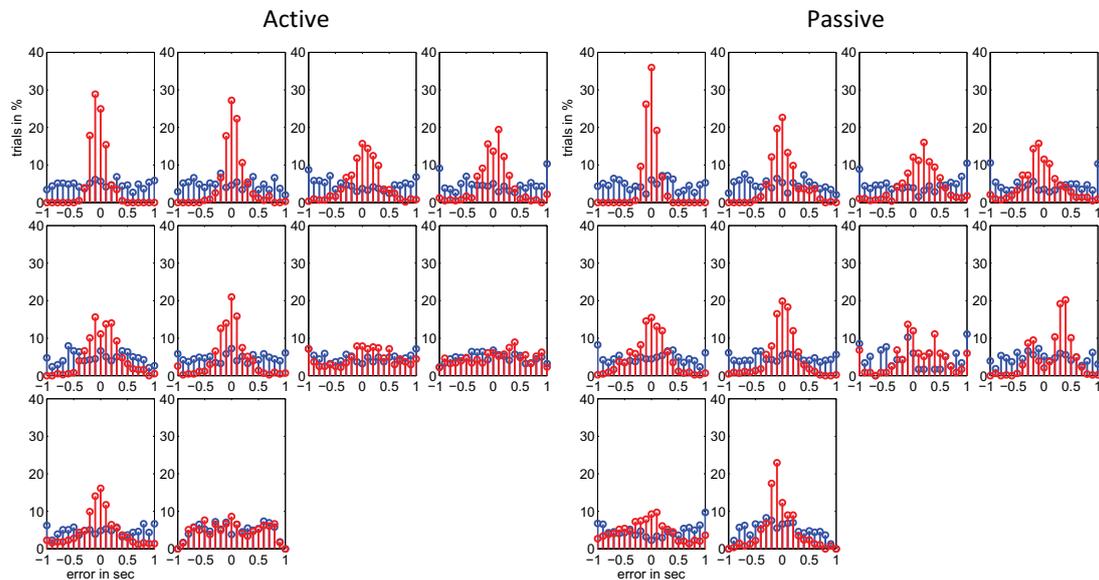


Figure 1: Error histograms for active (left) and passive (right) walking in red, for standing (chance level errors) in blue (S1-S10, top left to bottom right). Error distributions of the walking (red) conditions are centred around zero with an ideally small standard deviation for successful reconstructions. Standing condition errors (blue) seem to follow an uniform distribution.

## Acknowledgements

This work was supported by the European Union research project BETTER, BioTechMed and the Land Steiermark project BCI4REHAB.

## References

- [1] J. M. Antelis, L. Montesano, A. Ramos-Murguialday, N. Birbaumer, and J. Miguez. On the usage of linear regression models to reconstruct limb kinematics from low frequency EEG signals. *Plos One*, 2013.
- [2] T. Castermans, M. Duvinage, G. Cheron, and T. Dutoit. About the cortical origin of the low-delta and high-gamma rhythms observed in EEG signals during treadmill walking. *Neuroscience Letters*, 561:166–170, 2014.
- [3] J. Morlet, G. Arens, E. Fourgenau, and D. Glard. Wave propagation and sampling theory part I: Complex signal and scattering in multilayered media. *Geophysics*, 47:222–236, 1982.
- [4] A. Presacco, R. Goodman, L. Forrester, and J.L. Contreras-Vidal. Neural decoding of treadmill walking from noninvasive electroencephalographic signals. *Journal of Neurophysiology*, 106:1875–1887, 2011.
- [5] M. Seeber, R. Scherer, J. Wagner, T. Solis-Escalante, and G.R. Müller-Putz. EEG beta suppression and low gamma modulation are different elements of human upright walking. *Frontiers in Human Neuroscience*, 8:485, 2014.
- [6] J. Wagner, T. Solis-Escalante, P. Grieshofer, C. Neuper, and R. Scherer G.R. Müller-Putz. Level of participation in robotic-assisted treadmill walking modulates midline sensorimotor eeg rhythms in able-bodied subjects. *NeuroImage*, 63:1203–1211, 2012.

# A Probabilistic Graphical Model for Word-Level Language Modeling in P300 Spellers

Jaime F. Delgado Saa<sup>1,3</sup>, Adriana de Pestere<sup>2</sup>, Dennis McFarland<sup>2</sup> and Müjdat Çetin<sup>1</sup>

<sup>1</sup> Sabanci University, Signal Proc. & Info. Syst. Lab, Istanbul, Turkey  
delgado@sabanciuniv.edu, mctin@sabanciuniv.edu

<sup>2</sup> Wadsworth Center, Lab of Injury and Repair, Albany, NY, USA.  
adrianad@wadsworth.org, mcfarlan@wadsworth.org

<sup>3</sup> Universidad del Norte, Robotics and Intelligent Systems Lab, Barranquilla, Colombia.  
jadelgado@uninorte.edu.co

## Abstract

Motivated by P300 spelling scenarios involving communication based on a limited vocabulary, we propose a probabilistic graphical model-based framework and an associated classification algorithm that uses learned statistical prior models of language at the level of words. Exploiting such high-level contextual information helps reduce the error rate of the speller. The proposed approach models all the variables in the P300 speller in a unified framework and has the capability to correct errors in previous letters in a word given the data for the current one. The structure of our model allows the use of efficient inference algorithms, which makes it possible to use this approach in real-time applications. Our experimental results demonstrate the advantages of the proposed method.

## 1 Introduction

Recently there has been growing interest in the incorporation of statistical language models into P300 spellers with the intention to reduce the error rate or to increase typing speed [1, 2, 3, 4, 5]. These approaches learn marginal and conditional probabilities of letters in a language based on some corpus and use that information in the form of prior models in a P300-based brain-computer interface (BCI) system. In particular, such prior models are combined with measured electroencephalography (EEG) data in a Bayesian framework to infer the letters typed by the subject. The probabilistic structure in most of this work can be described by hidden Markov models, which are one of the simplest forms of probabilistic graphical models. Work under this theme involves filtering [2, 4], as well as smoothing methods [5]. The latter type of method allows the EEG data for the current letter to affect the decision on a previous letter through the language model, and possibly correct an erroneous decision that would be reached in the absence of such data. In most of this body of work, first a conventional classifier for the P300 speller is utilized and then the scores of that classifier are turned into probabilities to be combined with the language model for Bayesian inference.

In this work, we propose taking one further step incorporating higher-level, in particular word-level, language models into P300 spellers. Our motivation comes from BCI applications that involve typing based on a limited vocabulary. In a particular context, if it is known that a user is likely to type words from a dictionary of, say, a few thousand words, that provides very valuable prior information that can potentially be exploited to improve the performance of the BCI system. Based on this perspective, we propose a discriminative graphical model-based framework for the P300 speller with a language model at the level of words. The proposed model integrates all the elements of the BCI system from the input brain signals to the spelled

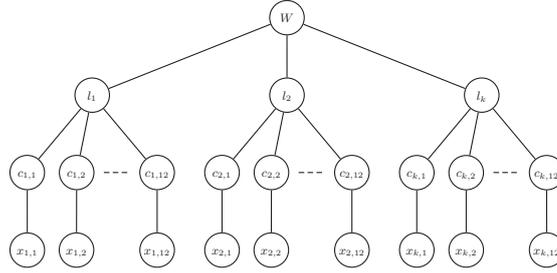


Figure 1: Proposed graphical model framework for the P300 speller.

word in a probabilistic framework in which classification and language modeling are integrated in a single model and the structure of the graph allows efficient inference methods making the system suitable for online applications. Results show that the proposed method provides significant improvement in the performance of the P300 speller by increasing the classification accuracy while reducing the number of flash repetitions.

## 2 Methods

**Proposed Graphical Model:** The proposed model is shown in Figure 1. In the bottom (first) layer, the variables  $x_{i,j}$  represent the EEG signal recorded during the intensification of each row and column of the spelling matrix. The index  $i$  denotes the ordinality of the letter being spelled and the index  $j$  represents a row or column ( $j = \{1, \dots, 6\}$  for rows and  $j = \{7, \dots, 12\}$  for columns). The second layer contains a set of twelve variables  $c_{i,j}$  indicating the presence or absence of the P300 potential for a particular flash. The third layer contains variables  $l_i$  representing the letter being spelled. The variables  $l_i$  are related to the variables  $c_{i,j}$  in the same fashion as in traditional P300 speller systems: the presence of a P300 potential in a particular row-column pair encodes one letter. The fourth layer contains the variable  $w$  which can be any member of a particular subset of valid words in the English language. A learned probability mass function for this variable constitutes the language model in this work.

The distributions of all the variables in the model ( $w, \mathbf{l} = \{l_1, \dots, l_k\}, \mathbf{c} = \{c_{1,1:12}, \dots, c_{k,1:12}\}$ ) given the observations ( $\mathbf{x} = \{x_{1,1:12}, \dots, x_{k,1:12}\}$ ) can be written as a product of factors over all the nodes and edges in the graph:

$$P(w, \mathbf{l}, \mathbf{c} | \mathbf{x}) = \frac{1}{Z} \Psi_4(w) \prod_i \left\{ \Psi_3(i, w, l_i) \prod_{j=1}^{12} \{ \Psi_2(j, l_i, c_{i,j}) \Psi_1(j, c_{i,j}, x_{i,j}) \} \right\} \quad (1)$$

where  $Z$  is a normalization factor and  $\Psi_4, \Psi_3, \Psi_2$  and  $\Psi_1$  are potential functions related to nodes and edges. The potential functions are defined as follows:

$$\begin{aligned} \Psi_4(w) &= e^{\theta_4 f_4(w)} & \Psi_3(i, w, l_i) &= e^{\theta_{3i} f_3(i, w, l_i)} \\ \Psi_2(j, l_i, c_{i,j}) &= e^{\theta_{2j} f_2(j, l_i, c_{i,j})} & \Psi_1(j, c_{i,j}, x_{i,j}) &= e^{\sum_{m=1}^d \theta_{1j,m} f_{1m}(j, c_{i,j}, x_{i,j,m})} \end{aligned} \quad (2)$$

where  $d$  is the dimensionality of the data. The parameter  $\theta_4$  is a vector of weights of length equal to the number of states of the node  $w$  (i.e., the number of words in the dictionary). The

product  $\theta_4 f_4(w)$  models a prior for the probability of a word in the language with the feature function  $f_4(w) = \mathbf{1}_{\{w\}}$ , where  $\mathbf{1}_{\{w\}}$  is a vector of length equal to the number of words in the dictionary, with a single nonzero entry of value 1 at the location corresponding to the argument of  $f_4$ . The product  $\theta_{3i} f_3(i, w, l_i)$  models a prior for the probability of a letter  $l_i$  appearing in the position  $i$  of a word with the feature function  $f_3(i, w, l_i) = \mathbf{1}_{\{w(i), l_i\}}$ . The product  $\theta_{2j} f_2(j, l_i, c_{i,j})$  measures the compatibility between the binary random variable  $c_{i,j}$  and the variable  $l_i$  with the feature function  $f_2(j, l_i, c_{i,j}) = \mathbf{1}_{\{C(l_i, j) = c_{i,j}\}}$  where  $C$  is a code-book that maps the intersections of rows and columns in the spelling matrix to letters. The product  $\theta_{1j,m} f_{1m}(j, c_{i,j}, x_{i,j,m})$  is a measure of the compatibility of the  $m$ -th element of the EEG signal  $x_{i,j} \in R^d$  with the variable  $c_{i,j}$ . Here, we use the feature function  $f_{1m}(j, c_{i,j}, x_{i,j,m}) = x_{i,j,m} \mathbf{1}_{\{c_{i,j}\}}$ . Learning in the model corresponds to finding the set of parameters  $\Theta = \{\theta_4, \theta_3, \theta_2, \theta_1\}$  that maximizes the log-likelihood of the conditional probability density function given in Equation 1. Given that the structure of the model does not involve loops, inference in the model can be performed using the belief propagation algorithm which can efficiently provide the posterior probabilities for the words:  $P(w|\mathbf{l}, \mathbf{c}, \mathbf{x}) = \sum_l \sum_c P(w, \mathbf{l}, \mathbf{c}|\mathbf{x})$ . We declare the word by maximizing the posterior density:  $\bar{w} = \arg \max_w P(w|\mathbf{l}, \mathbf{c}, \mathbf{x})$ . Note that the model allows computing other marginals of interest (e.g., letters) as well.

**Description of Experiments:** Two kinds of experiments are reported. In the first experiment, 8 subjects are instructed to spell a number of words one by one. In this scenario (which we call screening) the number of letters in the words typed in the testing session is known. In the second experiment 7 subjects write multiple words in a continuous fashion, using the character "-" to indicate the end of a word. The EEG signals were recorded at a sample frequency of 240 Hz using a cap embedded with 64 electrodes according to the 10-20 standard. All electrodes were referenced to the right earlobe and grounded to the right mastoid. From the total set of electrodes a subset of 16 electrodes in positions F3, Fz, F4, FCz, C3, Cz, C4, CPz, P3, Pz, P4, PO7, PO8, O1, O2, Oz were selected, motivated by the study presented in [6]. In total each subject spelled 32 letters (9 words). Training and testing sessions were held on different days for the same subjects. Signal segments of 600ms following the intensification of each row or column were calculated and filtered between 0.5Hz and 8Hz using a zero-phase IIR filter of order 8.

"The parameters  $\theta_4, \theta_{3i}, \theta_{2j}$  in Equation 2 are independent of the brain signals and can be learned prior to any EEG data collection. The parameters  $\theta_4$  are learned by calculating the relative frequency of each word in the dictionary of interest. The parameters  $\theta_{3i}$  take a positive or negative value depending of the presence of absence of a letter in the  $i$ -th position of a word. Statistics about the language were calculated using a text corpus with 450,000,000 words. The dictionary was then built using the 5000 words with highest frequency. The parameters  $\theta_2$  represent a set of decoding vectors that map rows and columns to letters in the spelling matrix. The parameters  $\theta_1$  are the set of parameters that maximize the potential function  $\Psi_1(j, c_{i,j}, x_{i,j})$  in Equation 2 for each class (i.e., P300 vs. not P300).

### 3 Results

Figure 2 shows the results for the screening scenario and for continuous decoding where a 3-gram method and a classifier based in Stepwise LDA have been used for comparison. The results show that the proposed method performs better than the other methods in terms of classification accuracy and requires fewer numbers of repetitions to achieve a particular level of accuracy. In order to verify the results of correct classification accuracy, a statistical test

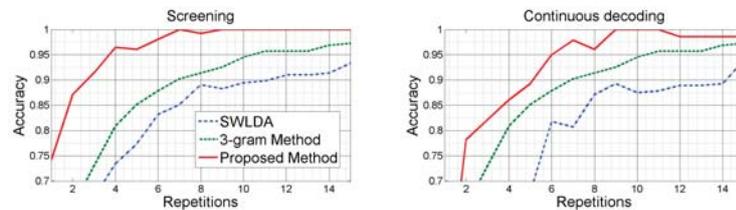


Figure 2: Average classification accuracy across subjects.

was performed. For the screening scenario, a repeated measures ANOVA on the performance results reveals significant difference ( $F(2, 14), \epsilon = 0.56, p = 0.0041$ ) between the three compared methods. Using a post hoc Tukey-Kramer test, the proposed method performs significantly better ( $p < 0.01$ ) than the 3-gram based method and than SWLDA. For the continuous decoding scenario the results are similar, the proposed method performs significantly better ( $p < 0.01$ ) than the 3-gram based method and than SWLDA.

## 4 Conclusion

We present a probabilistic framework as well as an inference approach for P300-based spelling that exploits learned prior models of words. While language models at the level of letters have previously been proposed for BCI, word-level language modeling is new. The structure of the model we propose enables the use of efficient inference algorithms, making it possible to use our approach in real-time applications. While our approach can in principle be used with word prior models learned from any corpus, we expect it to be of special interest for applications involving the use of a limited vocabulary in a specific context.

## Acknowledgments

This work was partially supported by Universidad del Norte (Colombia), the Scientific and Technological Research Council of Turkey under Grant 111E056, by Sabanci University under Grant IACF-11-00889, and by NIH grant EB00085605 (NIBIB).

## References

- [1] Tanja Mathis and Dennis Spohr. Corpus-Driven Enhancement of a BCI Spelling Component. In *RANLP'07: Int. Conf. on Recent Advances in Natural Language Processing*, September 2007.
- [2] Umut Orhan et al. Offline analysis of context contribution to ERP-based typing BCI performance. *Journal of Neural Engineering*, 10(6), 2013.
- [3] S.T. Ahi, H. Kambara, and Y. Koike. A Dictionary-Driven P300 Speller With a Modified Interface. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 19(1):6–14, 2011.
- [4] William Speier et al. Natural language processing with dynamic classification improves P300 speller accuracy and bit rate. *Journal of Neural Engineering*, 9(1), 2012.
- [5] C. Ulas and M. Cetin. Incorporation of a Language Model into a Brain Computer Interface based Speller through HMMs. In *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, 2013.
- [6] D.J. Krusienski, E.W. Sellers, D.J. McFarland, T.M. Vaughan, and J.R. Wolpaw. Toward Enhanced P300 Speller Performance. *Journal of Neuroscience Methods*, 167(1):15 – 21, 2008.

# Feasibility of using time domain parameters as online therapeutic BCI features

Osuagwu, A. Bethel<sup>1</sup> and Vuckovic, Aleksandra<sup>1</sup>

Centre for rehabilitation engineering, School of Engineering, University of Glasgow, Glasgow, UK  
1108643o@student.gla.ac.uk

## Abstract

Feature extraction and selection is a major issue in brain computer interfaces. In an electroencephalogram based brain computer interface bandpower features are widely used. Time domain parameters (TDP) are other features which have not been extensively tested in online brain computer interfaces. In this study with eight naive subjects, it is shown that the time domain parameters (TDP) are suitable online features for a motor imagery based therapeutic brain computer interface. ERD/ERS maps were compared between trials selected as motor imagery-active and those rejected when using TDP as features. The ERD/ERS maps of the trials selected with the TDP method show ERD mainly in the 8-12 Hz frequency band on the hemisphere corresponding to the hand the subjects were imagining to move. There were significantly stronger ERDs in the trials that were selected than in those rejected.

## 1 Introduction

In electroencephalogram (EEG) based brain computer interface (BCI), it is important to extract the appropriate feature suitable for classifying EEG arising from different cognitive tasks. Typically the bandpower features are extracted from the EEG and used in the online classification. These features are easy to compute and require minimum number of EEG electrodes minimizing setup time. This advantage makes the bandpower method suitable especially for therapeutic BCI where setup time must be minimized. The bandpower features are well established and researchers can target physiologically relevant frequency bands when using it as features for therapeutic BCI. However the need to select a user specific narrow frequency bands can be an issue. Firstly because the user specific bands are known to vary for the same user. Secondly due to the uncertainty principle, estimating narrow band spectral powers must be dubious within short time windows.

The time domain parameters (TDP) [2] are features which do not require the selection of narrow frequency bands. In addition TDP target the most time varying EEG features. TDP features have been shown to outperform bandpower [2]. However to be suitable for therapeutic BCI, TDP must be able to select physiologically relevant frequencies despite the use of a wide band filtered signal. Furthermore, the current authors could only find one report [2] on the use of TDP in online BCI; more investigation is required before using TDP on patients.

The aim of this study is to assess the suitability of using TDP for online therapeutic BCI by analysing the time-frequency components of EEG classified into two classes using TDP. It is shown that despite using wide band, TDP target the relevant frequency band.

## 2 Methods

Eight BCI naive subjects took part in this study after giving their informed consents. The study was approved by the university ethic committee.

To obtain an initial classifier, a subject performed the motor imagery (MI) of closing and opening of the left and the right hand (20 trials for each hand). During these tasks EEG was recorded, using the g.USBamp (GTEC, Austria) from three pairs of bipolar electrodes, namely Fc3-Cp3, Fcz-Cpz and Fc4-Cp4. The input signal from the amplifier was bandpass filtered (5<sup>th</sup> order Butterworth) online between 0.5 to 30 Hz. The TDP of the filtered data were computed.

The TDP features were computed in a similar way as that described by Vidaurre and colleagues [2]. This is shown graphically on Figure 1. The feature was used to compute initial

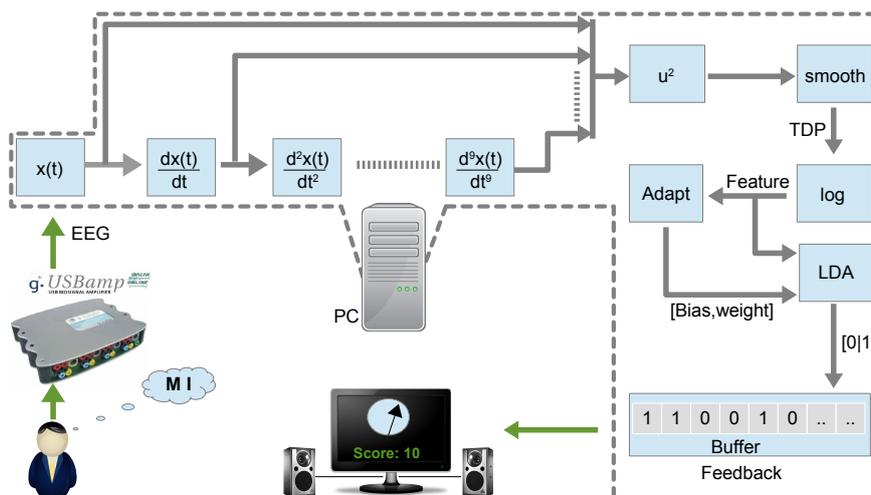


Figure 1: BCI setup showing the computation of TDP

linear discriminant analysis classifier to discriminate right hand movement MI from resting state. This initial classifier was saved for online use. These steps were all integrated into a graphical user interface.

In the online classification, TDP features were estimated using Simulink’s difference blocks (see Figure 1). Each sample of the signal in the feature space was binary classified either as an ‘Active’ or ‘Relaxed’ state using the initial classifier computed offline. The ‘Active’ state occurred when the subject attempted opening and closing of the right hand MI while the ‘Relaxed’ state corresponded to a resting period. The classifier output (‘1’ or ‘0’ for ‘Active’ or ‘Relaxed’ respectively) was then buffered for a variable length of time up to a maximum of 3 s or 768 samples. An ‘Active’ state was detected when the buffer was filled with a chosen percentage of ‘Active’ state. The length of the sub-buffer, usually 1.5-2 s long, was determined for the subject and optimized to significantly reduce false positives which was reported by the subject. The difficulty of the BCI was set to 50% determined using the equation,  $d = bf/B$ , where  $d$  is the difficulty,  $b$  is the sub-buffer length,  $B$  is the maximum buffer length (set to 3 s) and  $f$  is the percentage filling of the sub-buffer that activates the ‘Active’ state. For example when  $b = 2s$ ,  $f$  was set to 75%.

The initial classifier was updated online using the fixed rate supervised mean and covariance adaptation methods described elsewhere [1].

In the online BCI there were 30 trials in total divided into three runs of 10 trials each. A trial started at  $t = -3$  s with a cross sign on the screen facing the subject. At  $t = 0$  s an execution cue in form of an arrow pointing to the right was shown on the screen and the subject was instructed to perform right hand MI until a text and sound feedback were given. The feedback

included an acknowledgment text on the screen and a reward sound played to the subject when the ‘Active’ state was detected. If the ‘Active’ state was not detected after about 6.5 s following the execution cue onset, a text was shown and a sound was played to reflect the subject’s failure for that trial. The subjects were also provided with a continuous feedback in form of a scale that moves counter clockwise when imagery was detected.

Of interest were the time-frequency characteristics of the trials selected as ‘Active’ state (Detection) and those not selected (No Detection) when using TDP as BCI features. Therefore ERD/ERS analysis was carried out by first separating the trials from all subject and all runs into ‘Detection’ and ‘No Detection’ groups. The resulting ERD/ERS maps were compared between the two groups of trials using statistical nonparametric method with Holm’s correction for multiple comparisons at  $p=0.05$ .

### 3 Results and discussions

Table 1 shows the initial classifier accuracy, and the detection rate (true positive) per subject. The naive subjects improved the detection rate from run 1 and 2 to run 3 because they got better with experience. No false positive was reported because the subjects were imagining as soon as the execution cue was shown and the false positive was significantly reduced when the BCI difficulty was set.

Subjects	Initial classifier (% accuracy)	% rate (run 1 and 2)	% rate (run 3)
1	83	75	100
2	97	0	30
3	78	10	40
4	75	50	100
5	85	40	90
6	83	50	100
7	83	15	50
8	90	61	47

Table 1: The initial classifier accuracy and detection rate

Figure 2 shows the ERD/ERS maps of the right hand MI when ‘Active’ state was detected (first column) and when it was not detected (second column). The third column shows in frequency and time the statistical differences between column one and two for each channel. After the onset of the execution cue, ERD could be seen in all three channels suggesting that the subjects were performing MI. Visually inspecting the ERD/ERS shows that there are more ERD for the Detection group than for No Detection group. However the differences in ERD is only statistically significant in channel location Fc3-Cp3 as shown by the shaded area on the first row, third column in Figure 2. This is a desirable result because this channel is on the left hemisphere which represents the right hand which the subjects were imagining to move. The time-frequency statistical differences appears predominantly within the 8-12 Hz frequency range which is the so called  $\mu$ -band known to show ERD within the motor areas during MI. There are also small differences in the 12 -30 Hz range. It is an interesting result that the physiologically relevant frequency band is more active in the Detection group despite having not selected narrow frequency bands in this range. Despite the No Detection group showing ERD in the low frequency range they were still not selected. It was possible that the  $\mu$ -band has

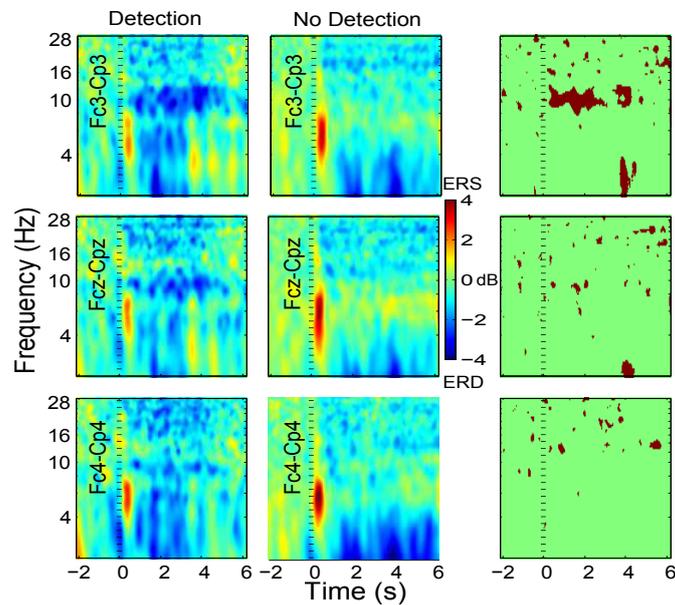


Figure 2: ERD/ERS maps of the right hand MI for when ‘Active’ state was detected (column 1) and when it was not detected (column 2). The column 3 shows the time-frequency statistical differences between column one and two ( $p=0.05$  with Holm’s correction). The execution cue was shown at  $t=0$  ms. (Generated with EEGLAB, <http://sccn.ucsd.edu/eeqlab>).

the most time varying activities during the cognitive tasks making it likely to be selected by the TDP. The statistical difference on Fc3-Cp3 at  $t=4$  s is due to the ERD in the low frequencies in the No Detection group.

There were many failed trials because the subjects were naive, no training was given, they had short time to perform the MI and only 20 trials were used to compute the initial classifier although it was updated online to compensate for the low number of trials. However this is a more realistic BCI as we tend to move it out of the laboratory and also use it in our rehabilitation programmes.

## 4 Conclusions

TDP eliminates the requirement to select a user specific frequency band allowing for a more generalised BCI classifier. It is easy to compute and require minimum EEG electrodes. The current result shows that TDP features are suitable in online therapeutic BCI because they target the physiologically relevant frequency band.

## References

- [1] Alois Schlögl, Carmen Vidaurre, and Klaus-Robert Müller. *Brain-Computer Interfaces: Revolutionizing Human-Computer Interaction*, chapter Adaptive Methods in BCI Research An Introductory Tutorial, pages 131–356. Springer, 2010.
- [2] Carmen Vidaurre, Nicole Krämer, Benjamin Blankertz, and Alois Schlögl. Time domain parameters as a feature for EEG-based brain-computer interfaces. *Neural Networks*, 22(9):1313–1319, 2009.

# Non-Linear and Spectral EEG Features during a Mental Calculation Task for Asynchronous BCI Control

Erik R. Bojorges-Valdez, Juan C. Echeverria and Oscar Yanez-Suarez

Universidad Autónoma Metropolitana Mexico D.F., Mexico  
erbv@xanum.uam.mx, jcea@xanum.uam.mx, yaso@xanum.uam.mx

## Abstract

In this paper a binary BCI control based on mental computation of arithmetic operations is evaluated. Data analysis was performed using information derived from five EEG channels, estimating the detrended fluctuation analysis scaling exponent and the power on  $\beta$  band. The strategy (task and data analysis) was validated on fifteen subjects realizing three experimental sessions on different days. Performance was measured using the area under receiver operating characteristic curve obtaining  $0.85 \pm 0.079$ ,  $0.87 \pm 0.071$  and  $0.87 \pm 0.065$  for each experimental session.

## 1 Introduction

Despite that the mechanism used to elicit a response for an asynchronous BCI application intends to be a “natural” mental activity (i.e. imagining to move a limb) the modest proportion of population capable of achieving a satisfactory control of these systems remains as major drawback [1]. Ono et al. [2] have shown that the mean accuracy values may be improved using a more realistic feedback for the training process. ERD/ERS responses can also be elicited during the realization of other activities like imagining a cube rotation, or by another type of cognitive activities such as performing mental calculation [3]. This last activity has been proposed for controlling a BCI, combining fNIRS and EEG techniques [4, 5]. Power et al., have reported accuracy values above 70%, using EEG window lengths of 5-20 seconds. Notwithstanding that these values suggest that it is possible to use such task in a BCI paradigm, the window length used produces a low information transfer rate [6]. This paper explores the feasibility of mental calculation, during shorter time windows, to be used as control for a BCI implementation.

## 2 Task and Data Collection

EEG data were recorded from fifteen subjects, seven men and eight women,  $25.3 \pm 3.47$  years old, all with completed high school studies. A 32 channel (Fp[z,1,2], AF[7,3,z,4,8], F[7,3,z,4,8], Fc[3,4], T[7,8] C[3,z,4], Cp[3,4], P[7,3,z,4,8], PO[3,4] and O[1,2,z]) montage was used and the signals were digitized at 512 sps with a g.tec USBamp amplifier system, using a bandpass filter between [0.1, 60] Hz and a notch filter at 60 Hz. All subjects were recorded for three sessions on different days (the lapse of time between sessions extended from 6-110 days, according to the time availability of volunteers); each session consisted of two or three runs. In each run the subject was instructed to pay attention to a screen and solve mentally fourteen sets of basic concatenated arithmetic operations. A set was formed by the concatenation of four to six randomly selected simple operations (“+”, “-”, “.”, “/”). The beginning and ending of each set was visually cued by an “X” and an “=” symbol, respectively. At the end of each set, the subject was asked to vocalize the answer.

As illustrated on figure 1, each set consisted of five different screens: *Cue*, *Begin*, *Operate*, *Answer* and *Rest*, presenting each one for two seconds with a random ISI duration within [625,750] ms. Subjects performed periods of continuous mental solving through the *Begin* and *Operate* screens, and idle periods at the *Rest* screens.

For data analysis, signals were conditioned by removing CAR and lowpass filtered at 40 Hz. Afterwards, signals were analyzed using windows of two seconds synchronized with the screen presentation for the channel selection process, and using sliding windows for an online BCI simulation. Epochs were labeled in relation to solving or resting periods, and only those sets with a correct answer were used for estimating the model and evaluating the performance.

Five different indexes were extracted from EEG data: the power spectral density (PSD) over the four classical bands,  $\delta_{PSD}$ ,  $\theta_{PSD}$ ,  $\alpha_{PSD}$  and  $\beta_{PSD}$  using the Welch Periodogram method, and the scaling exponent  $\alpha_{DFA}$  obtained using Detrended Fluctuation Analysis (DFA) [7, 8].

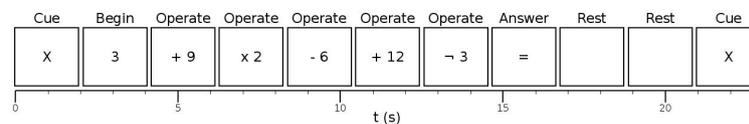


Figure 1: Example of a set of operations used for stimulation, infix notation:  $((((3 + 9) \cdot 2) - 6) + 12) / 3 = 10$ .

### 3 Methods

#### 3.1 Channel Selection

Channel selection was based on a statistical analysis applied to each electrode as follows. Valid epochs of each session, as synchronized with the screen presentation and correct arithmetic answer, were divided into two sets, *Operating* and *Idle*. To avoid including false positive channels not reflecting a real region of activity, only AF[3,z,4], F[7,3,z,4,8], FC[3,4], C[3,z,4] and CP[3,4] were considered. For each of these channels and each index ( $\delta_{PSD}$ ,  $\theta_{PSD}$ ,  $\alpha_{PSD}$ ,  $\beta_{PSD}$  and  $\alpha_{DFA}$ ), the area under the receiver operating characteristic (ROC) curve (AUC) was individually estimated using these two labels. The channels were sorted according to AUC value and the top five of each index were selected. Hence, the five most repeated channels were chosen as the final set.

#### 3.2 Performance evaluation

Using  $\alpha_{DFA}$  and  $\beta_{PSD}$  estimated from the five selected channels, evaluation of performance was carried out by both cross-validation (CV) and inter-session identification (iSI). In both cases, performance was measured with AUC values using a data set also labeled as subsection 3.1, but in this case epochs were extracted from sliding windows of two seconds length and overlapped by 70%. The CV process was applied to find support vector machine (SVM) parameters that optimized AUC values using repeated sub-sampling method with 20-folds. The mean and standard deviation were assessed and considered as the final CV performance value. Then, the classification model corresponding to each session was constructed and tested on the remaining two sessions. Therefore for each subject three values of CV were obtained and six related to iSI; Figure 2(a) presents box plots which summarize these results.

## 4 Results

Using sliding windows the mean AUC values for CV assessed for each session, were  $0.85 \pm 0.079$ ,  $0.87 \pm 0.071$  and  $0.87 \pm 0.065$ . The distribution of the individual values was above 0.7 for all sessions as can be seen on Figure 2(a). iSI evaluation shows a mean AUC value of 0.7 which is below the CV evaluation but acceptable for a BCI implementation; in particular when considering that for this process, the model used and data identified were recorded in different days. The individual distribution of iSI evaluation shows that about half of the population achieved a value above 0.7, and only two subjects performed below 0.6.

For illustration purposes, Figures 2(b) and 2(c) show two examples of SVM posterior probability estimates (SVMpp) and periods of mental activity. Given that SVMpp is proposed as the control signal for a BCI implementation, it should follow as is manifested in these examples the realization of the mental task, which means that it must be related with the decision to select the command or, for the case studied here, with the periods of mental activity due to mental calculation.

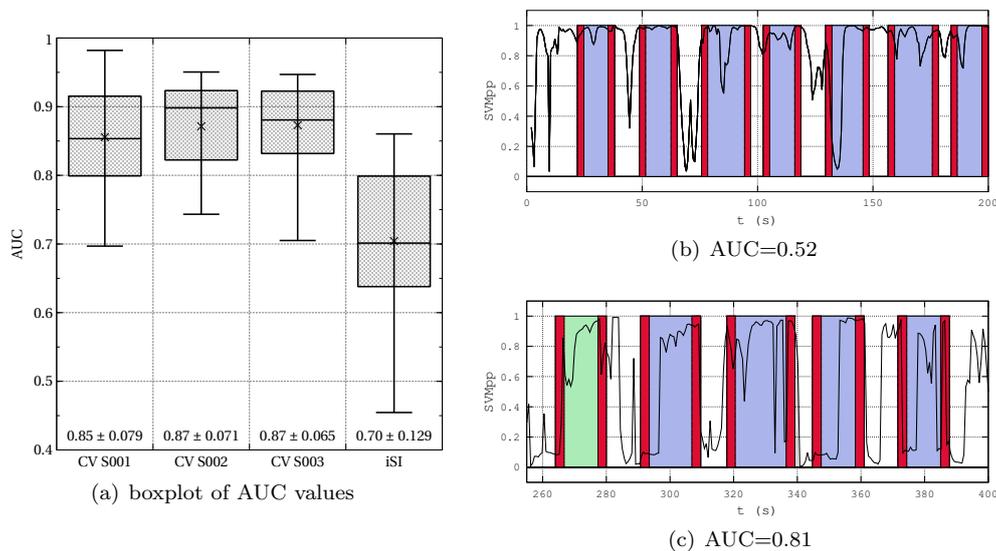


Figure 2: The distributions of the AUC values achieved for the three CV sessions and iSI are presented in subfigure (a), the mean and standard deviation values are included over the horizontal axis. Subfigures (b) and (c) depict two examples of SVMpp (black line) assessed for two subjects that achieved different AUC values, the light blue and green shadows indicate the periods of mental calculation realization with either correct or wrong answers, respectively, and the red ones indicate the periods for cue presentation and vocalization of the answer.

## 5 Discussion and Conclusions

Day-to-day activities could be fitted as stimulation paradigms for an asynchronous BCI. Here the utilization of mental calculation is demonstrated to be useful for such context. The results presented in this paper suggest that, for asynchronously controlling a BCI, it seems possible

to use mental computation with overlapped EEG windows of two seconds length. More experiments are required to assess the potential interference of other factors, such as tiredness, distraction or low attention. Importantly, the AUC values achieved were above 0.8; which appear better than the performance reported on previous publications that also use mental calculation tasks but under different experimental conditions [4, 5]. As can be seen on the examples of Figure 2, the classification output shows a rapid response to the changes of mental activity which is also a desirable condition for a BCI control. This encourages the utilization of mental calculation as a BCI activation paradigm.

## 6 Acknowledgments

First author acknowledges the National Council of Science and Technology (CONACyT) for support his PhD studies, grant 16848.

## References

- [1] C. Guger, G. Edlinger, W. Harkam, I. Niedermayer, and G. Pfurtscheller. How many people are able to operate an EEG-based brain-computer interface (BCI)? *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11:145–147, 2003.
- [2] Takashi Ono, Akio Kimura, and Junichi Ushiba. Daily training with realistic visual feedback improves reproducibility of event-related desynchronisation following hand motor imagery. *Clinical Neurophysiology*, 124:1779–1786, 2013.
- [3] C. Neuper and G. Pfurtscheller. Event-related dynamics of cortical rhythms: frequency-specific features and functional correlates. *International Journal of Psychophysiology*, 43:41–58, 2001.
- [4] Sarah D Power, Azadeh Kushki, and Tom Chau. Towards a system-paced near-infrared spectroscopy braincomputer interface: differentiating prefrontal activity due to mental arithmetic and mental singing from the no-control state. *Journal of Neural Engineering*, 8:14 pp, 2011.
- [5] Sarah D Power, Tiago H Falk, and Tom Chau. Classification of prefrontal activity due to mental arithmetic and music imagery using hidden Markov models and frequency domain near-infrared spectroscopy. *Journal of Neural Engineering*, 7:9 pp, 2010.
- [6] Jonathan R. Wolpaw, Niels Birbaumer, Dennis J. McFarland, Gert Pfurtscheller, and Theresa Vaughan. Brain-computer interfaces for communication and control. *Clinical Neurophysiology*, 113:767–791, 2002.
- [7] C.-K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger. Mosaic organization of DNA nucleotides. *Phys. Rev. E*, 49:1685–1689, Feb 1994.
- [8] E.R. Bojorges-Valdez, J.C. Echeverria, and O. Yanez-Suarez. Task identification using fluctuation analysis. In *Proceedings of the Fifth International Brain-Computer Interface Meeting 2013*.

# Discriminating Between Attention and Mind Wandering During Movement Using EEG

Filip Melinscak<sup>1,2</sup>, Luis Montesano<sup>2</sup> and Javier Minguez<sup>1,2</sup>

<sup>1</sup> Bit&Brain Technologies, S.L., Zaragoza, Spain

<sup>2</sup> University of Zaragoza, Zaragoza, Spain

E-mail: [filip.melinscak@gmail.com](mailto:filip.melinscak@gmail.com)

## Abstract

Attention, and conversely mind wandering, are believed to be important factors in physical rehabilitation. We propose an experimental protocol to investigate if it is possible to discriminate between attention and mind wandering during passive movements of lower limbs using EEG. We performed time-frequency analysis of the gathered data and designed a simple brain-computer interface (BCI) based on oscillatory features. The designed BCI achieved average accuracy of 75% in single trials, on a sample of five healthy subjects.

## 1 Introduction

Attention is believed to be an important cognitive factor in physical rehabilitation [7]. Conversely, repetitive tasks – which are typical of rehabilitation regimes – can induce inattentive cognitive states, termed *mind wandering* [6]. These considerations are relevant in, e.g., motor rehabilitation after stroke, where repetitive task-specific practice is a common intervention [3]. Therefore, an EEG-based brain-computer interface (BCI), capable of discriminating between attention and mind wandering, might prove useful in a number of different settings. In classical rehabilitation, the therapist could use such a device as a window into the patient's cognitive state, guiding the exercise regime. In the robot-assisted rehabilitation scenario, attention level could be used as an additional control signal, adhering to the human-in-the-loop concept. Additionally, such a device might allow to more rigorously answer research questions on the role of attention in rehabilitation.

We are building on the work described in [1]. The referenced work has shown that a distractor task (specifically, counting backwards by threes) modulates desynchronization of sensori-motor rhythms (SMR) and that a classifier can be designed to discriminate between attended and unattended passive movements of upper limbs, with an accuracy of around 75%.

Our approach differs from prior work in the way we defined the conditions: instead of contrasting attention with distraction by an artificial mental task, we contrast attention with a mind wandering state, more realistically modeling the rehabilitation scenario. Furthermore, we focus on a specific aspect of attention – the kinesthetic sensation of the movement – trying to disentangle it from other aspects like visual attention. Lastly, in this study we focused on the lower limbs instead of upper, envisioning the application of the developed BCI to a gait-rehabilitation robot.

## 2 Methods

### 2.1 Experimental Protocol and Data Collection

Five healthy male participants (age range: 23 – 28 years old) volunteered in the experiment. Subjects were seated in front of a screen with their feet strapped to an electrical mini bike that

was used to actuate passive movements. The view of participants' feet was obscured to prevent confounds with visual attention. EEG recording was carried out with 30 channels (positioned according to the 10-10 system). TMSi Refa 32 amplifier was used with 256 Hz sampling rate and linked ears average reference.

Each subject was exposed to two experimental conditions. In the first condition subjects were instructed to pay attention to the kinesthetic sensation of the passive movement – we denote this as the “Passive Movement with Attention” (PMA) condition. In the second condition (denoted PMR – “Passive Movement with Relaxation”) subjects were instructed to relax, ignore the movement and let their mind wander. For both conditions subjects were instructed not to make muscle contractions themselves and to not visualize or imagine the movements.

The recording sessions were divided into 6 blocks, each consisting of 10 consecutive trials of one, and 10 consecutive trials of the other condition. The ordering of conditions within blocks was semi-randomized, with 3 blocks having PMR, and 3 having PMA trials first. Before each block the subjects were informed what type of trials follows and could take a pause. The duration of trials was 15 seconds: for the first 5 seconds (baseline period) the message “Rest” was displayed; during the next 10 seconds the mini bike was working and a fixation cross was displayed. A sound cue, played 500 ms prior to the appearance of the fixation cross, prompted the experiment operator to turn on the bike. The sound was also audible to the subject. Recording time was around 50 minutes, and 60 trials per condition were collected.

## 2.2 Time-Frequency Analysis

To verify whether attention and mind wandering have an effect on sensori-motor rhythms, we performed time-frequency analysis of collected data. Before the time-frequency decomposition, data was preprocessed: zero-phase IIR filter with the pass band between 1 and 40 Hz was applied; data was segmented into 15 s long trials, with 5 s before the appearance of the fixation cross and 10 s after; artifactual trials were excluded on the basis of visual inspection (on average 20% of trials were rejected).

Event-related spectral perturbation (ERSP) [2] was calculated with Morlet wavelets. For baseline correction all the spectrum estimations were divided with the mean spectrum of the  $-3$  s to  $-1$  s period of all the trials. Using a divisive baseline, we are assuming a gain model of task activity: power in the task period is a modulation of power in the baseline period.

## 2.3 Classification

To check how well PMA trials can be discriminated against PMR trials we utilized a simple bandpower BCI design. In the preprocessing step signals were filtered with a causal FIR filter with the pass band from 8 to 35 Hz (capturing the alpha and beta bands that are relevant for SMR). Next, several different spatial filtering variants were applied: no spatial filter (all 30 channels used); surface Laplacian with 4 closest neighbors applied to all the channels; selection of 7 electrodes over the motor cortex (FC1, FC2, C3, Cz, C4, CP1, CP2); only channel Cz with a surface Laplacian. In feature extraction the logarithm of variance in the period of 1 to 8 s after the appearance of the fixation cross was calculated (resulting in 30, 7 or 1 feature, depending on the spatial filter). No artifact rejection was performed, i.e. all the recorded trials were used. An LDA classifier was then trained and tested on this data. The classifier was regularized using covariance shrinkage (with the regularization parameter determined analytically; see [5]). The classifier performance was then estimated using a 5-fold chronological cross-validation scheme [4], with 5 trials before and after the testing block omitted from the training set.

### 3 Results

In Fig. 1 we are showing the results of time-frequency analysis for all the subjects, for the electrode Cz which displayed strongest movement-related spectral changes. At a qualitative level we observe that for all the subjects passive movement produced a prominent desynchronization in the upper beta band with a spatial distribution concentrated over the central electrodes. For some subjects a desynchronization in the alpha band could also be observed, but with a more diffuse spatial distribution (possibly caused by the contribution of occipital or temporal rhythms in the alpha band, and not by the mu rhythm). For subject 3 we could also observe a synchronization at around 30 Hz that increased in bandwidth during the course of a trial.

Averaging ERSP values over 1-8 s time window and over alpha and beta bands (selected for each subject by visual inspection) yielded significant differences between the PMA and PMR conditions for subjects 3 and 4 in both bands (two-sample *t*-test at  $\alpha = 0.05$ , Bonferroni corrected).

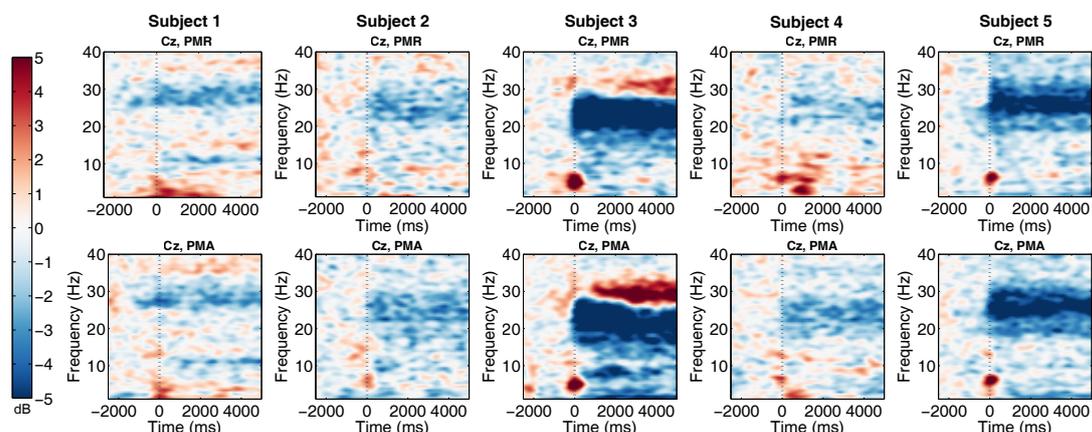


Figure 1: ERSP maps for all subjects, for electrode Cz. Upper plots show ERSPs for the PMR condition, and the lower ones for the PMA condition.

In Fig. 2 we present the cross-validated classification accuracy for different spatial filter choices. The results show that it was possible to discriminate between the PMA and PMR conditions significantly above chance level for all the subjects with a suitably chosen spatial filter. The average accuracy for the best performing design (with 30 channels and Laplacian spatial filter applied) was 75%. Given the small sample size, the average results should be taken with caution, but they do seem to suggest that it is beneficial to also include channels other than those over the motor cortex.

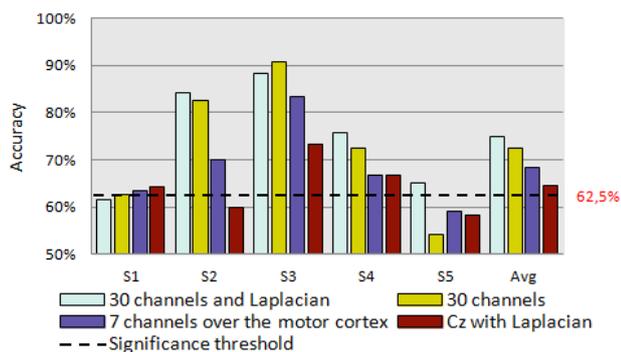


Figure 2: Mean cross-validation accuracy of classifiers with different spatial filters. The dashed line marks the significance threshold for a binomial test (at  $\alpha = 0.05$  significance level and with Bonferroni correction).

## 4 Discussion

For a sample of five subjects we found that it is possible to discriminate between attention and mind wandering during passive movement of lower limbs. With a simple BCI design based on bandpower features we were able to obtain average accuracy of 75% on single trials, in line with results for upper limbs by [1]. However, time-frequency analysis suggested different levels of SMR (de)synchronization for only two subjects.

While we tried to model realistically attention during rehabilitation, the “stop-go” nature of the trial-based experiments might not be very conducive to mind wandering. Also, unlike in our experiment, mind wandering is usually not intentional. Therefore, the experimental protocol we propose should be validated as a calibration session for an online BCI with continuous feedback. Our results also suggest that it is beneficial for classification to include features not only from the channels over the motor cortex. The question of whether the classifier is using class-specific information from other brain regions, or is using the additional features to cancel out class-unrelated noise, is left to be answered by future studies with source localization techniques. Another possibility is the existence of an uncontrolled confound in our experimental design – we find this explanation unlikely, due to the fact that the trials had the same external stimuli, and were semi-randomized.

Our future research efforts are motivated by several observations from this study: the subjects might have difficulties in complying with the protocol (inability to ignore the movement or to attend to it consistently); there might be discriminatory information in regions other than the motor cortex, and in frequency bands other than the alpha and beta bands; there is considerable variation in performance from subject to subject. We intend to address these questions, respectively, by using the proposed protocol as a calibration session for online BCI operation, by using optimized spatio-spectral features, and by analyzing subject-to-subject and session-to-session transfer of knowledge.

## Acknowledgements

This work was supported by the European Commission through the FP7 Marie Curie Initial Training Network 289146, NETT: Neural Engineering Transformative Technologies, the FP7 CORBYS IP project (GA 270219) and the Spanish Ministry of Science projects HYPER-CSD2009-00067 and DPI2011-25892. The authors would also like to thank Andreea Sburlea for valuable discussions.

## References

- [1] J. M. Antelis, L. Montesano, X. Giralt, A. Casals, and J. Minguez. Detection of movements with attention or distraction to the motor task during robot-assisted passive movements of the upper limb. In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pages 6410–6413, 2012.
- [2] R. Grandchamp and A. Delorme. Single-trial normalization for event-related spectral decomposition reduces sensitivity to noisy trials. *Frontiers in psychology*, 2:1–14, Jan. 2011.
- [3] P. Langhorne, F. Coupar, and A. Pollock. Motor recovery after stroke: a systematic review. *The Lancet Neurology*, 8(8):741–754, 2009.
- [4] S. Lemm, B. Blankertz, T. Dickhaus, and K.-R. Müller. Introduction to machine learning for brain imaging. *Neuroimage*, 56(2):387–399, 2011.
- [5] J. Schäfer 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), 2005.
- [6] J. Smallwood and J. W. Schooler. The Restless Mind. *Psychological Bulletin*, 132(6):946–958, 2006.
- [7] K. P. Tee, C. Guan, K. K. Ang, K. S. Phua, C. Wang, and H. Zhang. Augmenting Cognitive Processes in Robot-Assisted Motor Rehabilitation. In *Biomedical Robotics and Biomechanics, 2008. BioRob 2008. 2nd IEEE RAS & EMBS International Conference on*, pages 698–703, 2008.

# Motor Imagery Brain-Computer Interfaces: Random Forests vs Regularized LDA - Non-linear Beats Linear

David Steyrl, Reinhold Scherer, Oswin Förstner and Gernot R. Müller-Putz

Graz University of Technology, Institute for Knowledge Discovery, Laboratory of Brain-Computer Interfaces, Graz, Austria  
david.steyrl@tugraz.at, reinhold.scherer@tugraz.at, oswin.foerstner@student.tugraz.at,  
gernot.mueller@tugraz.at

## Abstract

Nowadays, non-linear classifiers are available that claim to generalize well at a low amount of data. Recently, we conducted an on-line study, where a random forest (RF) classifier successfully drove an electroencephalography (EEG) based sensorimotor rhythms (SMR) brain-computer interface (BCI) by classifying discrete Fourier transform (DFT) features. In this work, we re-analyse that data-set and simulate the use of common spatial patterns (CSP) features with a RF classifier and a shrinkage regularized linear discriminant analysis (sLDA). We found that the RF classifier could make better use of the CSP features and outperformed sLDA. The mean and median classification accuracy during the feedback period were improved by  $\sim 2\%$  and  $\sim 3\%$  when using a RF classifier. The effect is small, but statistically significant ( $p < 0.05$ ) and consistent over the participants. Therefore, we argue that the widespread view that linear methods are ideal for BCIs should be reconsidered and RF classifiers should be taken into account when choosing a classifier for SMR-BCIs.

## 1 Introduction

Thus far, linear machine-learning methods are considered ideal for the application in brain-computer interfaces (BCIs) [3]. Particularly with the main argument that simplicity should be preferred, especially when limited data are available as in BCIs. However, substantial progress has been made in the field of machine-learning. Nowadays, certain non-linear methods claim to generalize well when only limited amount of data are available. One such a method is the random forests (RF) classifier [2]. Our interest in this classifier is mainly based on his following properties: (1) RF classifiers provide a complex model which allows non-linear decision boundaries. (2) RF classifiers are “over-fitting” resistant, even with a large number of features. (3) RF classifiers are able to merge features originating from different statistical distribution into one model. Particularly hybrid BCIs and passive BCIs make use of such features. (4) RF classifiers are regularized by nature. (5) There exist efficient implementations of the RF classifiers which enables on-line operation. (6) RF classifiers are multi functional tools for data analysis. E.g. RF classifiers offer importance ratings of the features, they allow for analysis of features’ proximities and provide an estimate of the expected accuracy.

Recently, we conducted an on-line study, where a RF classifier was deployed in an electroencephalography (EEG) based sensorimotor rhythms BCI (SMR-BCI) [4]. Discrete Fourier transform (DFT) magnitudes were used as features for the classification. The on-line feedback results in 13 users demonstrate an classification accuracy competitive to other state of the art SMR-BCIs. Detailed results will be published elsewhere.

In the present work, we address two questions arising from the RF driven SMR-BCI: (1) The first question addresses the features for classification. Common spatial patterns (CSP)

filtering is a more powerful feature extraction method than DFT [5]. Hence, we hypothesize that replacing the DFT features by CSP features will boost the performance of the RF classifier driven SMR-BCI. (2) The second question addresses the impact of the RF classifier's non-linear model. According to literature, we hypothesize that the RF non-linear model outperforms a linear classification model although only a limited amount of data is available. The linear model is represented by an analytic-shrinkage-regularized linear discriminant analysis (sLDA) as LDA classifier are commonly used in BCIs [1]. For evaluating this two hypothesis, we conduct BCI simulations using the data of the on-line study mentioned above.

## 2 Methods

**Summary of the on-line studies set-up.** The paradigm was based on the cue-guided Graz-BCI training paradigm [4]. Hence, recording, training, and feedback was performed within a single session. The session consisted of eight runs, five of them for training and three with feedback for validation. One run was composed of 20 trials. Taken together, we recorded 50 trials per class for training and 30 trials per class for validation. Participants had the task of performing sustained (5 seconds) kinaesthetic motor imagery (MI) of the right hand and of the feet each as instructed by the cue. Feedback was presented in form of a white coloured bar-graph. The length of the bar-graph reflected the amount of correct classifications over the last second. EEG was measured with a biosignal amplifier and active Ag/AgCl electrodes (g.USBamp, g.LADYbird, Guger Technologies OG, Schiedlberg, Austria) at a sampling rate of 512 Hz. The electrodes placement was designed for obtaining three Laplacian derivations. Center electrodes at positions C3, Cz, and C4 and four additional electrodes around each center electrode with a distance of 2.5 cm, 15 electrodes total. The reference electrode was mounted on the left mastoid and the ground electrode on the right mastoid. The 13 participants were aged between 20 and 30 years, 8 naïve to the task, and had no known medical or neurological diseases.

**BCI simulation.** In this work, we want a balance between data-sets from naïve and non-naïve participants. We include all 5 data-sets of the non-naïve participants and 5 data-sets from naïve participants chosen by random. For the BCI simulation, each data-set is divided in two parts. The first part is used for CSP and classifier training, the second part for validation. The validation is carried out with a running classifier. The applied signal processing pipeline: (1) A filter bank of 8<sup>th</sup> order Butterworth band-pass-filters divides the EEG data into 15 sub-bands. Cut-off frequencies:  $[i, i + 2]$   $i = 6, 8, 10, 12$  in the  $\alpha$ -band and  $[i, i + 5]$   $i = 14, 17, 20, 23, 26, 29, 32, 35$  in the  $\beta$ -band. (2) We calculate a separate set of CSP filters for each sub-band [5]. The spatial filters according to the three highest and three lowest eigenvalues of each set of CSP filters are selected. Hence, one CSP calculation per sub-band and six filters per CSP results in 90 virtual channels. (3) The features used for classification are obtained by calculating logarithmic band-power for each of the 90 virtual channels. The logarithm changes the band-power features distribution to a normal distribution. Normal distributed features are not necessary for the RF classifier, but for the sLDA classifier. The band-power was estimated by squaring and subsequent averaging over a sliding window with a length of 1 s. (4) The classification was performed with a RF classifier on the one hand, and with a sLDA classifier on the other hand.

For training, we picked the features from the 1 s long window starting 2.5 s after the cue of each trial [6]. This implies a trials-to-features ratio of  $100/90 = 1.11$ . For validation, we performed a separate classification on each time point of each trial to obtain one course of classification accuracy per participant.

**Random Forests and analytic-shrinkage-regularized linear discriminant analysis.**

RF denotes for an ensemble classifier comprising of many decision trees. The decision trees are decorrelated by random processes during their construction. A majority voting of the trees defines the forests' decision. The voting is an important step as it reduces the variance of the forest which is commonly high for individual trees. This is a kind of regularization and improves the accuracy of a forests dramatically when compared with any single decision tree [2]. Due to our experience with the RF classifier, we chose to build 1000 trees per classifier and used the standard value for randomly drawn features per node ( $\sqrt{\#of\ features}$ ).

A comparison of the RF classifier with a non-regularized classifier is unfair since the RF classifier is regularized by nature. Hence, we chose an sLDA classifier for comparison. Shrinkage is a common remedy for achieving well conditioned covariance matrices even when the data is high-dimensional and only a few data points are given. For further information on sLDA, please see [1].

### 3 Results

For each participant, the peak, mean and median accuracies during the feedback period were calculated and are presented in Table 1. Peak means highest accuracy during the feedback period. Mean refereed the mean accuracy over the feedback period and median stands for the median accuracy over the feedback period. The approaches using CSP features significantly outperformed the approach using DFT features in terms of peak (82% < 89.67%, 87.83%;  $p < 0.05$ ), mean (66.82% < 79.30%, 77.15%;  $p < 0.01$ ), and median (67.67% < 80.42%, 77.83%;  $p < 0.01$ ) performance (paired  $t$ -tests, Bonferoni-Holm correction). The combination of CSP features with a RF classifier significantly outperformed the combination of CSP features with sLDA in terms of mean (79.3% > 77.15%;  $p < 0.05$ ), and median (80.42% > 77.83%;  $p < 0.05$ ) performance (paired  $t$ -test, Bonferoni-Holm correction).

ID	naive?	Online DFT+RF			Simulation CSP+RF			Simulation CSP+sLDA		
		peak	mean	median	peak	mean	median	peak	mean	median
P1	no	71.67	56.14	56.67	<b>81.67</b>	<b>74.56</b>	<b>76.67</b>	78.33	69.69	70.00
P2	no	86.67	67.27	68.33	<b>91.67</b>	<b>80.52</b>	<b>82.50</b>	90.00	79.22	80.00
P3	no	<b>100.00</b>	90.71	91.67	<b>100.00</b>	<b>99.30</b>	<b>100.00</b>	<b>100.00</b>	99.22	<b>100.00</b>
P4	no	76.67	64.39	65.00	<b>95.00</b>	<b>81.85</b>	<b>81.67</b>	93.33	79.09	77.50
P5	no	80.00	59.81	60.00	<b>81.67</b>	<b>65.44</b>	<b>65.00</b>	78.33	63.10	61.67
P6	yes	93.33	82.79	83.33	<b>96.67</b>	<b>88.46</b>	<b>88.33</b>	95.00	84.35	85.00
P7	yes	96.67	83.25	86.67	98.33	88.33	<b>92.50</b>	<b>100.00</b>	<b>88.61</b>	<b>92.50</b>
P8	yes	83.33	66.54	66.67	<b>95.00</b>	<b>83.28</b>	<b>85.83</b>	88.33	78.93	81.67
P9	yes	66.67	47.35	48.33	<b>88.33</b>	<b>76.64</b>	<b>78.33</b>	83.33	73.10	73.33
P10	yes	65.00	49.99	50.00	68.33	54.61	53.33	<b>71.67</b>	<b>56.22</b>	<b>56.67</b>
average		82.00	66.82	67.67	<b>89.67*</b>	<b>79.30***</b>	<b>80.42***</b>	87.83*	77.15**	77.83**
std		12.27	14.63	15.11	9.87	12.56	13.40	9.69	12.42	13.20

Table 1: Binary validation accuracies of the different BCI systems in %. Best performing method per participant is highlighted. \* significantly better than DFT+RF ( $p < 0.05$ ). \*\* significantly better than DFT+RF ( $p < 0.01$ ). \* significantly better than CSP+sLDA ( $p < 0.05$ ).

## 4 Discussion and Conclusion

By using CSP features instead of DFT features, the average peak, mean and median performance during the feedback period was significantly improved from 82% to 89.7%, from 66.8% to 79.3% and from 67.7% to 80.4%, respectively when combined with a RF classifier and from 82% to 87.3%, from 66.8% to 77.2% and from 67.7% to 77.8%, respectively when combined with an sLDA classifier. It is not surprising that an optimized spatial filtering outperforms the DFT features. CSP features have a higher signal-to-noise ratio and are therefore easier to classify. For example, the peak classification accuracy of one participant was improved by  $\sim 22\%$  (Table 1, P9). However, our results show that a RF classifier can make better use of CSP features than an sLDA classifier, at least for the present data. This is remarkable, as the RF classifier relies on a complex, non-linear model and the trials-to-features ratio is low ( $100/90 = 1.11$ ). The effect of using a RF classifier instead of a sLDA classifier is small (peak  $\sim 2\%$ , mean  $\sim 2\%$ , median  $\sim 3\%$ ), but statistically significant for mean and median performance and consistent over the participants. For 8 of the 10 participants the combination of a RF classifier with CSP features is the best performing method. For participant P7 the combination of sLDA with CSP features performed slightly better, but not in median performance. For participant P10, both methods failed, since the achieved performance is around 70% only. The enhancement of the sustained (i.e. mean and median) performance is of particular importance, as the Graz-BCI paradigm calls for these. Concluding, the present work on performance, in combination with previous work on RF classifiers as powerful tools for data analysis [7], underlines the potential of the RF classifier in the field of BCIs. Further, we argue that the widespread view that linear methods are ideal for BCIs should be reconsidered.

## Acknowledgements

This work is partly supported by the FP7 research projects BackHome (No. 288566) and ABC (No. 287774). This paper only reflects the authors' views and funding agencies are not liable for any use that may be made of the information contained herein.

## References

- [1] B. Blankertz, S. Lemm, M. Treder, S. Haufe, and K.-R. Müller. Single-trial analysis and classification of ERP components - a tutorial. *Neuroimage*, 56:814–825, 2011.
- [2] L. Breiman. Random forests. *Mach. Learn.*, 45:5–32, 2001.
- [3] K. R. Müller, C. W. Anderson, and G. E. Birch. Linear and nonlinear methods for brain-computer interfaces. *IEEE Trans. Neural Syst. Rehabil. Eng.*, 11:165–9, 2003.
- [4] G. Pfurtscheller and C. Neuper. Motor imagery and direct brain-computer communication. *Proc. IEEE*, 5(89):1123–1134, 2001.
- [5] H. Ramoser, J. Müller-Gerking, and G. Pfurtscheller. Optimal spatial filtering of single trial EEG during imagined hand movement. *IEEE Trans. Rehab. Eng.*, 8(4):441–446, 2000.
- [6] R. Scherer, G. Pfurtscheller, and C. Neuper. Motor imagery induced changes in oscillatory EEG components: speed vs. accuracy. In *Proc. of the 4th International Brain-Computer Interface Workshop and Training Course 2008*, pages 186–190, Graz, Austria, 2008. Verlag der Technischen Universität Graz.
- [7] D. Steyrl, R. Scherer, and G. R. Müller-Putz. Random forests for feature selection in non-invasive brain-computer interfacing. In *HCI-KDD 2013, Lecture Notes in Comput. Sci., 7947*, pages 207–216, Berlin Heidelberg, 2013. Springer.

# Improved Classification of Auditory Evoked Event-Related Potentials

Günther Bauernfeind<sup>1,6</sup>, Christoph Pokorny<sup>1,6</sup>, David Steyr<sup>1,6</sup>,  
Selina C. Wriessnegger<sup>1,6</sup>, Gerald Pichler<sup>2,6</sup>, Walter Schippinger<sup>2,6</sup>, Quentin  
Noirhomme<sup>3</sup>, Ruben G. Real<sup>4</sup>, Andrea Kübler<sup>4</sup>, Donatella Mattia<sup>5</sup> and  
Gernot R. Müller-Putz<sup>1,6</sup>

<sup>1</sup> Institute for Knowledge Discovery, Graz University of Technology, 8010 Graz, Austria

<sup>2</sup> Albert Schweitzer Clinic, 8020 Graz, Austria

<sup>3</sup> Cyclotron Research Center, University of Liège, 4000 Liège, Belgium

<sup>4</sup> Department of Psychology, University of Würzburg, 97070 Würzburg, Germany

<sup>5</sup> Fondazione Santa Lucia IRCCS, 00179 Rome, Italy

<sup>6</sup> BioTechMed-Graz, Austria

[g.bauernfeind@tugraz.at](mailto:g.bauernfeind@tugraz.at)

## Abstract

In this study we report on the improvement of classification accuracy in an auditory P300 paradigm, by using stepwise linear discriminant analysis (SWLDA) with an increased number of channels and analytic shrinkage-regularized LDA (sLDA) as classifier. The investigations were evaluated on recordings of 10 healthy subjects and 12 patients in a minimally conscious state. The results for healthy subjects were promising, significant improvements could be found by increasing the number of channels using SWLDA, as well as, for using sLDA instead of SWLDA. Single trial classification accuracies up to 85.1% could be achieved for healthy subjects. For the patients the results were less promising. However, for one patient improvement reaching up to an accuracy of 71.3% could be achieved.

## 1 Introduction

A promising approach, when considering the application of Brain-Computer Interface (BCI) systems to patients diagnosed with minimally conscious state (MCS), is the use of single-switch BCIs (ssBCIs) [6]. Such an ssBCI can be based, for example, on motor imagery or on responses to visual, tactile or auditory stimulation. However, while the visual ability might be considerably impaired in such class of patients, the auditory pathway is usually preserved. Therefore, the auditory system might be one of the last remaining channels usable for BCI-based communication [3]. Recently we proposed the concept of a novel auditory single-switch BCI and investigated the transition of a paradigm from healthy subjects (HS) to patients (PA) in MCS [4]. The paradigm was evaluated in 10 HS and applied to 12 PA. In 8 of the 10 HS significant single-trial classification accuracies up to 77.2 % could be reached using stepwise linear discriminant analysis (SWLDA). However, for MCS patients only a small number of classification results were above chance level and none of the results were sufficient for communication purposes. We speculated that the inclusion of all recorded channels instead of only three pre-selected channels might yield better classification results.

In the present work, we investigated this issue by using all recorded channels for SWLDA classification in HS and PA. However, due to the low number of trials available from patient measurements, there is a risk of overfitting the data when using all channels for classification. Therefore, we also used analytic shrinkage-regularized LDA (sLDA) which clearly outperforms SWLDA at low trial to feature ratios [1].

## 2 Methods

**Subjects, Experimental Paradigm and Data Recording:** We used the data of the 10 HS (mean age  $27.6 \pm 3.0$  SD years) and 12 PA ( $45.8 \pm 18.2$  years) in MCS, recorded in our previous study, using an auditory based P300 BCI [4]. Briefly summarized, two tone streams (low, LTS, at 396 Hz, and high, HTS, at 1900 Hz) with infrequently appearing deviant tones (297 Hz for the LTS and 2640 Hz for the HTS) at random positions were presented. The tones of both streams were intermixed (LHL\_LHL-, L..low tone, H..high tone, ...silent gap; for details and a schematic illustration see [4]). As in this way the LTS was twice as fast as the HTS, the percentage of deviant tones was different (20% for HTS and 10% for LTS) to generate the same absolute number of deviants in both streams. For the experiment, subjects were instructed visually, for HS, or auditory, for PA, to focus attention on one of the streams during one run. For the HS 80 runs (40 runs for each stream; with 4000/400 (standard/deviant) tones in the LTS and 1800/400 in the HTS) were recorded. Taking into account the reduced attention span of the patients, one to two sessions, each with 20 runs (10 for each stream), were recorded. In HS the EEG was recorded at 15 positions with a sampling rate of 512 Hz (filter setup: 0.5–100 Hz) using active electrodes. For PA recording a reduced channel set (9 positions) was used to facilitate measurements in a clinical environment. For the electrode setup see [4].

**Data Analysis and Classification:** Data recorded from HS and PA were analyzed in the same way. Raw signals were filtered with a  $3^{rd}$  order Butterworth low-pass filter (cut-off frequency at 10 Hz) and down-sampled to 64 Hz. In our previous work three EEG channels (Fz, Cz, Pz) were used for classification (SWLDA, with 10x10 cross-validation; for details see [4]). Only time points between 200 and 800 ms after tone onset were used as features.

In the present work, we performed two classification approaches: Firstly we used the data of all recorded channels for SWLDA classification. Secondly we investigated also the performance of sLDA instead of the SWLDA. In general, the classification performance of an LDA crucially depends on accurate estimates of the class means and the common covariance matrix. If enough data are available, an accurate estimate is unproblematic. However, in cases where insufficient data are available or the data are contaminated by outlier (e.g. during PA recordings), conventional estimation of the covariance matrix fails. In those cases, the covariance matrix is ill conditioned. One can improve the conditioning of the covariance matrix by regularization. An efficient regularization method is analytic shrinkage. This so-called analytic-shrinkage-regularized LDA (sLDA) is computationally very efficient and outperforms other methods at low trial to feature ratios [1]. For both approaches the classification was carried out as follows: 1) To detect the P300 the deviant tones were classified against the standard tones for each target stream separately, and 2) to investigate the attentional modulation of the P300, the target deviant tones were classified against the non-target deviant tones for each stream separately. By performing the second investigation, it should be possible to infer which stream was attended. For the first investigation random subsampling with 100 iterations was applied to account for the very different numbers of deviant and standard segments. Classification results were compared with the real level of chance [2] to identify random results.

## 3 Results

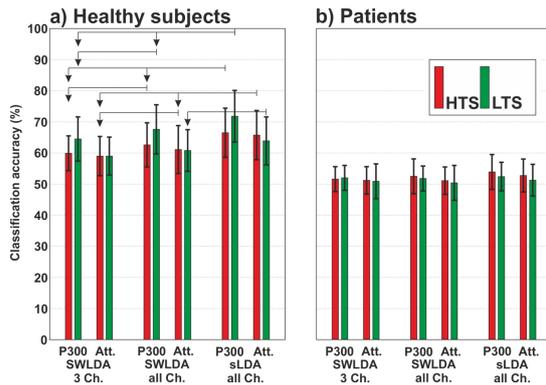
Results are shown for P300 and attentional modulation detection. Figure 1 depicts the averaged (mean  $\pm$  SD) classification accuracies for P300 and attentional modulation detection, using either SWLDA with 3 channels (SWLDA<sub>3</sub>), all channels (SWLDA<sub>all</sub>), or sLDA classifier with

all channels (sLDA<sub>all</sub>) for both groups (HS,PA). The results are shown for LTS and HTS. Table 1 summarizes the subject specific SWLDA and sLDA classification results of all 10 HS. The values in the table represent the mean accuracies over all cross-validation folds.

Subj.	P300						Attention					
	SWLDA <sub>3</sub>		SWLDA <sub>all</sub>		sLDA <sub>all</sub>		SWLDA <sub>3</sub>		SWLDA <sub>all</sub>		sLDA <sub>all</sub>	
	LTS	HTS	LTS	HTS	LTS	HTS	LTS	HTS	LTS	HTS	LTS	HTS
HS01	<i>57.3</i>	<i>61.2</i>	<i>58.0</i>	<i>62.5</i>	<b>64.0</b>	<b>68.2</b>	55.2	57.8	55.8	56.0	<b>58.6</b>	<b>59.3</b>
HS02	<i>70.7</i>	<i>61.8</i>	<i>73.5</i>	<i>63.6</i>	<b>76.7</b>	<b>65.2</b>	<i>62.6</i>	<i>62.8</i>	<i>61.9</i>	<i>62.9</i>	<b>65.3</b>	<b>69.4</b>
HS03	<i>77.2</i>	<i>72.0</i>	<i>78.2</i>	<i>73.3</i>	<b>85.1</b>	<b>77.7</b>	<i>69.5</i>	<i>70.4</i>	<i>70.9</i>	<i>69.2</i>	<b>75.9</b>	<b>74.8</b>
HS04	<i>62.2</i>	<i>57.2</i>	<i>64.3</i>	<i>57.9</i>	<b>67.2</b>	<b>61.4</b>	<i>56.9</i>	53.5	<i>59.4</i>	<i>57.4</i>	<b>61.1</b>	<b>61.2</b>
HS05	<i>63.9</i>	<i>60.7</i>	<i>65.2</i>	<i>61.5</i>	<b>69.2</b>	<b>65.8</b>	<i>63.5</i>	<i>61.0</i>	<i>65.4</i>	<i>65.5</i>	<b>67.2</b>	<b>71.2</b>
HS06	<i>71.5</i>	<i>64.4</i>	<i>78.7</i>	<i>75.0</i>	<b>81.3</b>	<b>79.7</b>	<i>65.5</i>	<i>65.5</i>	<i>71.6</i>	<i>76.7</i>	<b>76.5</b>	<b>78.8</b>
HS07	<i>61.9</i>	54.4	<i>62.9</i>	55.6	<b>67.2</b>	<b>59.8</b>	53.3	55.6	54.1	54.6	<b>56.7</b>	<b>58.1</b>
HS08	<i>67.2</i>	<i>59.7</i>	<i>74.0</i>	<i>65.1</i>	<b>80.7</b>	<b>72.1</b>	<i>58.9</i>	<i>60.2</i>	<i>61.0</i>	<i>62.2</i>	<b>65.9</b>	<b>69.3</b>
HS09	<i>56.8</i>	55.4	<i>59.6</i>	<i>57.6</i>	<b>62.6</b>	<b>58.6</b>	52.8	52.1	53.4	53.0	<b>54.7</b>	<b>55.9</b>
HS10	56.0	52.6	<i>60.9</i>	53.6	<b>64.4</b>	<b>56.7</b>	51.4	50.6	51.4	53.4	<b>57.0</b>	<b>60.3</b>
Mean	64.5	59.9	67.5	62.6	71.8	66.5	59.0	59.0	60.8	61.1	63.9	65.8
SD	7.1	5.6	7.8	7.1	8.3	7.9	6.1	6.3	6.7	7.7	7.7	7.9

**Table 1:** SWLA and sLDA classification accuracies (in %) for the HS for P300 and attentional modulation detection using either 3 or all channels. All results significantly better than random [2] ( $\alpha = 1\%$ ) are indicated in *italic*. Highest subject specific accuracies are indicated in **bold**.

In healthy subjects the P300 condition revealed a statistically significant difference in accuracy depending on which classification method was used,  $\chi^2_{(5)} = 39.94, p < 0.001$ . Post hoc analysis with Wilcoxon signed-rank tests was conducted with a Bonferroni correction applied, resulting in a significance level set at  $p \leq 0.012$ . There were significant differences between the SWLDA<sub>3</sub> and sLDA<sub>all</sub> (LTS:  $Z = -2.805, p = 0.005$ ; HTS:  $Z = -2.803, p = 0.005$ ) and SWLDA<sub>all</sub> and sLDA<sub>all</sub> (LTS:  $Z = -2.803, p = 0.005$ ; HTS:  $Z = -2.803, p = 0.005$ ). The attentional modulation condition showed also a significant difference in classification accuracies depending on the method used, ( $\chi^2_{(5)} = 38.11, p < 0.001$ ) and the follow up Wilcoxon signed-rank test showed significant differences between SWLDA<sub>all</sub> and sLDA<sub>all</sub> (LTS:  $Z = -2.805, p = 0.005$ ; HTS:  $Z = -2.803, p = 0.005$ ) and between SWLDA<sub>3</sub> and SWLDA<sub>all</sub> (LTS:  $Z = -2.49, p = 0.012$ ).



**Figure 1:** Mean ( $\pm$  SD) classification accuracies (in %) for a) healthy subj. and b) patients (for P300 and attentional modulation detection; separated for LTS and HTS) using SWLDA and sLDA classifier. Significant differences are indicated.

	P300					
	SWLDA <sub>3</sub>		SWLDA <sub>all</sub>		sLDA <sub>all</sub>	
	LTS	HTS	LTS	HTS	LTS	HTS
Mean	51.9	51.5	51.7	52.4	52.3	53.8
SD	4.0	3.9	4.1	4.2	4.6	5.6

	Attention					
	SWLDA <sub>3</sub>		SWLDA <sub>all</sub>		sLDA <sub>all</sub>	
	LTS	HTS	LTS	HTS	LTS	HTS
Mean	50.8	51.1	50.3	51.0	51.2	52.7
SD	5.6	4.4	5.6	4.4	5.1	5.3

**Table 2:** Mean ( $\pm$  SD) SWLDA and sLDA classification accuracies (in %) for the patient group for P300 and attentional modulation detection using either 3 or all channels.

For the patients (see Table 2 for mean  $\pm$  SD accuracies) the Friedman test reported significant differences between the classification methods only for the P300 condition,  $\chi^2_{(5)} = 13.28, p = 0.021$ . The follow up Wilcoxon signed-rank test revealed no significant results. Nevertheless, in one of the patients (PA09; CRS-r score: 18; Cause: Hemorrhagic stroke) a P300 could be classified above chance level. In more detail, by using sLDA, the accuracy for the HTS was improved up to 71.3% (SWLDA<sub>3</sub>: 63.0%; SWLDA<sub>all</sub>: 65.3%), which is significantly better than random [2] ( $\alpha = 1\%$ ). However, for the LTS all classification accuracies (SWLDA<sub>3</sub>: 64.8%; SWLDA<sub>all</sub>: 65.0%; sLDA<sub>all</sub>: 65.2%) remained below chance level.

## 4 Discussion and Conclusion

For the HS group the classification accuracies could be significantly improved by using sLDA and the inclusion of all recorded channels. However, unlike healthy subjects, for the patient group the results were less encouraging. Although in some PA the single-trial classification accuracies could be improved, they remained mainly below chance level. However, an accuracy significantly better than chance level [2] could be reached in one subject (71.3%).

Concluding, as stated by Blankertz et al. [1], "the use of a higher number of channels is potentially advantageous for ERP classification" and improves the classification accuracy in HS using either SWLDA or sLDA. For the PA recordings further studies should aim on investigating the use of additional electrode positions (especially frontal/fronto-lateral [5]). Furthermore, as concluded in [4], improvements on the used paradigm (e.g. include spatial information [5] if possible) are still required to take into account the specific needs and capabilities of the patients.

## 5 Acknowledgments

Supported by European ICT Programme Project FP7-247919 and "Land Steiermark" Project (ABT08-22-T-7/2013-17). The text reflects solely the views of its authors. The European Commission is not liable for any use that may be made of the information contained therein.

## References

- [1] B. Blankertz, S. Lemm, M. Treder, S. Haufe, and K.-R. Müller. Single-trial analysis and classification of erp components - a tutorial. *Neuroimage*, 56:814–825, 2011.
- [2] G. R. Müller-Putz, R. Scherer, C. Brunner, R. Leeb, and G. Pfurtscheller. Better than random? A closer look on BCI results. *Int. J. Bioelectromag.*, 10:52–55, 2008.
- [3] A. R. Murguialday, J. Hill, M. Bensch, S. Martens, S. Halder, F. Nijboer, B. Schoelkopf, N. Birbaumer, and A. Gharabaghi. Transition from the locked in to the completely locked-in state: a physiological analysis. *Clin Neurophysiol*, 122(5):925–933, 2011.
- [4] C. Pokorny, D. S. Klobassa, G. Pichler, H. Erlbeck, R. G. L. Real, A. Kübler, D. Lesenfants, D. Habbal, Q. Noirhomme, M. Riseti, D. Mattia, and G. R. Müller-Putz. The auditory P300-based single-switch BCI: Paradigm transition from healthy subjects to minimally conscious patients. *Artif Intell Med*, 59(2):81–90, 2013.
- [5] M. Schreuder, B. Blankertz, and M. Tangermann. A new auditory multi-class brain-computer interface paradigm: Spatial hearing as an informative cue. *PLoS ONE*, 5(4):e9813, 2010.
- [6] C. Zickler, A. Riccio, F. Leotta, S. Hillian-Tress, S. Halder, E. Holz, P. Staiger-Sälzer, E. J. Hoogerwerf, L. Desideri, D. Mattia, and A. Kübler. A brain-computer interface as input channel for a standard assistive technology software. *Clin EEG Neurosci.*, 42(4):236–244, 2011.

# Transferring Unsupervised Adaptive Classifiers Between Users of a Spatial Auditory Brain-Computer Interface

Pieter-Jan Kindermans<sup>1</sup>, Benjamin Schrauwen<sup>1</sup>, Benjamin Blankertz<sup>2</sup>,  
Klaus-Robert Müller<sup>3,4</sup> and Michael Tangermann<sup>5</sup>

<sup>1</sup> Electronics and Information Systems (ELIS), Ghent University, Ghent, Belgium  
`pieterjan.kindermans@UGent.be`

<sup>2</sup> Technical University of Berlin, Dept. Neurotechnology, Berlin, Germany

<sup>3</sup> Technical University of Berlin, Dept. Machine Learning, Berlin, Germany

<sup>4</sup> Department of Brain and Cognitive Engineering, Korea University, Seoul, Republic of Korea

<sup>5</sup> BrainLinks-BrainTools Excellence Cluster, University of Freiburg, Freiburg, Germany

## Abstract

The transfer of knowledge allows to rapidly set up a functioning BCI system for a novel user without user specific calibration. For spelling paradigms based on event-related potentials (ERP) of the EEG, a recent unsupervised classification method is able to train the classifier online by constantly adapting to the user on the fly. Hence, an explicit calibration phase can be avoided. However, due to the random initialization of the classifier, this method requires a warm-up time before it performs on the same level as a supervised trained classifier. This warm-up effect can be reduced by transfer learning. We present a thorough leave-one-user-out offline analysis (n=9 users) and additional preliminary online results from a spatial auditory ERP spelling study (AMUSE paradigm) on inter-subject transfer of an unsupervised adaptive classifier. For the online study, a classifier trained on data of n=8 previous users was transferred to two unseen users and further adapted online. The performance was evaluated in one online copy spelling session per user. Both, the offline simulations and the online results indicate that the transfer approach reduces the warm-up time by approx. 50 %.

## 1 Introduction

The auditory event-related potential (ERP) response of the electroencephalogram (EEG) is modulated by attention. This gave rise for the design of spatial auditory brain-computer interface (BCI) systems for communication [7, 1] to complement visual approaches.

Despite major improvements on paradigms and data analysis methods, one of the key limitations in BCI remains the dependency on a calibration session. Providing labeled data points, this recording is the prerequisite to train the EEG decoder with supervised machine learning methods. Unfortunately, this calibration session limits the time available to use a BCI online, which is a major problem for patients in need of such a BCI [5]. This problem arises especially for paradigms with long trial durations like motor imagery paradigms, and for those with intrinsic lower signal-to-noise ratio (SNR) like tactile or auditory ERP paradigms.

Hence, the BCI community has spent a considerable amount of effort to reduce the need for calibration data. In the field of ERP-based BCI systems, transfer learning approaches, which exploit information collected from previous users to facilitate decoding for a novel user, have been studied. The information can be transferred in the form of data [2] or as pre-trained classifiers [6]. The classifiers can be distinguished by their dependency onto labeled data from previous calibration recordings (supervised trained classifiers, e.g. [6]), or if unlabeled

online data is sufficient to pre-train them (unsupervised trained classifiers, see [4]). These basic transfer learning methods lead to a reasonably accurate classifier for novel data. However, transferred models typically can not compete with user-specific models. To tackle this problem, a combination of transfer learning and online unsupervised adaptation has been proposed. While this combination has already been investigated offline [4] based on high-SNR visual ERP data, a test with more challenging data and its application in online experiments is still lacking. Hence, this work contributes with offline results on challenging auditory ERP data and, more importantly, by presenting preliminary online results with the spatial auditory AMUSE [7] paradigm.

## 2 Methods

### 2.1 Auditory ERP Paradigm AMUSE

AMUSE utilizes six auditory stimuli originating from a ring of six speakers surrounding the user. The stimuli can be identified by their direction or pitch, have a duration of 40 ms and were presented in pseudo-randomized order with a stimulus onset asynchrony (SOA) of 175 ms and in 15 iterations per trial. To spell one of 36 possible symbols, the user selects a group of six symbols in a first step by focusing attention to one of the six stimuli. Subsequently the user can select one of the six symbols from the previously selected group. Contrary to [7], a correction of wrong letters was not foreseen. Nevertheless, for all other details we closely replicated the setup detailed by Schreuder and colleagues.

### 2.2 Unsupervised Transfer Learning

We evaluated the unsupervised transfer learning model from [4]. This probabilistic model is applicable for multi-stimulus ERP paradigms. It assumes that only a single one of the six stimuli can result in a target ERP response, while the other stimuli must elicit a non-target response. On top of that, it assumes that EEG features can be projected into a single dimension, where the projected data is Gaussian with class-dependent mean and shared variance.

These assumptions allow for an unsupervised training of the model. Hence a calibration recording and labeled data are no longer required. The training is based on the Expectation Maximization algorithm and allows the model to learn how to decode ERP features while the user is interacting with the BCI. A known limitation is the initially unreliable classification, which is caused by a random initialization. Therefore, even though this model is an improvement over supervised systems with a calibration recording, this so-called warm-up period may limit the usability of the unsupervised system in the presence of tight time constraints.

To alleviate the warm-up phase, we have extended the model with inter-subject transfer learning [4]. The concept of transfer learning is introduced in the model by placing a shared hyper-prior on the user-specific mean of the prior on the weight vector. This corresponds to regularizing the subject-specific model towards the general shared model in place of regularizing towards a zero vector, which is a common approach to limit the model complexity. Due to space limitations, we refer the reader to [4] for the update equations and the training procedure of the transfer learning model.

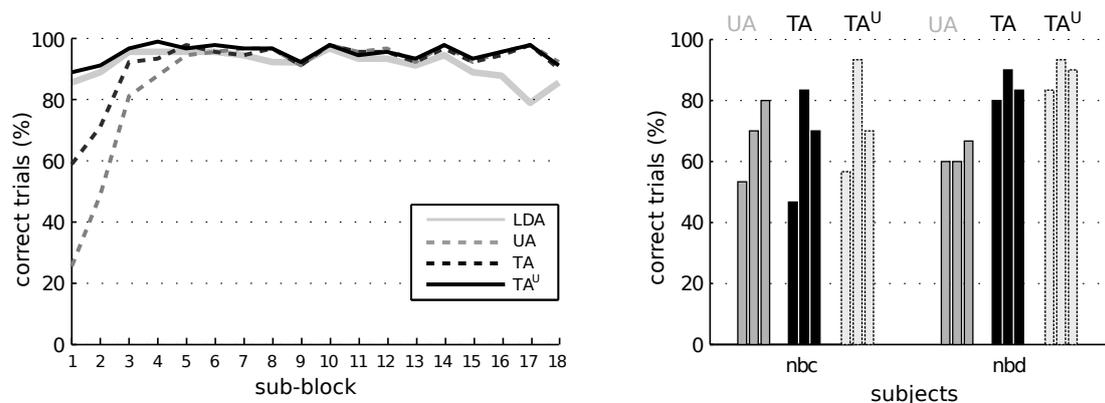


Figure 1: **Left:** Simulated grand average performance over 18 sub-blocks of ten trials each. **Right:** Online selection accuracy for two users and three blocks of 30 trials each.

### 2.3 Experiments and Data

An offline simulation was conducted on existing data from [3] to assess the feasibility of unsupervised adaptive transfer learning for auditory ERP data with challenging SNR. In the original study, unsupervised learning was compared to a standard supervised approach (LDA with shrinkage-regularization on the estimated covariance matrix). Each user had performed a calibration recording comprising 30 trials and six online evaluation blocks of 30 trials each. Supervised decoding was used for three blocks, and unsupervised for the remaining three blocks, resulting in 180 trials. For the simulation in the present work, this data was re-used trial-by-trial to train up a subject-specific unsupervised adaptive model per user ( $UA$ ). It is compared with the fully updated unsupervised model (having seen all 180 trials) after processing the entire dataset ( $UA^U$ ). Furthermore, data is combined in a leave-one-user-out transfer learning model. The simulated trial-by-trial application and updating of this transfer model to data of the left-out user is denoted  $TA$ . The final, fully updated version of this model is  $TA^U$ .

The novel online transfer experiment made use of  $TA$ . Two users of the original study were re-invited for a second session. They each performed three blocks of 30 trials of online copy-spelling spelling using the transfer learning model  $TA$ . Please note that it was reset at the beginning of each block in order to study the warm-up behavior. In addition, the fully updated model  $TA^U$  was available at the end of each block. In addition, the performance of a non-transferred, randomly initialized unsupervised classifier  $UA$  was evaluated by a post-hoc simulation. We would like to stress that the transferred models used in this work all are completely unsupervised. Label information was not used at any point in the models' training. Hence, it is sufficient to use data from free spelling sessions to build the transfer model.

## 3 Results and Discussion

Results from the simulated experiment on nine users are given in the left plot of Figure 1. Offline performance is estimated on sub-blocks of ten trials (corresponding to five symbols each). As a baseline, the standard supervised LDA model (trained on data from a calibration session of 30 trials in the original study) is compared to the unsupervised models  $UA$ ,  $TA$  and the fully adapted model at the end of the session  $TA^U$ .

During the first two sub-blocks the subject-specific supervised LDA model performs best. Furthermore it is relatively stable over most sub-blocks, but it displays a performance drop at the end of the experiment (sub-blocks 17 and 18), while this is not observed for any of the unsupervised adapted models. The randomly initiated unsupervised model  $UA$  exhibits typical warm-up behavior during the first two experimental blocks (25 % and 47 % accuracy, with 16.6 % chance level). As it observes more and more data, it becomes more reliable. The transfer learning model  $TA$  was able to significantly reduce the warm-up period and achieves 60 % selection accuracy in the first block and 70 % in the second block. While it does not yet performs at the same level as the supervised LDA model, we must point out that the LDA model had already observed ten trials more data at this point (30 trials in total). Moreover, the re-evaluation of the transfer learning model  $TA^U$  can correct many of the initial mistakes committed by  $TA$  and is as reliable as the LDA model even in the first blocks. In previous work, we have shown that the revised predictions of the  $UA^U$  model are quite similar to those of  $TA^U$ . The online results given in the right plot of Figure 1 provide details on the warm-up performance for three repeated blocks of 30 trials (corresponding to 15 symbols each). They not only show the technical feasibility of the transfer approach, but are a first indication that the simulation results reported above can be replicated online. Obviously, the online results so far are limited by the fact that they comprise two users only.

However, the offline simulation in combination with these first online results of unsupervised transfer learning indicate that it is not only useful to share decoding knowledge between users, but that an additional subject-specific further online adaptation quickly leads to a very reliable model for the novel user. This study will be extended in future work and naturally, if successful, taken to patients.

## Acknowledgment

This work was partly supported by: BMBF Germany (FKZ 01GQ0850, 16SV5839); Ghent Univ. Special Res. Fund (BOF-GOA project Home-MATE); German Research Found. (grant number EXC 1086); Nat. Res. Found. of Korea. We thank Johannes Höhne and Martijn Schreuder for support with the experimental setup.

## References

- [1] J. Höhne, M. Schreuder, B. Blankertz, and M. Tangermann. A novel 9-class auditory ERP paradigm driving a predictive text entry system. *Front. Neuroscience*, 5:99, 2011.
- [2] J. Jin, E.W. Sellers, Y. Zhang, I. Daly, X. Wang, and A. Cichocki. Whether generic model works for rapid ERP-based BCI calibration. *J. Neurosc. Meth.*, 2012.
- [3] P.-J. Kindermans, M. Schreuder, B. Schrauwen, K.-R. Müller, and M. Tangermann. True zero-training brain-computer interfacing — an online study. *PLoS One*, (in press).
- [4] P.-J. Kindermans, M. Tangermann, K.-R. Müller, and B. Schrauwen. Integrating dynamic stopping, transfer learning and language models in an adaptive zero-training ERP speller. *J. Neural Eng.*, 2014.
- [5] K. Li, V. N. Raju, R. Sankar, Y. Arbel, and E. Donchin. Advances and challenges in signal analysis for single trial P300-BCI. In *Proc. 6th Int. Conf. Foundations of Augm. Cogn.*, FAC'11, pages 87–94, Berlin, Heidelberg, 2011. Springer-Verlag.
- [6] Y. Li, C. Guan, H. Li, and Z. Chin. A self-training semi-supervised SVM algorithm and its application in an EEG-based brain computer interface speller system. *Pattern Recogn. Lett.*, 29(9):1285 – 1294, 2008.
- [7] M. Schreuder, T. Rost, and M. Tangermann. Listen, you are writing! speeding up online spelling with a dynamic auditory BCI. *Front. Neuroscience*, 5, 2011.

# Latency correction of error-related potentials reduces BCI calibration time

Iñaki Iturrate<sup>1,2</sup>, Ricardo Chavarriaga<sup>1</sup>, Luis Montesano<sup>2</sup>,  
Javier Minguez<sup>2,3</sup> and José del R. Millán<sup>1</sup>

<sup>1</sup> Chair in Non-Invasive Brain-Machine Interface (CNBI), EPFL, Switzerland  
[inaki.iturrate@epfl.ch](mailto:inaki.iturrate@epfl.ch), [ricardo.chavarriaga@epfl.ch](mailto:ricardo.chavarriaga@epfl.ch), [jose.millan@epfl.ch](mailto:jose.millan@epfl.ch)

<sup>2</sup> Instituto de Investigación en Ingeniería de Aragón (I3A), DIIS, Universidad de Zaragoza, Spain  
[montesano@unizar.es](mailto:montesano@unizar.es), [jminguez@unizar.es](mailto:jminguez@unizar.es)

<sup>3</sup> Bit&Brain Technologies SL, 50018 Zaragoza, Spain

## Abstract

Calibration of brain-machine interfaces exploiting event-related potentials has to be performed for each experimental paradigm. Even if these signals have been used in previous experiments with different protocols. We show that use of signals from previous experiments can reduce the calibration time for single-trial classification of error-related potentials. Compensating latency variations across tasks yield up to a 50% reduction the training period in new experiments without decrease in online performance compared to the standard training.

## 1 Introduction

Successful decoding of event-related potentials (ERP) for brain-machine interfacing requires adequate models of the signal of interest. Considering the variability of EEG signals, calibration of these models is done through the acquisition of a large number of trials. Therefore, a considerable amount of time has to be spent before a system can be operated in online manner. Different approaches have been proposed to overcome this issue by applying adaptive classifiers [7] or using previous information from multiple subjects [5, 6].

Remarkably, the recalibration process has to be performed for every protocol, even if ERPs elicited by the same cognitive processes have previously been used with other experimental setups (e.g. different feedback stimuli or final application). Recent works have tried to exploit ERP similarities in these cases [2, 4]. For instance, it has been shown that variations of error-related potentials (ErrP) across different experimental protocols can be largely explained by changes in their latency [2]. We claim that these variations can be compensated in order to exploit available data from previous experiments for the calibration of new experimental protocols. This work reports an online evaluation of this approach, showing that it can effectively reduce the calibration time with respect to the standard practice without degrading the online recognition performance.

## 2 Methods

### 2.1 Experimental protocols

Twelve participants performed three experimental protocols of increasing complexity as shown in Fig. 1. They were seated on a comfortable chair facing the visual displays of the protocols approximately one meter away and asked to restrict eye movements and blinks to specific

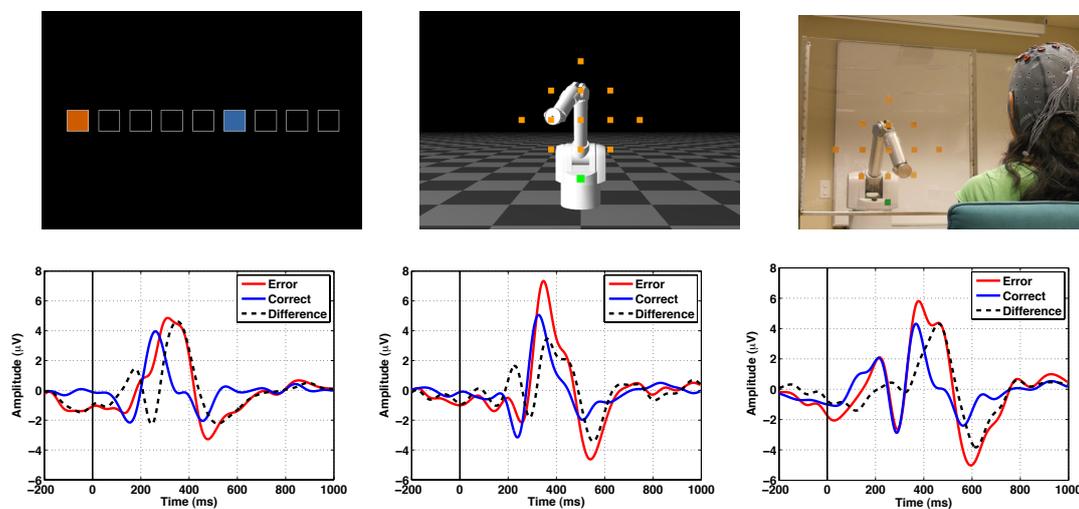


Figure 1: (Top) Experimental protocols. (Bottom) ErrPs at the FCz electrode obtained in each experimental protocol. (Left) One-dimensional cursor movement ( $E_1$ ). (Center) Two-dimensional movements of a simulated robotic arm ( $E_2$ ). (Right) Real robotic arm ( $E_3$ ).

resting periods. In all experiments they were asked to evaluate whether a device moves towards a given target location. The device moved in discrete steps and the time between movements was randomly chosen within the range [1.7 4.0] s. There was a probability of moving in the wrong direction of about 30%. Experiments were always performed in the same order from the simplest to the most complex one. The first experiment,  $E_1$ , consists of a cursor that moves in discrete steps (either left or right) towards a target [1]. In the second protocol,  $E_2$ , the user monitors a simulated robotic arm that moves on a 2D plane (allowed movement directions were left, right, up and down). The third experiment,  $E_3$ , consists of the same task using a real robotic arm. A detailed explanation of the protocols and methods is provided in [2].

Each experiment started by a calibration phase. This phase had a variable length depending on the obtained performance. Calibration stopped whenever the mean accuracy (ten-fold cross-validation) on the training data exceed 75%. Then the classifier parameters were fixed, and performance was tested on an online phase lasting 400 trials.

EEG was recorded at 256 Hz with a gUSBamp amplifier (gTec GmbH, Austria) with 16 active electrodes (Fz, FC3, FC1, FCz, FC2, FC4, C3, C1, Cz, C2, C4, CP3, CP1, CPz, CP2, and CP4 according to the 10-10 system). Ground and reference were placed on the forehead and the left earlobe. Data was notch filtered at 50 Hz, and zero-phase band-pass filtered at [1, 10] Hz. Prior to classification, we applied common-average reference and downsampled the signal to 64 Hz. Features from eight fronto-central channels were selected in the window [200 800] ms using a spatiotemporal filter [3]. On average  $45 \pm 10$  features were selected based on their  $r^2$  score. Single-trials were classified as erroneous or correct using linear discriminant analysis (LDA).

## 2.2 Training paradigm using latency-correction

As mentioned above, differences between ErrPs elicited in different experiments can be largely explained by latency variations. These variations can be easily estimated by computing the

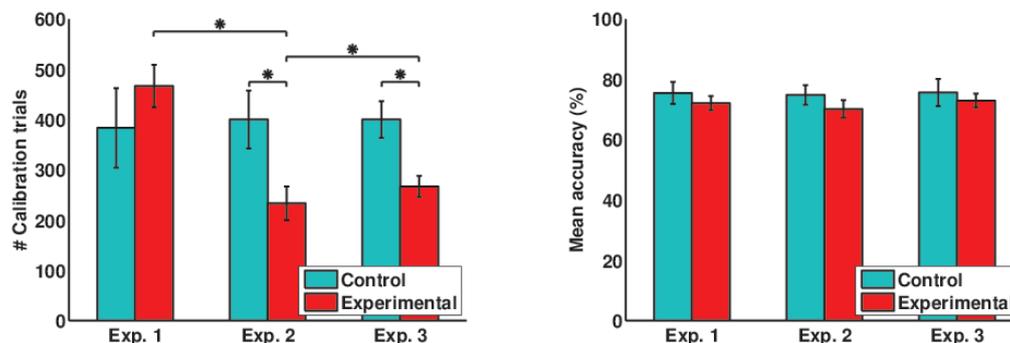


Figure 2: (Left) Number of calibration trials (mean  $\pm$  SEM) required to reach the calibration criterion (\*:  $p < 0.05$ ). (Right) Mean classifier accuracy during the online control phase.

cross-correlation between the grand-average ERPs for each experiment. In brief, given a previous experiment  $E_i$ , from which data is available, and a new experiment  $E_j$ , the ERP latency variation  $d_{E_i E_j}$  will correspond to the shift that yields the maximum cross-correlation. Then, ERP data from  $E_i$  can be shifted in time by  $d_{E_i E_j}$  and used along the available (few) trials from the new experiment  $E_j$  to train a classifier.

Complementing previous reports [2], we compare whether this latency correction mechanism effectively reduces the calibration time. To that end, we defined two groups of participants depending on the training procedure. The *control* group (N=6, one female, mean age  $27.33 \pm 2.73$  years) followed a standard calibration approach, i.e., based only on data from the current experiment. The *experimental* group (N=6; two females, mean age  $27.17 \pm 4.07$  years) used latency-corrected trials from the previous experiment to build the classifier for the current task. That is, standard calibration was followed for  $E_1$ , while data from that experiment was used during the calibration period of  $E_2$ . Similarly, during calibration for  $E_3$  the data from  $E_2$  was used. The latency between experiments was estimated based on the cross-correlation of the difference ERP (error minus correct condition) of channel FCz within the window  $[0, 500]$  ms.

Mixed two-way ANOVAs (within factor: experiments; between factor: group of subjects) were performed to test whether (i) the number of calibration trials in the experimental group decreased across experiments; (ii) the number of calibration trials was significantly different between groups; and (iii) the online accuracies of both groups were not different. Post-hoc one-tailed Bonferroni-corrected t-tests were performed to assess statistically significant differences.

### 3 Results

The obtained ERPs can be seen in Figure 1. Latency variations were about  $60 \pm 25$ ms between  $E_1$  and  $E_2$ , and about  $41 \pm 13$  ms between  $E_2$  and  $E_3$ . Figure 2 shows the number of calibration trials needed in each experiment to reach the stopping criteria. The calibration period for the control group was similar for all experiments. In contrast, the experimental group exhibit a large reduction on the required calibration trials in  $E_2$  and  $E_3$  when previous information was re-used. The ANOVA test revealed a significant interaction between the experiment and group ( $F_{2,20} = 8.65, p = 0.002$ ). Post-hoc tests showed that significant differences were found between groups in experiments 2 and 3 (one-tailed unpaired t-tests,  $p < 0.05$ ), and also significant differences within the experimental group between experiments 1 and 2 (one-tailed paired t-

test,  $p = 0.004$ ), and between experiments 1 and 3 ( $p = 0.004$ ).

No significant difference was found in the accuracies for all the experiments and subjects ( $p > 0.85$ ). These results indicate that, provided data from previous experiments, knowledge from these protocols can be transferred to the new task using the latency correction algorithm.

## 4 Conclusion

Our results confirm that compensating for latency variations across protocols allows the use of previous data to shorten the calibration phase in new applications. In the reported experiments the use of data from the first experiment enabled users to reach the training criteria for the second one in about half of the trials required with the standard approach. Importantly, no significant difference was observed between the two training paradigms in the online performance. Similar reductions were also observed for the third experiment with the real robot.

The latency correction mechanism used in this work relied on a simple measure based on the cross-correlation in a single channel. However, ERP variations across experiments may follow more complex patterns both spatially (i.e. across channels) and temporally (i.e. individual ERP components). Multiple factors affect the ERP waveforms including the feedback modality, the inter-stimulus interval, as well as subject dependent variability. It remains to be validated how suitable this correction mechanism is to other experimental paradigms and signals. Moreover, further research is required to explore more sophisticated techniques to better model these ERP variations (e.g. dynamic time warping).

## 5 Acknowledgments

This work was supported by the EU Project FP7-224631; the Swiss funded NCCR Robotics and the Spanish Ministry of Science projects HYPER- CSD2009-00067 and DPI2009-14732-C02-01, DPI2011-25892, CAI Programa Europa, DGA-FSE, and grupo T04.

## References

- [1] R. Chavarriaga and J. del R. Millán. Learning from EEG error-related potentials in noninvasive brain-computer interfaces. *IEEE Trans Neural Syst Rehabil Eng*, 18(4):381–388, Aug 2010.
- [2] I. Iturrate, R. Chavarriaga, L. Montesano, J. Minguez, and J.d.R. Millán. Latency correction of error potentials between different experiments reduces calibration time for single-trial classification. *J Neural Engineering*, 11:036005, 2014.
- [3] I. Iturrate, L. Montesano, R. Chavarriaga, J. del R. Millán, and J. Minguez. Spatio-temporal filtering for EEG error related potentials. In *5th Int Brain-Computer Interface Conference, 2011*, pages 12–15, Graz, Austria, 2011.
- [4] S.K. Kim and E.A. Kirchner. Classifier transferability in the detection of error related potentials from observation to interaction. In *IEEE Int Conf Systems, Man, and Cybernetics (SMC), 2013*, pages 3360–3365, Oct 2013.
- [5] F. Lotte and C. Guan. Learning from other subjects helps reducing brain-computer interface calibration time. In *Int. Conf. Audio Speech and Signal Proc (ICASSP)*, pages 614–617, 2010.
- [6] W. Samek, F.C. Meinecke, and K.-R. Müller. Transferring subspaces between subjects in brain-computer interfacing. *IEEE Trans Biomed Eng*, 60(8):2289–2298, 2013.
- [7] C. Vidaurre, M. Kawanabe, P. von Büna, B. Blankertz, and K.-R. Müller. Toward unsupervised adaptation of LDA for brain-computer interfaces. *IEEE Trans Biomed Eng*, 58:587–597, 2011.

# Asynchronous detection of error potentials

Jason Omedes<sup>1</sup>, Iñaki Iturrate<sup>2</sup> and Luis Montesano<sup>1</sup>

<sup>1</sup> I3A, DIIS, University of Zaragoza, Zaragoza, Spain

<sup>2</sup> Chair in Non-Invasive Brain-Machine Interface (CNBI), EPFL, Switzerland  
jomedes@unizar.es, inaki.iturrate@epfl.ch, montesano@unizar.es

## Abstract

Recent developments in brain-machine interfaces (BMIs) have proposed the use of error-related potentials as cognitive signal that can provide feedback to control devices or to teach them how to solve a task. Due to the nature of these signals, all the proposed error-based BMIs use discrete tasks to classify a signal as correct or incorrect under the assumption that the response is time-locked to a known event. However, during the continuous operation of a robotic device, the occurrence of an error is not known a priori and thus it is required to be constantly classifying. Here, we present an experimental protocol that allows to train a decoder and detect errors in single trial using a sliding window.

## 1 Introduction

EEG-measured error-related potentials (ErrPs) are one class of event-related potential (ERP) elicited in the user brain when the outcome of an event differs from the user expected one. These potentials have been observed, in particular, when a user observes a machine committing incorrect actions or operations. Recent works have proposed the incorporation of these signals into brain-machine interfaces (BMI) to correct the classifier or as rewards to control or to teach devices [5]. As with all event-related potentials, the neural response associated to ErrPs is triggered in response to an exogenous event. Consequently, most of the developed works try to distinguish whether an action is correct or erroneous based on the knowledge of the time of its occurrence [1, 4]. On the other hand, real applications (such as executing a trajectory with a robotic arm or a mobile robot) imply the use of continuous actions where the classifier has to asynchronously differentiate between erroneous events and the background EEG.

This work presents a method to detect error potentials during the continuous operation of a device. Showing that using events introduced as abrupt changes of direction, it is possible to train a classifier to later on asynchronously classify among a complete trajectory achieving detection rates comparable to those obtained in discrete tasks.

## 2 Methods

### 2.1 Experimental Protocol

Two healthy subjects (mean age 28 years) participated in the study. The experimental setup consisted of a virtual cursor that had to reach a target position by moving at a fixed speed towards it. The initial cursor and target positions of each trial were randomly generated. One trial consisted of a trajectory performed by the device and lasted between 3 and 5 seconds. Trajectories were correct in 70% of the trials, which consisted in a straight line between the start and goal locations, Figure 1a. The remaining 30% of the trials were erroneous trajectories that started as the correct ones but executed an abrupt change of direction in a random instant between the 20% and 80% of the path, Figure 1b. Six rounds composed of 40 trials each were recorded, obtaining around 70 erroneous and 170 correct trials per participant.

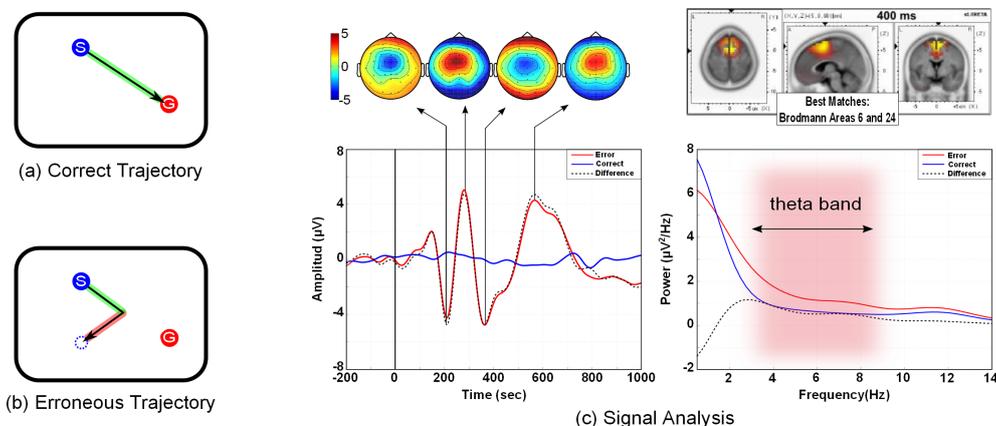


Figure 1: (a-b) Designed experimental setup. Starting (S) and goal (G) positions of the device are marked in blue and red respectively. Correct and wrong directions of movements are shadowed in green and red respectively. (c) Neurophysiology Analysis: grand averages in temporal (bottom, left) and frequency (bottom, right) domain, topographical representation (top, left) and source localization (top, right).

EEG and EOG activity were recorded at 256 Hz using a gTec system with 32 electrodes according to the 10/20 international system (including 8 fronto-central channels), with the ground on FPz and the reference on the left earlobe; for the EOG, 6 monopolar electrodes were recorded [3], with the ground on FPz and the reference on the left mastoid. Recorded data were power-line notch filtered, and band-pass filtered at [1, 10] Hz. The EEG was also common-average-reference (CAR) filtered. Additionally, EOG was removed from the EEG using a regression algorithm, and those trajectories where EOG magnitude higher than  $40 \mu\text{V}$  was detected, were rejected.

## 2.2 Single-trial continuous classification

Temporal domain features have been widely used to detect ErrPs under discrete tasks. However, for continuous detection, temporal features result in a large number of false positives as EEG oscillations resemble ErrP patterns. On the other hand, it has been shown that frequency features can be more robust under specific changes in the ErrP signals [7]. In this work, we propose the combination of temporal and frequency features extracted from the most relevant common spatial patterns associated (CSPs) for the continuous detection of ErrPs. To train a classifier, the onset of the erroneous events was selected at the instant in which the device performed the abrupt change of direction. On the other hand, onsets of correct trials were selected at random instants of time within the execution of a correct trajectory. Training data from error and correct conditions was used to extract the CSPs, and the two first CSPs were retained. For each retained CSP, the temporal features were the EEG voltages within a time window of [0, 1000] ms downsampled to 64 Hz, forming a vector of 128 features. The power spectral density (PSD) was computed taking a window interval from 0 to 1000 ms as it gave the best trade-off between frequency resolution and capturing the signal of interest. Frequency features were selected as the power values of each channel from the theta band ([4, 8] Hz)  $\pm 1$  Hz [2] leading to a vector of 14 features. Finally, both set of features were concatenated and normalized within the range [0, 1].

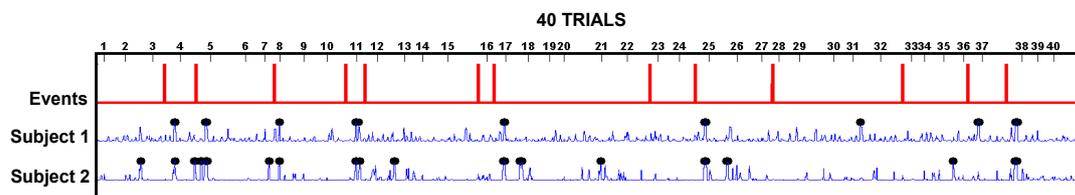


Figure 2: Representative example of the sliding window results. The error events are plotted as red spikes indicating a change in direction, while the probability of detecting an error ( $p_e$ ) for each of the 2 subjects is plotted in blue. Black dots over the probability values indicate the time instant when the classifier detected an error ( $p_e > 0.8$ ).

After balancing the error and correct datasets, a RBF-SVM was trained using the previous features [6]. During classification, we retained the classifier output for each window evaluation as the probability estimate that the current EEG was an error,  $p_e$ . For the asynchronous classification we performed a 6-fold cross-validation where each fold was composed by a complete recorded round. We continuously classified every 62.50ms over the testing set using a sliding window of one-second width. To ensure a low false positive rate, the detection of error events was only considered when  $p_e > 0.8$ , value selected through cross-validation. For the sliding window, all inter-trial data was removed. Classification performance was computed as follows: the erroneous trials where the classifier detected an error, in a one second window after the direction change, were considered as true positives. When an error was not detected the trial was a false negative. True negatives were those correct trials where no error was detected. And, when an error was detected on correct trials, they were considered false positives.

### 3 Results

Figure 1(c) shows the error, correct and difference grand averages for channel FCz averaged for the two subjects in temporal and frequency domain, next to topographic interpolation of 4 representative peaks and the source localization for the most prominent negative deflection. Regarding to the results of applying the sliding window, Figure 2 displays the detection level obtained for both subjects in a representative round of the experimental condition. Here, the 40 trials that compose the round are concatenated removing the inter-trial resting periods. It can be seen that most of the trials are properly classified. Also notice that the correctly detected ErrPs are delayed with respect the onset of the event. This delay was on average  $867.13 \pm 99.33$  ms after the onset of the erroneous action, corresponding to the time needed for the appearance of the most relevant peaks and the maximum spectral power activation used as features.

The performance rates achieved for the entire test set reached an 89.09% of true negative and 64% of true positives. At the same time, the number of false positives and false negatives hovered the 11% and 36% respectively. On average, these values were around 10% less performance

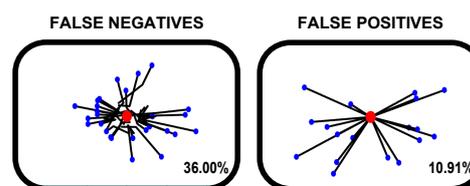


Figure 3: Trajectories of false positives and false negatives reprojected to have the goal at center of the image (red dot). The starting position is marked in blue.

than those obtained with standard ErrP protocols that only classify in the window of the event [1]. This decay is reasonable considering that classifying with a sliding window conveys a higher chance of detecting a false positive during non event intervals. Trajectories executed by the device according to their classification as false positives or false negatives are depicted in Figure 3. Furthermore, it can be seen that correct trials are mostly well detected independently of the direction and distance covered by the device, and only few of them are detected as erroneous. However, the number of erroneous trials not detected was higher, since it was preferable to miss the detection of an error than detect errors where was not intended. Also, notice that most of the true negatives ended up very close to the goal. Indeed, 40% of them ended less than 50 pixels from it, which lead us to think that the subjects may have not interpreted them as erroneous.

## 4 Conclusions and future work

This paper studies the on-line asynchronous detection of error potentials during continuous trajectories. The results obtained for the proposed experimental protocol show that the error potentials appear when the user monitors a continuous target reaching task and that they can be detected in single trial using a sliding window, obtaining results comparable to those achieved in discrete tasks. These promising results are a first step towards the use of this type of cognitive information to control or teach robotic devices in realistic and complex tasks. There exist several opportunities for future work. Currently, we are extending the study to more users and more types of trajectories containing errors. Also we are testing the usage of this kind of events on real devices obtaining promising results.

## 5 Acknowledgments

This work has been supported by Spanish projects DPI2011-2589, and DGA-FSE (grupo T04).

## References

- [1] R. Chavarriaga and J.d.R. Millán. Learning from EEG error-related potentials in noninvasive brain-computer interfaces. *IEEE Trans Neural Syst Rehabil Eng*, 18(4):381–388, 2010.
- [2] Michael X. Cohen. Error-related medial frontal theta activity predicts cingulate-related structural connectivity. *NeuroImage*, 55(3):1373–1383, 2011.
- [3] RJ Croft and RJ Barry. EOG correction of blinks with saccade coefficients: a test and revision of the aligned-artefact average solution. *Clinical neurophysiology*, 111(3):444–51, March 2000.
- [4] I. Iturrate, R. Chavarriaga, L. Montesano, J. Minguez, and J. del R Millan. Latency correction of error potentials between different experiments reduces calibration time for single-trial classification. In *Int Conf of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2012.
- [5] I. Iturrate, L. Montesano, and J. Minguez. Shared-control brain-computer interface for a two dimensional reaching task using eeg error-related potentials. In *Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2013.
- [6] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi. A review of classification algorithms for EEG-based brain-computer interfaces. *J Neural Eng*, 4(2):R1–R13, June 2007.
- [7] J. Omedes, I. Iturrate, L. Montesano, and J. Minguez. Using frequency-domain features for the generalization of eeg error-related potentials among different tasks. In *Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2013.

# In dialogue with the real experts: technical, ethical, legal and social requirements for BCIs as access technologies

Femke Nijboer<sup>1</sup>

<sup>1</sup> University of Twente, Enschede, the Netherlands.  
Femke.nijboer@utwente.nl

## Abstract

This preliminary study describes the assessment of BCI technology by people with neuromuscular disorders (N=7) and high spinal cord injuries (N=9) from a technical, ethical, legal and societal perspective. It becomes clear that people with disabilities could and should greatly contribute to the design of assistive technology.

## 1 Introduction

The field of brain-computer interfacing (BCI) is rapidly expanding its application areas. Scientific and technological endeavors typically focus on feasibility, validity and reliability of such emerging technologies. However, the transfer of BCI technology to the clinic may be slow or difficult if practical issues as well as ethical, legal and social issues are not properly and timely addressed.

Previously, Nijboer and colleagues interviewed rehabilitation specialists (N=28) what type of users could be potential target users for BCIs as access technologies and what design requirements BCIs should fulfil to be usable and pleasant technologies for such users [1]. Recommended target users are only those who can hardly or not at all use alternative access technologies. People in the locked-in state (resulting from late-stage amyotrophic lateral sclerosis, multiple sclerosis, spinal muscular atrophy type II or classical or total locked-in syndrome) and people with high spinal cord injury (C1/C2) could be target users. Specialists caution engineers and developers that these users may have many concurring problems such as sensory or cognitive impairments and epileptic seizures. In addition, transferring BCIs from the lab to the daily life of such target users will need a grounded consideration for ethical, legal, social and cultural issues.

In this study we investigate the experience and opinions of people with neuromuscular and muscular diseases (N=7) and people with high spinal cord injuries (N=9) on such issues after they 1) have been educated about BCIs, 2) have had the opportunity to experience a BCI and 3) had the opportunity to discuss the technology with each other. These user groups are interesting to compare. One group, with disorders which have progressed sometimes from childhood, has a lifelong experience with assistive technologies. The people with SCI have had years of experience with access technologies for able-bodied people (keyboard and mouse) and now find themselves having to use access technologies for disabled people. In addition, people with SCI do not have to anticipate further functional decline, whereas people with progressive disorders do.

## 2 Methods

Seven participants with progressive neuromuscular or muscular diseases (NM group) and nine participants with high spinal cord injuries (SCI group) were recruited through the Dutch Neuromuscular Diseases Association and the Association Spinal Cord Injury (see Table 1). Participants were invited to attend workshops (one workshop for each group of users) entitled ‘The possibilities and impossibilities of Brain-Computer Interfaces’. The two workshops were prepared together with the directors of the associations and one of the participants to ensure that the program was relevant and satisfactory for participants. Workshops were free of charge and held in accessible buildings in a central location of the Netherlands to ensure equal access. Participants were reimbursed for travel costs and provided informed consent before the workshop for their opinions to be recorded on audio and video and pictures to be made.

Participant	Diagnosis	Age	Verbal Communication	Wheel chair	Artificial ventilation	Access technology in use
<b>NM group</b>						
1	SMA type II		Slurred	Yes	Yes	Sip & puff device
2	ALS	52	Speech synthesizer	No	No	Keyboard/mouse
3	SMA type II	38	Normal	Yes	Yes	Finger switch
4	SMA type II	33	Slurred	Yes	Yes	Eye tracking
5	Distal SMA	60	Normal	Yes	No	Keyboard/mouse
6	SMA type II	23	Normal	Yes	No	Keyboard/mouse
7	SMA type II	21	Normal	Yes	Nighttime	Keyboard/mouse
<b>SCI group</b>						
1	T5	61	Normal	Yes	No	Keyboard/mouse
2	C4/C5	45	Normal	Yes	No	Keyboard/mouse, mouth switch
3	C3/C4	39	Normal	Yes	No	Head switch
4	C3/C4	51	Normal	Yes	Nighttime	Chin joystick
5	C4/C5	61	Normal	Yes	No	Keyboard/mouse
6	C3/C5	38	Hushed	Yes	No	Keyboard/mouse
7	C5/C6	30	Normal	Yes	No	Adapted keyboard
8	C5/C6	52	Normal	Yes	No	Keyboard/mouse
9	C6/C7	60	Normal	Yes	No	Keyboard/mouse

**Table 1:** Overview of participants. NM = neuromuscular; SCI = Spinal Cord Injury; SMA = Spinal Muscular Atrophy; ALS = Amyotrophic Lateral Sclerosis. Hushed = cannot produce normal voice loudness.

The workshop followed a 3-step format which we dubbed the “give-and-take” approach. First, participants were educated about the technical components of BCIs, available neuroimaging techniques and types of applications. They were also presented with some of the major challenges the field of BCI faces (e.g. universal design, reliability issues, limitations to information transfer rate, sensors). Second, two participants in each workshop were offered to try out a BCI as an access technology to operate a commercially available computer access software (The Grid 2 from Sensory Software, see for a description [1] and for a video of a participant trying the demo: <http://youtu.be/gf3C1AHT8U>). The rest of the group watched and asked questions. Third and finally,

we held a focus group interview with the participants to assess the technology they had seen in the demonstration and the BCI field at large. Participants were prompted with predefined questions, but also encouraged to bring up other topics they found (more) important. The interviews were transcribed verbatim. Here we offer a preliminary descriptive overview of take home messages given by participants.

### 3 Results

The focus group interviews provided recommendations from the participants on 4 different levels: technical, ethical, social and legal (policy) issues. In addition, a philosophical outlook on human identity was discussed.

In general, participants were positively surprised by the state of the art of the BCI field, the effectiveness and the feeling of the BCI prototype. On a technical level, participants in both groups agreed that they could not yet see the added functional value of current BCIs over existing communication aids, although some participants in the NM group could imagine that BCIs could have added value for them in about a year, because they anticipated a further decline in function such that operation of their current aid would no longer be possible. Both groups indicated that the usability of the system needs improvement. In both groups there was a strong interest for operation of a robot arm rather than a communication aid, which seemed to be motivated, firstly, by the availability of reliable communication aids in the NM group and the ability to speak in the SCI group and, secondly, by an interest to expand or enhance current functionality rather than just repair functionality.

This was often accompanied by a wish to have electrodes implanted in the brain for practical reasons: 1) sensors would be always in place (no need to bother caregivers), 2) the BCI would be less bulky and less prone to get damaged by outside factors, and 3) invisible sensors are more esthetically appealing.

Many participants seemed reluctant to discuss ethical issues and rather discussed practical issues, but when prompted with questions such as “what would you be afraid of?” participants recognized that ethical problems can be practical as well and indicated that agency was important. They want to feel sure that the BCI does not “*go out of control*” or “*takes over control*”. Also, the surgery for implanted sensors was perceived as risky, in particular for people in the NM group, since anesthesia is often very complicated if not impossible for these people. Nevertheless, a few participants said they “*would risk it*” if it gave them better functionality. Participants with neuromuscular disorders advise to do the surgery in an early stage of the disease and, if possible, in combination with the surgery for other life-sustaining measures, such as the tracheotomy.

Participants did not identify legal issues such as liability problems or see the need for special laws, but they did discuss reimbursement policies. For example, mainstream technologies, such as an iPad, are not reimbursed even if these technologies would sometimes be the best solution for some people with disabilities. Instead, they have to make a choice out of special technologies for disability. Participants foresee difficulty in obtaining financial reimbursement for BCI technology. One participant said: “*It is also a money thing of course. As long as I can use a puff-and-sip device, we’ll use that. It is cheaper than a BCI*”.

Societal issues related to BCI technology and assistive technology were discussed at length. Many participants indicated that they did not want to wear a cap with sensors on their head unless it was disguised as conventional headwear. The technology that surrounds people with disability, such as communication aids and wheelchairs often scare people in society to such an extent that they literally turn their back when a person with a disability comes in sight. Hence, any BCI technology that would attract even more attention to persons with disabilities could potentially exclude them further rather

than include them in society as intended. Participants stress the fact that esthetical BCI design is therefore of utmost importance. On a more philosophical level we shortly discussed if BCIs could change the boundaries between humans and technology, but participants mostly agree that such boundaries have long changed for them since they are so dependent on technology anyway. They are not afraid of the cyborg idea. Rather they would rather embrace it.

## 4 Discussion

Compatible with previous studies [2-4], the preliminary study presented here showed that participants have stringent technical requirements concerning the usability of the overall BCI system (efficiency, robustness, esthetics of the cap, multiple users). After reviewing the current state of the art of BCI technology, participants did not yet see an added value of BCIs for themselves. However, they strongly endorse further development and would like to be involved in the process.

User involvement in technology design should not necessarily be restricted to the definition of technological requirements. This study shows that participants also have ethical requirements about agency and (timing of) risks. Agency is essential for feeling safe, for self-determination and for dignified living, especially for users with severe disabilities whose lives literally depend on technology. This study also shows that users with disabilities are interested in implanted electrodes. However, people with progressive neuromuscular disorders recommend that surgery should happen at an early stage in disease onset when the risk of anesthesia is lower. This view contrast the view of most scientists and ethical committees that surgery for a BCI should be a 'last option'. A recommendation could be to investigate scenarios in which surgeries happen at an early stage. Compatible with [5], most participants felt no need for special BCI regulations. Legal and policy recommendations focus on reimbursement issues (cost-effectiveness and robustness related to maintenance costs). Finally, end-users know – as no other - how society reacts to disability and assistive technology. When the appearance of BCIs does not improve, BCI technology risks to exclude people from society rather than to include them.

In conclusion, it becomes clear that people with disabilities could greatly contribute to the design of assistive technology and provide expertise into the ethical, policy and societal issues that must be addressed for successful technology transfer to the market. Thus, it is strongly recommended that potential users are involved – as experts – in the earliest stages of research and development of BCIs.

## References

1. Nijboer, F., et al., *Design requirements and potential target users for brain-computer interfaces – recommendations from rehabilitation professionals*. Brain-Computer Interfaces, 2014: p. 1-12.
2. Zickler, C., et al., *Brain Painting: Usability testing according to the user-centered design in end users with severe motor paralysis*. Artificial intelligence in medicine, 2013. **59**(2): p. 99-110.
3. Zickler, C., et al., *A Brain-Computer Interface as Input Channel for a Standard Assistive Technology Software*. Clinical Eeg and Neuroscience, 2011. **42**(4): p. 236-244.
4. Blain-Moraes, S., et al., *Barriers to and mediators of brain-computer interface user acceptance: focus group findings*. Ergonomics, 2012. **55**(5): p. 516-25.
5. Grüber, G., et al., *Psychosocial and Ethical Aspects in Non-Invasive EEG-Based BCI Research—A Survey Among BCI Users and BCI Professionals*. Neuroethics, 2013: p. 1-13.

# Platform for analyzing multi-tasking capabilities during BCI operation

Robert Leeb<sup>1</sup>, Kiuk Gwak<sup>2</sup>, Dae-Shik Kim<sup>2</sup> and José del. R. Millán<sup>1,3</sup>

<sup>1</sup> Center for Neuroprosthetics, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland

`robert.leeb@epfl.ch`

<sup>2</sup> Brain Reverse Engineering and Imaging Laboratory, Korea Advanced Institute of Science and Technology, Daejeon 305-701, Korea

`kgwak@kaist.ac.kr`, `dskim@ee.kaist.ac.kr`

<sup>3</sup> Chair on Non-Invasive Brain-Machine Interface, Institute of Bioengineering, School of Engineering, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland

`jose.millan@epfl.ch`

## Abstract

Operating brain-actuated devices requires split attention between the interaction of the device with its environment and the Brain-Computer Interface (BCI) feedback. In case of screen-based applications it is possible to merge BCI feedback and application, but not in case of controlling devices, like wheelchairs or exoskeletons. Recently we demonstrated that BCI feedback could be provided via tactile stimulation and no performance difference between tactile and visual BCI feedback could be found. In this work we want to present a platform for characterizing these multi-tasking capabilities during BCI operation. Thereby, the subjects will have to perform a visual engaging and complex task in a game like setup while interacting with the BCI in a coordinated multi-tasking manner.

## 1 Introduction

Controlling a brain-actuated device like a wheelchair requires the participant to look at and to split his attention between the interaction of the device with its environment and the status information of the Brain-Computer Interface (BCI). Such parallel visual tasks are partly contradictory, with the goal of achieving a good and natural device control. Recently we demonstrated that it is possible to free the visual channel from one of these tasks, by providing feedback for a motor imagery based BCI via a spatially continuous tactile stimulation [1]. Thereby, we could demonstrate that the participants were able to perceive this type of tactile feedback well and no statistical degradation in the online BCI performance could be identified between visual and tactile feedback conditions. Nevertheless, an open issue is to demonstrate, that the freed visual channel can now be fully devoted to the device while controlling the BCI.

In this work we want to present a platform for analyzing these multi-tasking capabilities during BCI operation. Thereby, the subjects will have to perform a visual engaging and complex task in a game like setup while interacting with the BCI in a coordinated multi-tasking manner with the presented stimuli. This platform is giving us the possibility to evaluate different parameters in a reproducible design, which would not be possible with a brain-actuated device.

## 2 Methods

The goal of this work is the evaluation of multi-tasking capabilities while exploiting visual and tactile BCI feedback. Therefore, the participant will have to perform 2 tasks in parallel.

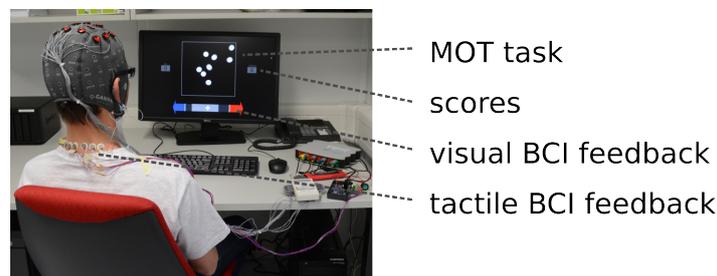


Figure 1: Experimental setup including the multi-object tracking game and the BCI feedback. Visual and tactile BCI feedbacks are used alternately.

The first task is the use of a 2-class BCI, which allows the delivery of a right or left BCI command via the imagination of hand movements (MI). E.g. the imagination of a left hand movement moves the BCI feedback to the left and the imagination of a right hand movement to the right. The instructions when to move the BCI feedback and in which direction (cue) is embedded in the parallel task described below.

The second task is to track several ball like objects on the screen (multi-object tracking game [3], see Figure 2). At the beginning of each trial one, two or three targets out of 10 possible balls are highlighted and the participant is instructed to keep track of them. Furthermore, only the target(s) is/are slightly moving twice for 1 second either to the left or right, to indicate the target direction (the BCI cue for the motor imagery) for this run.

After that, all targets are starting to move around in a physical realistic way. They can bounce at the wall or at each other, but are never occluded. After approximately 10 seconds half of the balls are highlighted (changing the color) for 4 seconds. The participants should perform a BCI command during the time (while still keeping track of the targets). If one of the tracked targets is highlighted, the participant should perform the BCI command, which was indicated by the cue during the first 2 seconds of the trial. If the tracked target is not highlighted the opposite BCI command should be applied. After these 4 seconds all balls return to white. After another random period of around 10 seconds again half of the balls are highlighted and the same actions as mentioned above have to be performed, dependent if one of the targets is highlighted or not.

The trial finishes after around 30 seconds and all balls stop moving. They are immediately numbered and the participant has to mark the one(s), s/he thinks are the targets. Furthermore, they can report at which time the lost (if at all) one of the targets, and/or delivered a wrong BCI command. After that the users receives feedback about his/her tracking performance. Finally a new trial with new targets and BCI cue direction is starting.

The visualization contains two parts (see Figure 1). The balls are displayed in the top center part of the screen. But also information about the current state of the 2-class BCI is presented to the participant, which gives information of how much the BCI thinks that the participant is applying e.g. the imagination of a left or right hand movement. Two different feedback conditions are performed: This feedback is either displayed as a visual bar on the bottom part of the screen (as used in normal BCI runs). Or in the second conditions this BCI feedback is presented via 6 tactile stimulators on the neck (see Figure 1), whereby different stimulators are activated to present the information similar to the position of the bar before. In this condition no visual BCI feedback is shown, but the tracking task is performed in the same visual manner.

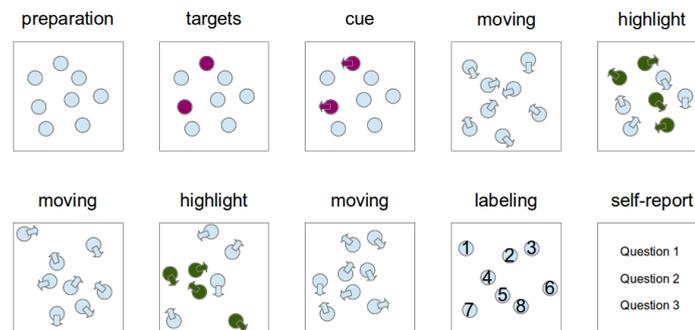


Figure 2: Time-line of the experimental protocol.

The experiment contains 10 trials in a series (called run), after that a short break is given. In total 2–4 runs are played with each of the 2 conditions and with different levels of difficulties.

### 2.1 Applied BCI setup

The brain activity was acquired via 16 EEG channels placed over the sensori-motor cortex (Fz, FC3, FC1, FCz, FC2, FC4, C3, C1, Cz, C2, C4, CP3, CP1, CPz, CP2 and CP4 with reference on the right ear and ground on AFz). The EEG was recorded with a g.USBamp (g.tec medical engineering, Austria) at  $f_s=512$  Hz, band-pass filtered 0.5–100 Hz and notch filter set to 50 Hz. From the Laplacian filtered EEG, the power spectral density (4–48 Hz) was calculated. Canonical variate analysis was used to select subject-specific features, which were classified with a Gaussian classifier. Decisions with low confidence on the probability distribution were filtered out and evidence was accumulated over time. More information about our BCI is given in [2].

In online experiments the output of the BCI is translated in a movement of the feedback, which informs the subjects about their current brain status. In the case of visual feedback, the horizontal bar moves on a screen. In the case of tactile feedback, the motors vibrate accordingly.

## 3 Results

Three experienced BCI participants (between 27 and 40 years, all male, all participated already in study [1]), performed the first tests with the goal of tracking 1 or 2 balls out of 10 with either visual or tactile feedback (conditions are called v1, v2, t1 or t2). As performance measures we define the tracking score (TS) as the percentage of correctly tracked balls, the BCI true positive rate (TP) as the percentage of correct BCI commands delivered in the highlighting period (maximum 1 command possible per period) and the BCI false positive rate (FP) as the ratio between BCI commands delivered in the free moving period (intentional non-control, where no commands should have been delivered) compared to the number of highlighting periods.

Exemplary for one subject the detailed results are: TS varied between 100 % in t1, 75 % in v1, 55 % in t2 and 40 % in v2, with a TP of 73 % in t1, 75 % in v1, 65 % in t2 and 50 % in v2, and with a FP of 40 % in t1, 55 % in v1, 15 % in t2 and 70 % in v2. The other subjects showed similar behaviors, but with different performance levels (e.g. TP of 30 %). Overall in all subject, the TS was reduced in the visual conditions compared to the tactile ones, and further reduced if the number of targets was increased. The TP ratio was stable over the two feedback types, but seems to be reduced if more targets have to be tracked.

Participants reported that especially an existing bias in the BCI classifier, was dragging too much attention away from the search task, since the subject had to fight all the time against the misclassification. Otherwise if the tracking task was performed with higher priority the number of FP or wrong commands increased. Furthermore, the BCI decision thresholds and the duration of the highlighting period have to be mutual adjusted, so that there is more and longer chance to deliver correct commands. Currently some correct commands have been delivered shortly after the end of the highlighting period (which would be FP) since it took longer to identify the correct command side and then bring the BCI feedback towards it.

Nevertheless, we realized that the original goal of tracking 3 balls out of 10, while delivering BCI commands in (randomized) defined orders is very challenging and mostly too complex. Subjects even reported that they forgot the target class (cue) within one run, since the workload was too high, which of course is contra productive to the goal of the game.

Therefore, we are currently modifying our setup according to following parameters:

1. Keeping the BCI target class (which has to be delivered if a tracked ball is highlighted) predefined for the whole run or even session and not changing it every trial. This will not have any effect on the balance of BCI commands to be delivered, since anyway just half of the time the tracked balls are highlighted.
2. Characterization of the maximum possible balls to be tracked and the maximum speed of the balls, without BCI control but while delivering faked BCI commands. The idea thereby is, that the subject is experiencing the same BCI feedback characteristics (as it would be during brain control), but the subjects are not delivering mental commands. Instead they are sending the commands via the keyboard (to have a 100% success rate), which are then influencing the BCI feedback towards the left or right side.

## 4 Discussion and Future Work

The presented platform is giving us the possibility to evaluate different multi-tasking parameters. The strong link between the challenging visual multi-object tracking task and BCI feedback in coordination with the presented stimuli is a perfect simulation environment of a real brain-actuated device.

Unfortunately, the initial experiments were too challenging and complex, so that modifications and a deeper characterization of the game parameters are necessary. Furthermore, an unbiased classifier should be applied and an additional standard BCI online recording before and after the experiment should be performed. Based on the outcome of the characterization, we will modify the tracking time and paradigm, the number of balls and the used speed, so that there is a chance to succeed in the task during easier levels, while it will be very challenging at higher levels, and then re-do the experiments.

## References

- [1] R. Leeb, K. Gwak, D-S. Kim, and J.d.R. Millán. Freeing the visual channel by exploiting vibrotactile BCI feedback. In *Conf Proc IEEE Eng Med Biol Soc*, 2013.
- [2] R. Leeb, S. Perdikis, L. Tonin, A. Biasiucci, M. Tavella, A. Molina, A. Al-Khodairy, T. Carlson, and J.d.R. Millán. Transferring brain-computer interfaces beyond the laboratory: Successful application control for motor-disabled users. *Artif Intell Med*, 59(2):121–132, 2013.
- [3] Z.W. Pylyshyn and R.W. Storm. Tracking multiple independent targets: evidence for a parallel tracking mechanism. *Spatial Vision*, 3(3):1–19, 1988.

# Comparing BCI performance using scalp EEG- and inverse solution-based features

Mohit Kumar Goel, Ricardo Chavarriaga and José del R. Millán

Chair in Non-Invasive Brain-Machine Interface (CNBI), EPFL, Switzerland  
{mohit.goel,ricardo.chavarriaga,jose.millan}@epfl.ch

## Abstract

Several studies have proposed the use of inverse solutions based features to improve the decoding performance of brain-computer interfaces. Most of these studies have compared the performance of inverse solutions features over scalp activity in a small set of electrodes. However, the estimated sources are indeed a linear combination of scalp-wide activity. Therefore, this comparison may be biased against surface EEG. Performance comparison in three ERP-based protocols show that classifiers combining larger sets of EEG electrodes may perform comparably, and previous reports may have overestimated the advantages of using inverse solution based features.

## 1 Introduction

Brain-Computer interfaces (BCI) rely on the single-trial recognition of neural signals corresponding to different conditions. Currently, the most common BCI implementations use electroencephalography (EEG) signals elicited after presentation of different stimuli (i.e., event-related potentials, ERP). However, EEG-based applications have low spatial resolution of the signals due to poor skull conductivity which smears the electrical activity originated in cortical sources. This reduces the signal-to-noise ratio (SNR) posing serious challenges for BCI use.

Inverse solution methods, allowing to estimate the intra-cranial sources that generate the scalp measured potentials, have been proposed as a potential method to improve classification performance compared to surface EEG [3, 8, 7]. The rationale is to increase the signal spatial resolution by projecting scalp potentials onto a higher dimensional (source) space. The estimated activity of these sources is expected to have better SNR as they represent unmixed scalp potentials captured by the EEG electrodes. Correspondingly, from a classification perspective, the features extracted from the estimated cortical activity are expected to yield better discrimination between the BCI classes.

Previous studies on the use of inverse solutions for BCI typically compared the performance using source-related features against features from individual or a small number of EEG channels. However, since sources are estimated by linearly combining information from all the EEG channels, is not surprising that source-based features provide higher discriminant information compared to individual scalp-electrodes. Therefore, results may be biased against against surface EEG. In this paper we address this issue by comparing source-based classification with features based on linear combination of scalp electrodes aiming to provide a fair assessment of their capabilities in ERP-based BCIs.

## 2 Methods

We compared the classification performance using scalp- and source-based features. The first classifier, henceforth termed *Cortical*, uses the estimated cortical current density (CCD) of

intracranial sources [3]. Features are extracted from the temporal activity of each source within a given time window (see below) and at each time-point the 100 most discriminant CCD features are selected using the Fisher score. We train individual classifiers for each of these selected features that compute the probability of likelihood to the means and covariance computed on the training data. The output of these classifiers are combined using naïve Bayes rule. The motivation for this ensemble method is that using individual classifiers per feature reduces the possibility of overfitting when limited number of training trials are available

In the case of EEG features, we compute the classification performance using a Fisher linear discriminator (*FLD*). This classifier combines linearly the activity from all the electrodes. Thus mathematically, it is comparable to the projection used to compute one intracranial source. However, projection weights in the FLD are optimized for separating the classes while the inverse solution optimize localization capabilities. Since the *Cortical* classifier combines several projections (i.e. sources) at each time point while the FLD only uses one, we also tested two more classifiers that combine multiple linear projections of the scalp activity. In the first one these projections were obtained through bagging, i.e. multiple FLD ( $N=100$ ) are built from subsets of the training data. Each subset was obtained by randomly selecting 50% of trials from the training set. This classifier is referred to as *EEG N-FLD* to signify multiple FLD projections whereas the earlier classifier is referred to as *EEG 1-FLD*. Furthermore, since the bagging approach may lead to similar projections, we also tested a method, termed *Orthogonal*, that enforces the use of orthogonal FLD projections [5]. Last but not least, as it has been done in previous works on decoding using inverse solutions [3, 8, 4], we also tested classifiers trained on a small subset of electrodes, here referred to as *EEG-Channel* classifier.

These classifiers were compared offline in three standard ERP-based BCI experiments: Rapid Serial Visual Presentation (RSVP) [10], P300-speller [6] and Error-related potentials (ErrP) [2]. In the RSVP study ( $N=15$ ), sequences of images are rapidly presented (4 images per second) and the evoked EEG activity is decoded looking for signatures of target and distractor images. Training data was acquired during 4 search tasks, each one composed of two sequences of 200 images. The testing phase was performed on three search tasks using different target objects than for training. Both training and test phases were performed on the same day. 64 EEG channels were acquired at 2048 Hz, then filtered in the range 1-10 Hz, spatially filtered using CAR and downsampled to 32Hz. Classification was based on the data in the time window [200-700] ms. Following previous works, channels Cz, Pz and Oz were used for the *EEG-Channel* classifier [9].

The second experiment ( $N=8$ ) consists on the standard P300 matrix speller. The experiment was performed over two days (average separation of 12 days) and each day an average of 5 runs were performed per subject. Each run requires the writing of 5 characters. Data of the first day was used for training and classifiers were tested on the data for the second day. 61 EEG channels were recorded at 250Hz, then filtered in the [1 20] Hz range, downsampled to 50 Hz and CAR referenced. Signal in the window [100 600] ms was selected for classification. The *EEG-Channel* classifier used channels Fz, Cz and Pz [1].

Finally, in the ErrP ( $N=6$ ) experiment subjects monitor a cursor that moves horizontally in discrete steps towards a target location. 20% of the time it goes in the opposite direction to the target. The experiment was performed over two days with a separation ranging from 2 months to 2 years. Classifiers were trained on day 1 and tested on the second day. 64 EEG channels were acquired at 2048 Hz, then filtered in the range 1-10 Hz, spatially filtered using CAR and downsampled to 32Hz. Classification was based on the data in the time window [200-450] ms. Features from channels Fz, Cz and Pz were used by the *EEG-Channel* classifier [2].

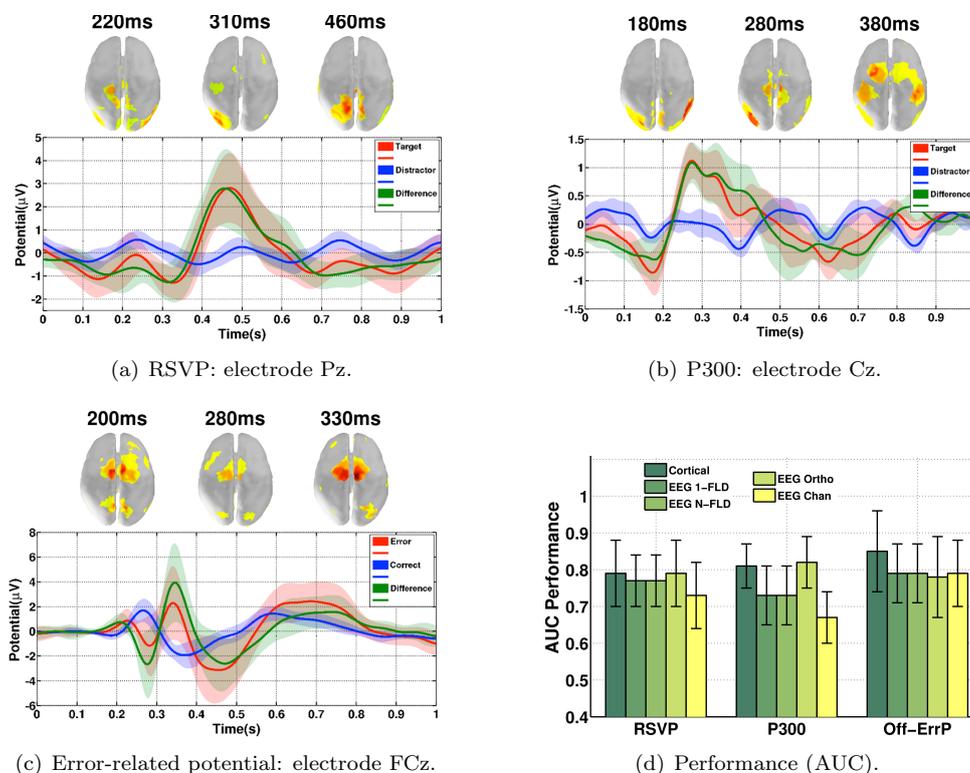


Figure 1: (a-c) Event-related potentials at discriminant electrodes. Topographic Localization of discriminant sources (Top view) at selected time points is shown in the top row. (d) Classification performance on the three ERP protocols.

### 3 Results

Fig 1(a-c) shows representative ERPs and discriminant intracranial sources for the three protocols. Sources are color coded from green to red to show how often they were found to be discriminant across subjects (red color indicates features that were discriminant in all subjects). Classification performance (area under the ROC curve, AUC) is shown in Figure 1(d). In general *EEG-Channel* classifier exhibits the lowest performance. In the RSVP protocol the *Cortical* classifier outperforms the *EEG-Channel* but failed to show statistically significant differences ( $p = 0.08$ , Wilcoxon). It also performs comparably to the other three classifiers, all using linear combinations of EEG channels. The performance obtained with inverse solution is similar to a previous study using a full brain head model based inverse solution [9]. In contrast, in the P300 experiment the *Cortical* and *Orthogonal* methods significantly outperform the other three classifiers ( $p < 0.05$ , Wilcoxon). The results with our classifiers are comparable to previous EEG-based studies [6]. In the case of the ErrP protocol, on average the *Cortical*, *1-FLD*, and *N-FLD* show higher performance. Nonetheless, no significant differences were found with respect to the *Orthogonal* and *EEG-Channel*. Overall, performance of the *1-FLD* and *N-FLD* classifiers was comparable across protocols. This implies that the bagging procedure yielded similar projections and therefore no improvement with respect to the initial classifier.

## 4 Conclusion

We found that inverse solution based classifiers have consistent classification performance for all the experimental protocols and the results matches with the state of the art methods. Contrasting with previous reports, most of them on SMR-based BCIs, surface EEG based classifiers yielded similar performance, in particular when multiple channels are considered. There is thus, at least for ERP-based BCIs, the risk of overestimating the advantages of source-based classification if the proper validation is not performed. Noticeably, bar a few exceptions [7], previous works on SMR classification do not report such comparison.

Nevertheless, the *Cortical* classifier outperformed the *FLD* scalp-based classifiers. This suggests that these classifiers can indeed bring some advantages. Remarkably, testing data in two of the experiments were obtained on a different day than for training, suggesting that source estimation can still be reliable despite removal and relocation of the EEG electrodes across sessions. Further studies are required to better identify those cases where the use of inverse solutions may be more suitable for BCI systems than scalp-recorded signals.

## 5 Acknowledgments

This study is supported by the Swiss National Science Foundation (grant 200021-120293, www.snf.ch). We thank Fabio Aloise and the IRCCS Fondazione Santa Lucia, Rome for providing the P300 dataset.

## References

- [1] C Bledowski, D Prvulovic, and K Hoechstetter. Localizing P300 generators in visual target and distractor processing: A combined event-related potential and functional magnetic resonance imaging study. *Journal of Neuroscience*, 24:9353–9360, 2004.
- [2] R. Chavarriaga and J. del R. Millán. Learning from EEG error-related potentials in noninvasive brain-computer interfaces. *IEEE Trans Neural Syst Rehabil Eng*, 18(4):381–388, 2010.
- [3] F. Cincotti, D. Mattia, F. Aloise, S. Bufalari, L. Astolfi, F. Fallani, A. Tocci, L. Bianchi, M.G. Marciani, S.Gao, J.dR. Millán, and F. Babiloni. High-resolution EEG techniques for brain-computer interface applications. *J Neurosci Methods*, 167(1):31–42, 2008.
- [4] M. Congedo, F. Lotte, and A. Lécuyer. Classification of movement intention by spatially filtered electromagnetic inverse solutions. *Physics in Medicine and Biology*, 51(8):1971–1989, 2006.
- [5] U. Hoffmann, J.-M. Vesin, and T. Ebrahimi. Spatial filters for the classification of event-related potentials. In *14th European Symposium on Artificial Neural Networks*, 2006.
- [6] D.J Krusienski, E.W Sellers, F. Cabestaing, S.Bayoudh, D.J. McFarland, T.M Vaughan, and J.R Wolpaw. A comparison of classification techniques for the P300 speller. *J Neural Eng*, 3(4):299–305, 2006.
- [7] F. Lotte, A Lécuyer, and B. Arnaldi. FuRIA: an inverse solution based feature extraction algorithm using fuzzy set theory for brain-computer interfaces. *Trans. Sig. Proc.*, 57(8):3253–3263, 2009.
- [8] Q. Noirhomme, R. I. Kitney, and B. Macq. Single-trial EEG source reconstruction for brain-computer interface. *IEEE Transactions on Biomedical Engineering*, 55:1592–1601, 2008.
- [9] P. Poolman, R. M Frank, P. Luu, S. M Pederson, and D. M Tucker. A single-trial analytic framework for EEG analysis and its application to target detection and classification. *Neuroimage*, 42(2):787–798, 2008.
- [10] M. Uscumlic, R. Chavarriaga, and J.dR. Millán. An iterative framework for EEG-based image search: Robust retrieval with weak classifiers. *PLoS ONE*, 8:e72018, 2013.

# Plasticity following skilled learning and the implications for BCI performance

Natalie Mrachacz-Kersting<sup>1</sup>, Ning Jiang<sup>2</sup>, Ren Xu<sup>2</sup>, Kim Dremstrup<sup>1</sup>  
and Dario Farina<sup>2</sup>

<sup>1</sup> University of Aalborg, Denmark

<sup>2</sup> Georg-August University, Goettingen, Germany

nm@hst.aau.dk, ning.jiang@bccn.uni-goettingen.de, ren.xu@bccn.uni-goettingen.de, kdn@hst.aau.dk, dario.farina@bccn.uni-goettingen.de

## Abstract

Brain Computer Interfaces (BCIs) rely on robust detection and classification algorithms for stable and accurate device control. Here we investigate the effects of skilled motor learning on the stability of movement related cortical potentials (MRCP). Eight subjects were trained for 32 min on a task known to induce plasticity. Significant increases in performance were accompanied by significant increases in the variability of the EEG signal from two to one second prior to movement onset. The implications of these results for the use of the MRCP for online device control are discussed.

## 1 Introduction

Current research on brain computer interfaces (BCIs) has been almost exclusively limited to a single session trial where the performance is of acceptable accuracy (Wolpaw, 2013). Essentially, the user is trained to fit into the performance of the BCI governed by a classifier or a detector. But, the brain is undergoing continuous adaptation through learning and shifts in attention to task (Sanes & Donoghue, 2000). While the human nervous system is adaptable, current BCIs lack this capability. There is a need for adapting the BCI to the user, thus paving the way for bidirectional adaptation of either the BCI or the user, depending on the reasons for alterations in brain activity. Any design of a BCI must use the changing brain and thus the principles related to how the brain acquires, improves and maintains its natural function as a guide.

For the past 10 years, our research group has successfully implemented the slow cortical potential arising during movement or its intention as a control signal, also called the MRCP. We have shown that a MRCP-based BCI can detect movement or movement intention at high (above 75%) accuracy in the online mode (Xu *et al.*, 2014) and due to the possibility of such detection prior to movement onset, it provides a fast device control. In this study we aim to quantify the alterations in the MRCP morphology after healthy subjects participate in a novel skill learning motor task. We predict that since skill-learning is associated with significant plasticity at the level of the motor cortex, MRCP will be affected significantly.

## 2 Methods

Eight healthy, volunteers (3 females, 5 males; 25-43 years) with no prior history of neurological conditions participated in this study. All procedures were approved by the Scientific Ethics Committee of Northern Jutland (Reference number: N-20130039) and subjects gave their written consent.

### 2.1 EEG and EMG recordings

Ten channels of monopolar EEG were collected using the EEG electrode system and g.USBamp amplifier from gTec, GmbH at a sampling frequency of 256 Hz. Electrodes were placed on FP1, Fz, FC1, FC2, C3, Cz, C4, CP1, CP2 and Pz, according to the standard international 10-20 system. The ground electrode and reference electrode were placed on Fpz and the right earlobe, respectively.

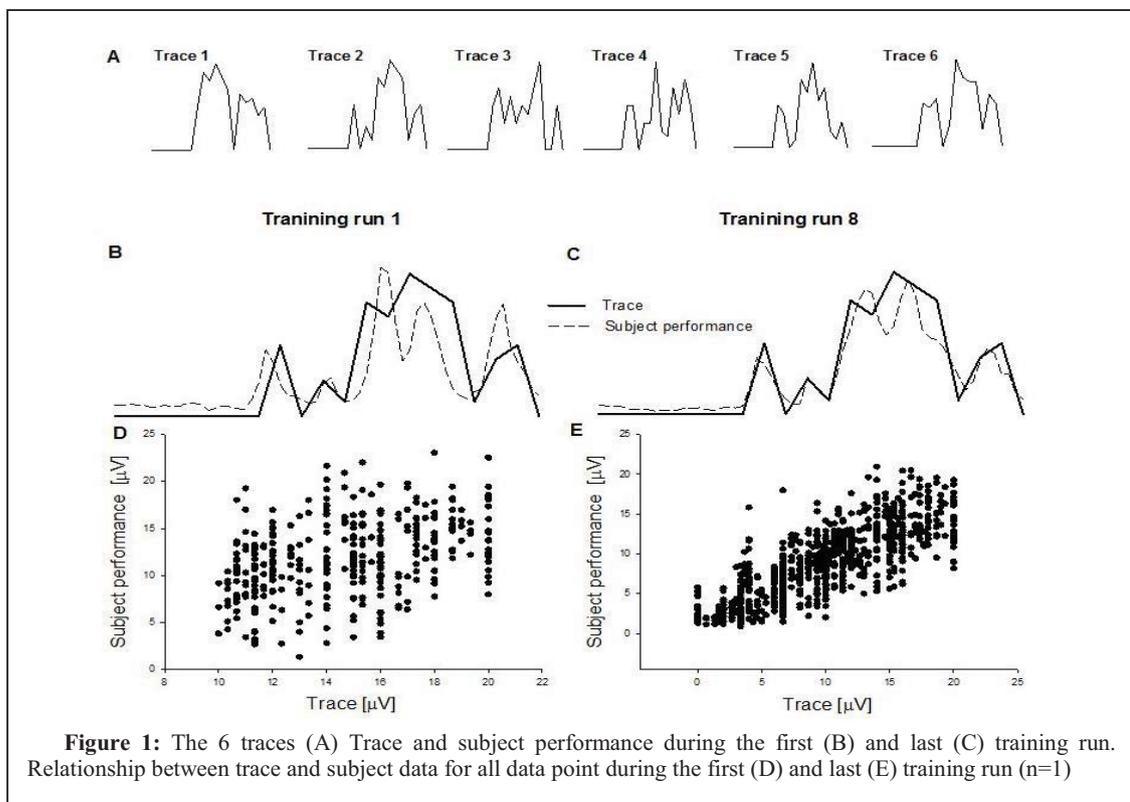
Surface electrodes (20 mm Blue Sensor Ag/AgCl, AMBU A/S, Denmark) were used to record the electromyographic (EMG) activity of the tibialis anterior (TA) muscle of the dominant leg for all aspects of the experiments. EMG data were collected at a sampling frequency of 2000 Hz using a custom made Labview program (Follow-me, Knud Larsen, Aalborg University).

### 2.2 The learning protocol

Subjects were seated comfortably in an armchair with the dominant leg flexed in the hip ( $120^\circ$ ), the knee ( $160^\circ$ ) and the ankle ( $110^\circ$  of plantarflexion). The foot was resting on a foot-plate and a computer screen was positioned approximately 1.5 m in front of the subjects at eye level. The learning task was comprised of a series of six randomized figures each sketching a different series of combinations of dorsi- and plantarflexion movements (Figure 1A). Upon appearance of the trace on the screen, a countdown of 3 s was visually shown and on the word 'go', the activity of the subjects EMG was displayed in real-time overlaying the respective trace. Subjects were instructed to follow these traces by controlling the activation level of their TA muscle. Each trace lasted 3-4s and a single training run lasted for 4 min followed by a 2 min rest period. A total of eight training runs were completed leading to a total training time of 32 min.

### 2.3 Data analysis

Analysis of the EEG data was performed for training run one and eight and only from the Cz channel since Cz is located over that area of the motor cortex that has direct connections to the target muscle TA. It is here where most alterations due to the learning paradigm are expected based on past studies using a similar paradigm. EEG were band pass-filtered (0.05-10 Hz, 2nd order Butterworth filter) and continuous EEG data divided into epochs of 4 s (from 2 s before to 2 s after the onset of EMG in the TA). The onset of the TA EMG activity was used as time zero since it is at this time-point when the subjects commenced to track the traces and where an MRCP is expected to occur. Following appropriate segmentation, the following parameters were extracted from single trials during the first and last training run respectively: (i) average EEG signal from -2 to -1 s prior to task onset, (ii) standard deviation of the signal from -2 to -1 s prior to task onset, (iii) the peak negative value timing within a 500 ms window on either side of the task onset and (iv) the amplitude of the peak negative value (Figure 2A). To quantify any alterations in performance of the task, the correlation between the trace and the actual subject performance was calculated for the first and the last 4 min of training. Paired t - test was used to compare the  $r^2$  values, the PN latency, amplitude and the mean EEG signal and its SD for the first and last training run respectively. Significance level was set to  $p < 0.05$ .



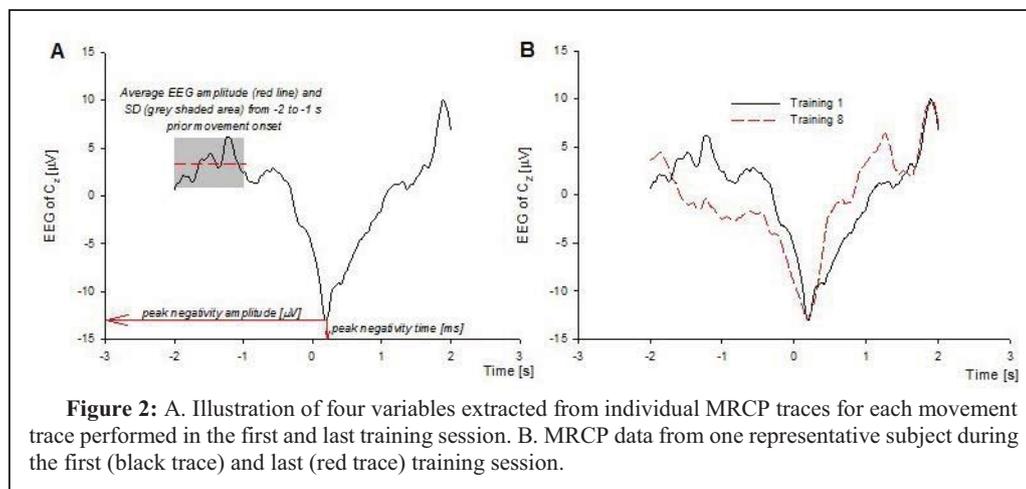
### 3 Results

#### 3.1 Motor performance

The effect of the training for one subject is illustrated in Figure 1B-E. Figure 1B and 1C show the data for one of the traces selected for analysis during training 1 and training 8, while Figure 1D and 1E shows the scatter plot between the exerted TA activation and the target level for each interval of time during training runs 1 and 8. For this subject the correlation between shown traces and actual subject performance increased from 0.84 to 0.91. Across all subjects there was a significant increase in correlation between the first and last session ( $p=0.0005$ ) indicating an improved performance.

#### 3.2 Movement related cortical potential (MRCP) during the tracking task

Four parameters were extracted from the MRCPs generated during training session 1 and 8 (Figure 2A). Figure 2B shows the average MRCP for training session 1 and 8 for one subject. The average EEG amplitude within 1-2 s prior to movement onset remained relatively stable (session 1:  $-4.23 \mu\text{V}$  vs session 8:  $-3.39 \mu\text{V}$ ;  $p=0.42$ ). However, the variability of the EEG activity during this time (the SD within the time window), increased significantly (session 1:  $1.96 \mu\text{V}$  vs session 8:  $2.54 \mu\text{V}$ ;  $p=0.01$ ). Both the time of peak negativity in relation to task onset (session 1:  $-40 \text{ ms}$  vs session 8:  $-50 \text{ ms}$ ,  $p=0.35$ ) and its amplitude (session 1:  $-17.07$  vs session 8:  $-17.92$ ,  $p=0.36$ ) did not change significantly.



## 4 Discussion

Results demonstrate that significant improvements in task performance are accompanied by a significantly greater variability of the MRCP 2-1 s prior to movement onset. This has important implications in the design of any MRCP-based BCI for online detection of movement. Here the algorithm has a role in identifying when the EEG activity attains a specific threshold level. Once detected, the algorithm proceeds to implement device control. For a BCI designed for neuromodulation this may take the form of turning on an orthotic device that performs a desired movement or an electrical stimulator that triggers nerve stimulation to induce contraction of specific muscles. Typically such a use of BCI intends to induce plasticity at the level of the motor cortex. Data presented here suggested that the performance of the algorithm could be affected as plasticity is induced. Previous studies have shown significant plasticity within the motor cortex and likely also connected sites during the acquisition of a new motor skill. Such plasticity results in an expansion of the motor maps – i.e. those areas of the motor cortex that when stimulated will result in a contraction of the target muscle. As the motor skill manifests itself, the motor maps tend towards their normal size. However evidence also exists that motor maps differ between skilled versus novice athletes. Any BCI designed to induce neuromodulation must ensure that shifts in EEG patterns are taken into account with the acquisition of new or indeed the relearning of old motor skills. The MRCP seems to be robust to effects of motor learning at least in healthy subjects. However, since online detecting relies on threshold values for the period prior to movement onset, the findings here suggest that BCI performance will decline with skill acquisition unless the algorithm is adapted to the new level of activity. This is one of our current areas of investigation.

## References

- Sanes, J. & Donoghue, J. (2000). Plasticity and Primary Motor Cortex. *Rev Neurosci* **23**, 393-415.
- Wolpaw, JR (2013). Brain-computer interfaces. *Handb Clin Neurol* **110**, 67-74.
- Xu R., Jiang N., Mrachacz-Kersting N., Lin C., Asin G., Moreno J., Pons J., Dremstrup K., & Farina D. (2014). A Closed-Loop Brain-Computer Interface Triggering an Active Ankle-Foot Orthosis for Inducing Cortical Neural Plasticity. *IEEE Trans Biomed Eng*. In press. (DOI: 10.1109/TBME.2014.2313867)

# Dual-Frequency SSVEP Stimulation Using a Stereoscopic Display

Leandra Webb<sup>1</sup>, Jozef Legény<sup>2</sup>, Robert Leeb<sup>3</sup> and Anatole Lécuyer<sup>2</sup>

<sup>1</sup>MRC/UCT Medical Imaging Unit, Division of Biomedical Engineering, Department of Human Biology, University of Cape Town, South Africa, <sup>2</sup>Inria, Rennes, France,

<sup>3</sup>École Polytechnique Fédérale de Lausanne, Switzerland

[lwebb@csir.co.za](mailto:lwebb@csir.co.za), [robert.leeb@epfl.ch](mailto:robert.leeb@epfl.ch), [anatole.lecuyer@inria.fr](mailto:anatole.lecuyer@inria.fr)

## Abstract

This study aimed to investigate the feasibility of dual-frequency Steady-State Visual Evoked Potential (SSVEP) stimulation using a 3-D display and stereoscopic glasses. Participants were exposed to a repetitive visual stimulus flashing at different frequencies in the left and right views and the SSVEP responses were observed. The results suggest that the two stimulation frequencies can still be evident in the SSVEP response. In addition, the participant ratings showed no significant differences in fatigue, annoyance, discomfort or strangeness of the stimulation compared to conventional forms of stimulation. These results pave the way for further studies using stereoscopic dual-frequency stimulation and its potential for use in virtual reality and 3-D videogames.

## 1 Introduction

Steady-State Visual Evoked Potentials (SSVEP) have received much attention in Brain-Computer Interface (BCI) research. However, a challenge still faced in SSVEP BCIs is the limited number of unique frequencies available for stimulation [1].

To counteract this problem, researchers have begun to investigate the use of dual-frequency stimulation, since this allows more targets to be created using a small number of frequencies. This has been done either by overlaying two targets oscillating at different frequencies [1, 2] or displaying two separate targets at different frequencies [3, 4].

In this paper, a novel method of providing dual-frequency SSVEP stimulation is investigated, which involves using a stereoscopic display, i.e. 3-D screen and stereoscopic glasses. Stereoscopic vision lends itself well for use with dual-frequency stimulation since the views to the left and right eyes can be controlled independently, which may make the concept visually more transparent to the user. Furthermore, the method is easily implementable due to the availability of low-cost equipment and easily extendable to three-dimensional videogames and virtual reality BCIs [6].

## 2 Method

To test the concept of stereoscopic dual-frequency SSVEP stimulation, an offline experiment was conducted.

**Population:** Ten participants (aged 24-33, 8 male, 2 female) participated in the experiment. All of these participants reported normal or corrected-to-normal vision. One participant was excluded from the final analysis as no SSVEP response was evident in the EEG trace.

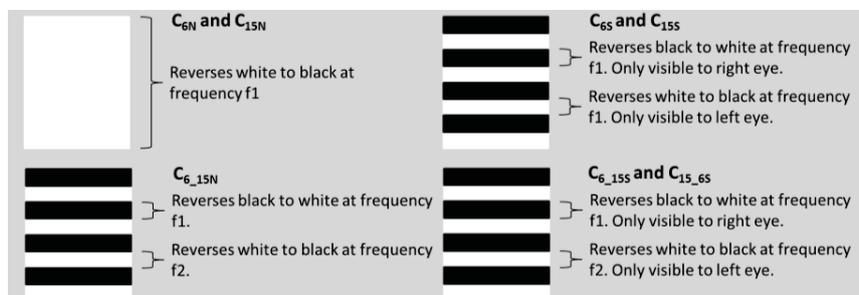
**Experimental Setup:** The software used for creating and running the experiments was *OpenViBE* (<http://openvibe.inria.fr/>). An LCD 3-D television screen was used for displaying the SSVEP targets. Stereoscopic polarized glasses were worn by the participants for all of the stereoscopic conditions. For acquiring data the *g.Tec* USBamp was used to record data from thirteen electrodes (CPz, POz, PO7, PO8, O1, O2, Oz, Iz, Pz, PO3, PO4, O9 and O10).

**Procedure:** Participants were seated 1.5 m in front of the screen. They received six different forms of stimulation, as shown in Table 1. Frequencies of 6 Hz and 15 Hz were chosen as both showed high SSVEP responses in pre-testing and were possible frequencies of the refresh rate of the screen. Two normal vision and two stereoscopic vision single-frequency conditions were carried out as control conditions (Conditions  $C_{6N}$ ,  $C_{15N}$ ,  $C_{6S}$  and  $C_{15S}$ ), in which a block on the screen was inverted from black to white at a single frequency. For the stereoscopic single-frequency conditions ( $C_{6S}$  and  $C_{15S}$ ) half of the image was displayed to the right eye and the other half to the left eye through the use of the 3-D display as illustrated in Figure 1. To simulate previous work in dual-frequency stimulation, the block was divided into lines, with half flashing at 6 Hz and half at 15 Hz ( $C_{6,15N}$ ). To achieve dual-frequency stimulation with stereoscopic vision, the block was again divided into lines but each eye was only exposed to half of the lines (and thus only one of the frequencies) through the use of stereoscopic glasses and the 3-D mode of the screen. This was repeated with the frequencies to the left and right eyes reversed to overcome any biasing due to ocular dominance ( $C_{6,15S}$  and  $C_{15,6S}$ ).

For each condition, eight runs of stimulation were performed for eight seconds each, with three second breaks. The order of the conditions was rotated pseudo-randomly among participants to prevent a fatigue effect. Between each condition there was a rest period of approximately two minutes. The conditions were also divided into stereoscopic and normal vision blocks and rotated. Half of the participants started with a normal vision block while the other half started with a stereoscopic block. After each condition participants were asked to rate the condition using a 10-point Likert scale based on four different criteria: Fatigue, Annoyance, Discomfort and Strangeness.

**Table 1:** Stimulation frequencies for the different conditions of the experiment.

Condition	Target
$C_{6N}$	6 Hz normal vision
$C_{15N}$	15 Hz normal vision
$C_{6S}$	6 Hz stereoscopic vision
$C_{15S}$	15 Hz stereoscopic vision
$C_{6,15N}$	6 Hz and 15 Hz normal vision
$C_{6,15S}$	6 Hz (left) and 15 Hz (right) stereoscopic vision
$C_{15,6S}$	15 Hz (left) and 6 Hz (right) stereoscopic vision



**Figure 1:** Illustrations of the various forms of stimulation of the experiment.

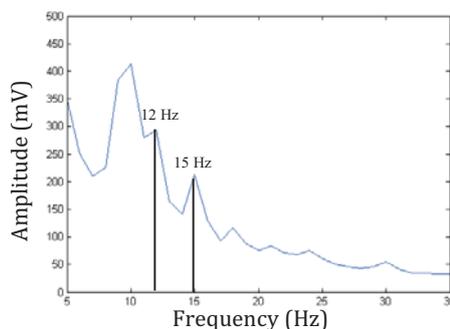
### 3 Results

**EEG Data Results:** The signals were divided into non-overlapping two-second epochs on which a 1024 point Fast Fourier Transform (FFT) was carried out. The FFTs for all the epochs were then averaged. Figure 2 shows the spectrum averaged over nine participants and averaged over electrodes O1, O2 and Oz for the dual-frequency stereoscopic conditions  $C_{6\_15S}$  and  $C_{15\_6S}$ . The two spectrums were averaged in order to prevent any biasing due to ocular dominance.

Power values were calculated from the FFT and then converted to relative power values, in order to cancel differences in magnitude, as follows:

$$Power_{Relative} = 3 * Power(f) / (Power(f-1) + Power(f) + Power(f+1))$$

where  $Power(f)$  is the average power at the relevant frequency and  $Power(f - 1)$  and  $Power(f + 1)$  are the average power at frequencies 1 Hz above and below the relevant frequency respectively.



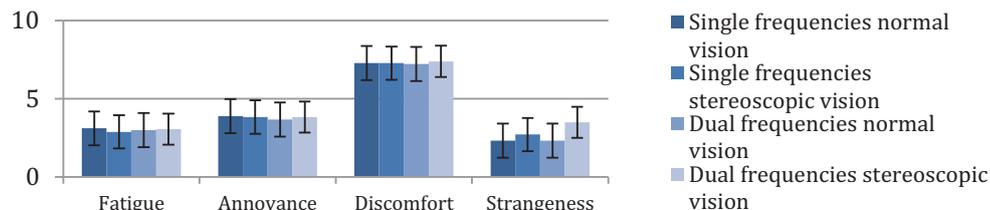
**Figure 2:** Averaged spectrums for the stereoscopic dual-frequency conditions ( $C_{6\_15S}$  and  $C_{15\_6S}$ ) using 6 Hz and 15 Hz stimulation, with peaks at 12 Hz and 15 Hz indicated.

Paired t-tests were carried out between the relative power at each frequency of interest and the average relative power of the two neighbouring frequencies. The results are shown in Table 2. The single-frequency control conditions show statistically significant increases in power at the stimulation frequencies. For the dual-frequency stereoscopic stimulation ( $C_{6\_15S}$  and  $C_{15\_6S}$ ), the 6 Hz stimulation still shows a statistically significant increase in power ( $p < 0.05$ ) at 12 Hz (the first harmonic). For the 15 Hz stimulation, the increase in power is tending towards statistical significance ( $p < 0.1$ ). This suggests that the two stimulation frequencies can still be identified in the dual-frequency stereoscopic case. A possible reason that not all power differences were significant is the decreased luminosity due to the use of stereoscopic glasses with tinting, or the fact that each eye was only viewing half of the lines of the image as opposed to all lines in the normal vision cases, as shown in Figure 1.

**Table 2:** Results of paired t-tests using relative power at relevant frequencies and neighbouring frequencies.

Condition	Comparison	t value	Significance
$C_{6N}$	12 Hz to average of 11 and 13 Hz	6.611	< 0.001*
$C_{15N}$	15 Hz to average of 14 and 16 Hz	5.634	0.001*
$C_{6S}$	12 Hz to average of 11 and 13 Hz	5.608	0.001*
$C_{15S}$	15 Hz to average of 14 and 16 Hz	6.844	< 0.001*
$C_{6\_15N}$	12 Hz to average of 11 and 13 Hz	1.587	0.164
	15 Hz to average of 14 and 16 Hz	2.802	0.031*
$C_{6\_15S}$	12 Hz to average of 11 and 13 Hz	2.954	0.025*
	15 Hz to average of 14 and 16 Hz	2.151	0.075
$C_{15\_6S}$	12 Hz to average of 11 and 13 Hz	3.492	0.010*
	15 Hz to average of 14 and 16 Hz	2.211	0.063

**Participant Questionnaires:** The averaged participant ratings are shown in Figure 3, grouped according to the type of stimulation. A Friedman test revealed no significant differences in any ratings ( $p > 0.05$ ). This suggests that the participants did not find the stereoscopic dual-frequency stimulation more fatiguing, annoying, uncomfortable or strange than the other forms of stimulation. The strangeness rating did show an increase tending towards significance ( $p < 0.1$ ), thus the participants may have noticed a slight difference in appearance but not one that affected their other perceptions.



**Figure 3:** Participant ratings of Fatigue, Annoyance, Discomfort and Strangeness of each form of stimulation.

## 4 Conclusions

The aim of the experiment was to determine whether stereoscopic dual-frequency stimulation is feasible. The results indicated that the two stimulated frequencies are evident in the SSVEP response. Changes in the experimental setup may be required to achieve more significant responses, for example, by increasing the luminosity of the targets. In addition, further combinations of frequencies should be tested. The concept could then be extended to 3-D videogames and virtual reality setups. The participants did not rate the stereoscopic dual-frequency stimulation as more fatiguing, annoying, uncomfortable or strange, suggesting that this form of stimulation is also feasible from the point of view of participant preferences.

**Acknowledgements:** This work was made possible in part by funding from a UCT Scholarship for International Travel and the National Research Foundation (NRF), South Africa.

## References

- [1] D. Zhu, J. Bieger, G. Garcia Molina, and R. M. Aarts, "A survey of stimulation methods used in SSVEP-based BCIs.," *Comput. Intell. Neurosci.*, vol. 2010, Jan. 2010.
- [2] M. Cheng, X. Gao, S. Gao, and D. Xu, "Multiple Color Stimulus Induced Steady State Visual Evoked Potentials," in *23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2001.
- [3] H. Hwang, D. H. Kim, C. Han, and C. Im, "A new dual-frequency stimulation method to increase the number of visual stimuli for multi-class SSVEP-based brain – computer interface (BCI)," *Brain Res.*, 2013.
- [4] S. T. Morgan, J. C. Hansen, and S. A. Hillyard, "Selective attention to stimulus location modulates the steady-state visual evoked potential.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 93, no. 10, pp. 4770–4, May 1996.
- [5] S. P. Kelly, E. C. Lalor, R. B. Reilly, and J. J. Foxe, "High-Density SSVEP Data for Independent Brain – Computer Communication," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 13, no. 2, pp. 172–178, 2005.
- [6] A. Lécuyer, L. George, and M. Marchal, "Toward Adaptive VR Simulators Combining Visual, Haptic, and Brain-Computer Interfaces," *IEEE Comput. Graph.*, vol. 33, no. 5, pp. 18–23, 2013.

# Spatial masking might increase classification accuracy in tactile ERP-BCIs

Andreas Herweg<sup>1</sup>, Tobias Kaufmann<sup>1,2</sup> and Andrea Kübler<sup>1</sup>

<sup>1</sup> University of Würzburg, Würzburg, Germany

<sup>2</sup> University of Oslo, Oslo, Norway

andreas.herweg@uni-wuerzburg.de

## Abstract

Tactile ERP-BCIs provide unique advantages over visual ERP-BCIs but suffer from lower classification accuracies due to lower discrimination between target and non-target ERPs. In this study we analyzed whether spatial masking can reduce non-target ERPs to improve discrimination. Two different tactile stimulation setups designed to produce simultaneous stimulation with and without masking were compared to a classic oddball stimulation in three healthy subjects. Spatial masking decreased the average number of sequences required to achieve 100% classification accuracy. Preliminary data indicate that spatial masking might have a beneficial effect on classification accuracy.

## 1 Introduction

Whenever possible assistive technology should strive to be unobtrusive for the end-user. Visual ERP-BCIs generally restrict the end-users' perception of his environment and are difficult to conceal due to the necessary display devices. Tactile ERP-BCIs on the other hand allow the user to retain his visual and auditory senses and tactile stimulation devices (tactors) can easily be hidden below layers of clothing. While information transfer rates of most tactile BCIs are lower than those of comparable visual BCIs [1], the unobtrusive nature of tactile ERP-BCIs might be a crucial advantage for certain applications such as wheelchair control [3]. In daily life, vibrotactile stimuli occur rarely compared to visual and auditory stimuli. Therefore, tactile stimuli are difficult to ignore and are most likely perceived as odd stimuli. Consequently tactile stimulation elicits significant ERPs for target and non-target stimuli [3]. While tactile ERPs from target and non-target stimuli are sufficiently different for classification, reduced non-target ERPs may increase classification accuracies. For this purpose, we investigated whether we can accustom a user to the tactile stimulation to reduce non-target ERPs. We applied spatial masking to allow for continuous stimulation on all tactors while still maintaining an oddball task. When multiple tactile stimuli are applied at the same time, but one stimulus is significantly stronger than the remaining ones, the strongest stimulus can mask the other stimuli [7]. Therefore, only one stimulus is perceived despite all locations being stimulated. In this study, we investigated the effect of spatial masking on classification accuracy.

## 2 Methods

Different stimulation patterns were applied to N=3 healthy subjects using 8 vibrate transducers (C2 tactors; Engineering Acoustic Inc., USA). EEG signals were recorded from 16 passiv Ag/AgCl electrodes and amplified using a g.USBamp amplifier (g.tec Engineering GmbH, Austria). Stimulation was applied to the lower and upper arms (see Figure 1) with a stimulus-duration of 240ms and an inter-stimulus-interval of 400ms.

Three different stimulations setups were compared in this study (see Figure 1):

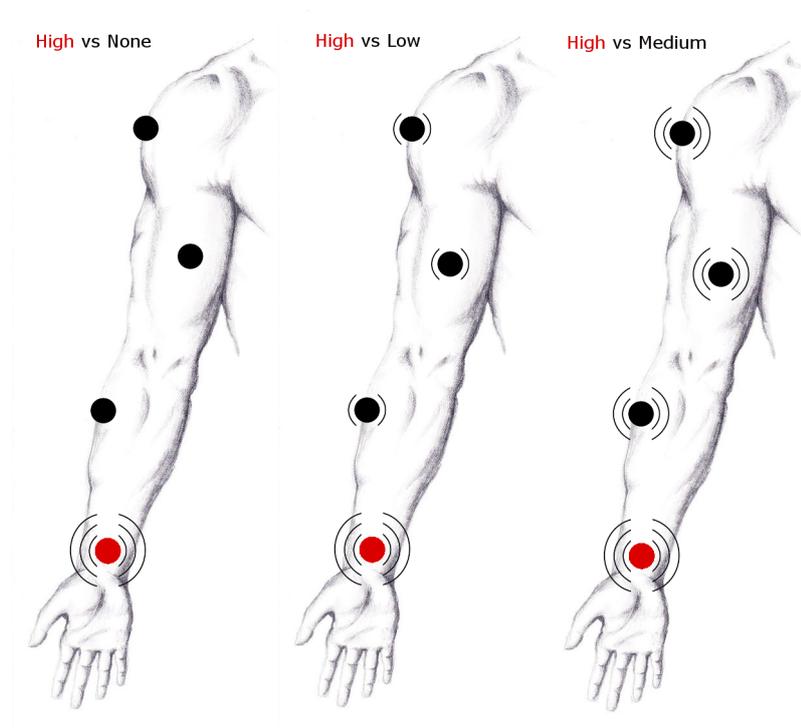


Figure 1: Overview of the three different stimulation setups. Factors were placed along both (right and left) lower and upper arms (red and black dots). The **High vs None** setup was a classical oddball where the target stimulus (red) was vibrating with high intensity and non-target stimuli (black) were inactive. For the **High vs Low** setup the target stimulus (red) was vibrating with high intensity and non-target stimuli (black) were vibrating with low intensity. For the **High vs Medium** setup the target stimulus (red) was vibrating with high intensity and non-target stimuli (black) were vibrating with medium intensity. Target stimuli were presented in randomized order and all stimulations were above detection threshold.

#### **High vs None**

Classical oddball where the target stimulus (odd) is vibrating with high intensity while non-targets (frequent) are inactive.

#### **High vs Low**

Modified oddball where the target stimulus (odd) is vibrating with high intensity while non-targets (frequent) are vibrating with low intensity.

#### **High vs Medium**

Modified oddball where the target stimulus (odd) is vibrating with high intensity while non-targets (frequent) are vibrating with medium intensity.

Subjects were not informed about the difference between the three stimulation setups, but it was ensured that low stimulation was above detection threshold for all subjects. Participants had to perform a calibration-task for each stimulation setup, using 15 sequences, meaning each factor was the target stimulus 15 times. Afterwards they had to perform three copy-spelling tasks, selecting each of the 8 body locations once, with each setup, using 10, 5 and 3 sequences.

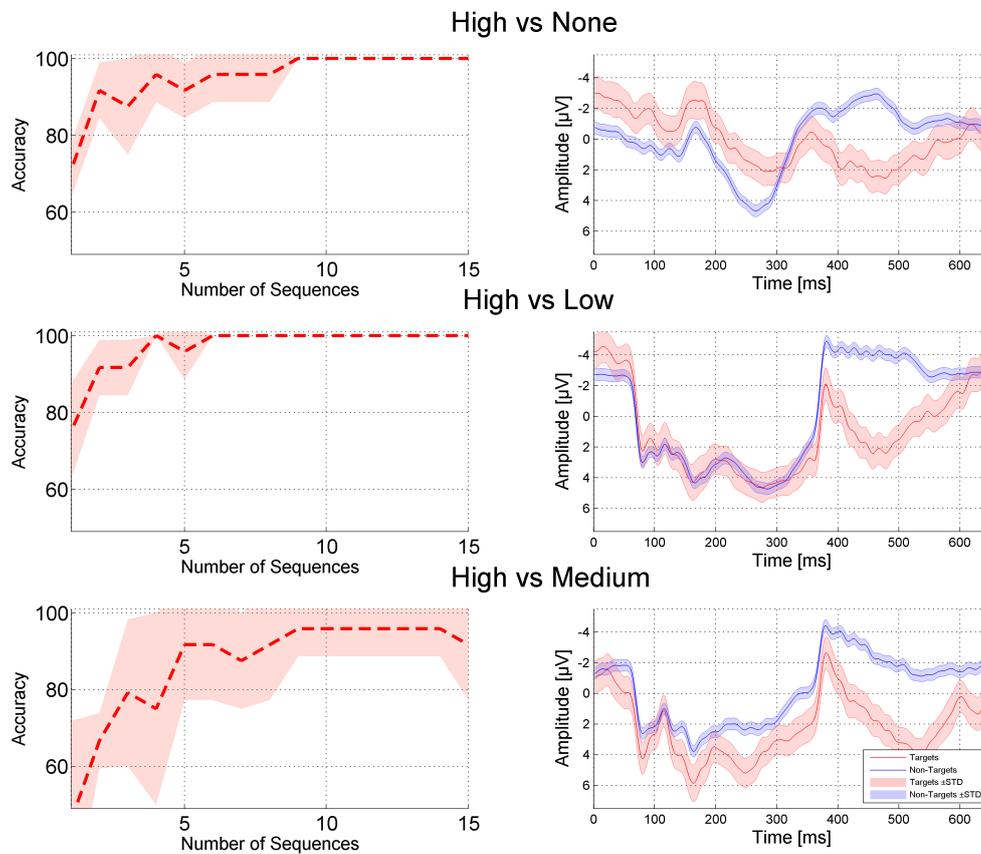


Figure 2: Calibration task results for different stimulation setups. Left: classification accuracy as a function of number of sequences is shown. Right: averaged ERPs for position Cz. Red: Targets, Blue: Non-targets.

### 3 Results

After the experiment, participants were asked to report the felt difference between the **High vs None** and **High vs Low** setup. All participants stated that there was either no difference at all or that they could feel an additional active factor from time to time. None reported that they could reliably feel multiple active factors. For the **High vs Medium** setup all participants reported feeling multiple factors simultaneously.

All participants achieved 100% classification accuracy in 9 sequences or less with the **High vs None** setup (see Figure 2). In the **High vs Low** setup all participants were able to reach 100% classification accuracy with 6 sequences or less. Two participants achieved 100% classification accuracy in the **High vs Medium** setup (4 and 8 sequences needed), one participant was unable to achieve 100% classification accuracy. All stimulation setups elicited event-related-potentials in all participants. Simultaneous activation of all 8 factors resulted in stimulation artifacts in the **High vs Low** and the **High vs Medium** setup (see Figure 2). Performance in the copy-spelling tasks appeared to be equal for all stimulation setups, but due to the small sample size no statistics were calculated (see Table 1).

Table 1: Mean performance in the copy-spelling task

	<b>10 Sequences</b>	<b>5 Sequences</b>	<b>3 Sequences</b>
<b>High vs None</b>	87.5%	70.8%	45.8%
<b>High vs Low</b>	83.3%	75%	50%
<b>High vs Medium</b>	83.3%	70.8%	50%

## 4 Discussion

These preliminary results suggest that spatial masking (**High vs Low**) might have a beneficial effect as it reduced the required number of sequences to reach 100% classification accuracy below that of the default setup (**High vs None**). However no performance difference could be observed in the copy-spelling task for this limited sample. Due to stimulation artifacts a direct comparison between the ERPs of **High vs None** and **High vs Low** condition was difficult. The Source of the artefacts was identified as an unintended flow of current between tactors and skin and will be eliminated for future studies. More data is needed to verify whether tactile masking provides beneficial effects on target or non-target ERPs. Spatial masking is only one out of a multitude of different spatial, temporal and spatio-temporal effects relevant for tactile stimulation [5]. Further research is needed to assess whether additional effects such as apparent location [4], temporal summation [6] and tactile illusions [2] have to be circumvented when designing a tactile BCI or might even prove to be useful for good control of a BCI driven application.

### 4.1 Acknowledgments

This study was funded by the European Community for research, technological development and demonstration activities under the Seventh Framework Programme (FP7, 2007-13), project grant agreement number 288566 (BackHome). This paper reflects only the authors views and funding agencies are not liable for any use that may be made of the information contained herein.

## References

- [1] F. Aloise and I. Lasorsa. Multimodal stimulation for a p300-based BCI. *International Journal of Bioelectromagnetism www.ijbem.org*, 9:128–130, 2007.
- [2] F. A. Geldard and C. E. Sherrick. The cutaneous "Rabbit": a perceptual illusion. *Science*, 178(4057):178–179, October 1972.
- [3] T. Kaufmann, A. Herweg, and A. Kübler. Toward brain-computer interface based wheelchair control utilizing tactually-evoked event-related potentials. *Journal of NeuroEngineering and Rehabilitation*, 11(1):7, January 2014.
- [4] C. E. Sherrick, R. W. Cholewiak, and A. A. Collins. The localization of low- and high-frequency vibrotactile stimuli. *The Journal of the Acoustical Society of America*, 88(1):169–179, July 1990.
- [5] J. van Erp. Guidelines for the use of vibro-tactile displays in human computer interaction. In *Proceedings of Eurohaptics 2002*, 2002.
- [6] R. T. Verrillo and G. A. Gescheider. Enhancement and summation in the perception of two successive vibrotactile stimuli. *Perception & Psychophysics*, 18(2):128–136, March 1975.
- [7] R. T. Verrillo, G. A. Gescheider, B. G. Calman, and C. L. Van Doren. Vibrotactile masking: effects of one- and two-site stimulation. *Perception & psychophysics*, 33(4):379–87, 1983.

# Attention and working memory influence on P300-based BCI performance in people with amyotrophic lateral sclerosis

A. Riccio<sup>1,2</sup>, L. Simione<sup>2,3</sup>, F. Schettini<sup>1,4</sup>, A. Pizzimenti<sup>5</sup>, M. Inghilleri<sup>5</sup>, M. Olivetti Belardinelli<sup>2</sup>, D. Mattia<sup>1</sup>, F. Cincotti<sup>1,4</sup>.

<sup>1</sup>Fond Santa Lucia, Rome, Italy; <sup>2</sup>Dep of Psychology, Sapienza University of Rome, Italy; <sup>3</sup>Institute of Cognitive Sciences and Technologies, CNR, Rome, Italy.; <sup>4</sup>DIAG, Sapienza University of Rome, Italy; <sup>5</sup>Department of Neurology and Psychiatry, Sapienza University of Rome, Italy.  
a.riccio@hsantalucia.it

## Abstract

The purpose of this study was to investigate the support of attentional and memory processes in controlling a P300-based brain-computer interface (BCI) in people with amyotrophic lateral sclerosis (ALS). Eight people with ALS performed two behavioral tasks: i) a rapid serial visual presentation (RSVP) task, screening the temporal filtering capacity and the speed of the update of the attentive filter, and ii) a change detection task, screening the memory capacity and the spatial filtering capacity. The participants were also asked to perform a P300-based BCI spelling task. We found that only the temporal filtering capacity was a predictor of both the P300-based BCI accuracy and of the amplitude of the P300 elicited performing the BCI task. We concluded that the ability to keep the attentional filter active during the selection of a target influences performance in BCI control.

## 1 Introduction

The Brain Computer Interface (BCI) exploits neurophysiological signals to control external devices for a range of applications (Wolpaw & Wolpaw 2012). Reasons for performance variability across people with Amyotrophic Lateral Sclerosis (ALS) and performance predictors of P300-based BCI were not fully investigated. Indeed, the knowledge about the cognitive capabilities reflecting a successful use of P300-based BCI is limited. In particular, people suffering from ALS, included in the range of potential users of BCIs, could have cognitive dysfunctions in association with motoneuron failure, mostly regarded as attention, concentration and verbal fluency (Ringholz et al. 2005). In this study we investigated the influence of the attentive and memory features of people with ALS on performances in a P300-based BCI task.

## 2 Methods

### 2.1 Participants

We recruited the participants at the ALS center of the Policlinic “Umberto I” of Rome. We included in the study a total of nine volunteers, all naïve to BCI training, (3 women; mean age=59.7±12.3) with definite, probable, or probable with laboratory support ALS diagnosis (mean ALSFRS scores: 32.4± 8.2; Brooks et al. 1996). Inclusion criteria required that participants were able

to communicate with or without the help of an AT device which also included an eye tracker, thus all participants had their eye-gaze control preserved. Due to the fact that one participant did not perform the behavioral tasks, only the data of 8 participants out of nine (3 women; mean age=58±12; mean ALSFRS scores: 31.8±8.6) were reported in this article.

The study was approved by the ethic committee of Fondazione Santa Lucia, Rome and all participants provided an informed consent.

## 2.2 Experimental protocol

The experimental protocol consisted in two sessions performed on two different days. In the first session participants were asked to copy spell seven predefined words (5 characters each) by controlling a P300-Speller (Farwell & Donchin 1988). The EEG was recorded using 8 active electrodes (Fz, Cz, Pz, Oz, P3, P4, Po7, Po8). All channels were referenced to the right earlobe and grounded to the left mastoid. EEG was amplified using an 8 channel EEG amplifier (gMobilab, g.tec Austria) and recorded by the BCI2000 software.

During the second session, temporal attention capabilities of participants were screened using a rapid serial visual presentation (RSVP) task: two targets were embedded in a stream of distracter stimuli, all presented at central fixation at a presentation rate of 100ms each. Distracters were black capital consonants. The first target (T1) was a green letter. The second target (T2) was a black capital "X". In 20% of trials T2 was not present, whereas it followed T1 with no (lag 1), one (lag 2), three (lag 4) or five (lag 6) intervening distracters in 20% of trials for each condition. Subject was asked to decide whether the green letter was a vowel and whether the black X was contained in the stimulus stream (Kranzioch et al. 2007).

Memory capacity and attentional spatial filtering capacity were screened by means of two change detection (CD) tasks: a baseline task and a selection task (Vogel et al. 2005). In the baseline task the memory array consisted of three or four rectangles of the same color (all blue or all red) with one out of four possible orientations (vertical, horizontal and two diagonals), presented for 100 ms. The memory array was followed by a retention interval of 900ms and then by a second array of rectangles (test array). The participants were asked to report if the orientation of the rectangles in the test array was identical to the one in the memory array. In the selection task, each memory array consisted of six or eight rectangles. Half of the rectangles were blue and the other half were red. Participants were instructed to memorize the rectangles of one color and to ignore the rectangles of the other color. They were then asked to report if the spatial orientation of the memorized rectangles in the test array was identical to the one in the memory array.

Due to the possible motor disabilities of the experimental group, participants were asked to give a binary response (yes or no) to the operator with the residual communication channel.

## 3 Data analysis

EEG data was high pass and low pass filtered with cut off frequencies of 0.1 Hz and 10 Hz respectively using a 4th order Butterworth filter. In addition, a notch filter was used to remove 50 Hz contamination. Data was divided into 1000 ms long epochs starting with the onset of each stimulus. The amplitude of the P300 potential in Cz was defined as the highest value of the difference between target and non-target average waveforms in the time interval 250-700ms (P300amp).

To provide an estimate of the classifier accuracy we considered the binary classification problem target vs. non-target (binary accuracy, BA; Blankertz et al. 2011). A 7-fold cross-validation was used to evaluate the binary accuracy of the classifier on each participant's dataset by applying a Stepwise Linear Discriminant Analysis (SWLDA) on the testing dataset (6 words) and assessing the binary accuracy on the training dataset (the remaining word).

The detection accuracy of T1 (T1%) in the RSVP task was considered as an index of the temporal attentional filtering capacity of the participants. Because the detection accuracy of T2 (T2%) was considered as an index of the capability to adequately update the attentive filter, only trials in which T1 had been correctly identified were selected in order to determine T2%.

To investigate the memory capacity, according to Cowan (2001), we defined the number of items held in memory (K) as  $K=S(H-F)$ , where S is the size of the array (highest number of item to memorize,  $S=4$ ), H is the observed hit rate and F is the false alarm rate. We calculated the K index for the baseline task ( $K_b$ ) and for the selection task ( $K_s$ ). To screen for the attentional spatial filtering capacity ( $\alpha$ ), of the participants (that is the capacity to efficiently filter the distracters) we subtracted the  $K_s$  from the  $K_b$  ( $\alpha=K_b-K_s$ ).

Because variables were normally distributed, Pearson's correlation coefficient of T1% and T2% with the BA and the P3amp was computed. Because  $K_b$  and  $\alpha$  violated the assumption of normality, they were correlated with BA and P3amp by means of the non-parametric Spearman correlation test. For the parameters whose correlation was statistically significant we performed two regression analyses in which attentional parameter (T1%) was considered as the independent variable and the BA and the P300amp variables were considered as dependent variables.

## 4 Results

A significant positive correlation was observed between T1% and the offline BA,  $r=.79$ ,  $p<0.05$ . To estimate the predictive value of T1% on the binary accuracy we computed a regression analysis which resulted in an  $F=8.341$  with a  $p<0.05$ , indicating that the variance of the binary performance was predictable by the participant temporal filtering capacity, with  $\beta=0.79$ . A significant positive correlation was found between T1% and P300amp ( $r=.84.5$ ,  $p<0.05$ ) showing that participants with higher T1% had a larger P300amp. As a result of the linear regression, T1 accuracy was significantly predictive of P300amp ( $F=16.23$  with a  $p<0.05$ ) with  $\beta=0.87$ . Offline mean value of BA obtained during the BCI task was 87.4% (SD = 2.4%, range = 84.5–92.3 %). The mean amplitude for P300 was 3.3  $\mu V$  (SD = 1.6, range = 1.1–6.5  $\mu V$ ). Mean accuracy of detection was 77.2% (SD = 10.4%, range = 65–96.25%) and 67.7% (SD= 14.1%, range = 50.3–87.1%) for T1 and T2, respectively. No significant correlation was found between T2%,  $K_b$ ,  $\alpha$  the offline binary performance and P300amp.

## 5 Discussion and conclusion

The detection rate of T1 in the RSVP task can be interpreted as an index of selective attention: it represents indeed the capacity to detect a target within a stream of stimuli, to create a memory trace and to retain it. We demonstrated that such capability influences the performances in the BCI task.

As the accuracy of detecting T2 is an index of the speed of attentive update, the missed correlation with the considered BCI variables leads us to speculate that the capacity of dynamically updating the attentive filter is less likely to be a cognitive substrate supporting the BCI control.

The lack of relationship between BCI parameters and the variables measured with the change detection task did not confirm the hypothesis that the memory capacity and the attentional spatial filtering capacity were associated with the capability of the participants to control the P300-speller. We can speculate that the allocation of attention resources on the selected item during the BCI task might not be based on spatial (being the location of the target letter static) or feature characteristics but on symbolic aspects (e.g. semantic aspects of the target letter).

The data reported in the present paper partly clarify the cognitive substrate related to BCI control in people with ALS. This issue could allow future speculations on the factors underlying BCI control

failure observed in potential user groups. The awareness about the processes and the clinical features of BCI potential end-users influencing the BCI performance, would allow developing flexible systems, adaptable to different clinical profiles. We can conclude that top-down mechanism to keep an attentive map active (Huang & Pashler 2007) is crucial to control a BCI speller task, allowing the user to set up and maintain the proper attentional map throughout a trial and thus to select the desired letter.

## References

- Farwell, L.A. & Donchin, E., 1988. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and Clinical Neurophysiology*, 70(6), pagg.510–523.
- Huang, L. & Pashler, H., 2007. A Boolean map theory of visual attention. *Psychological review*, 114(3), pagg.599–631.
- Krancioch, C. et al., 2007. Temporal dynamics of access to consciousness in the attentional blink. *NeuroImage*, 37(3), pagg.947–955.
- Riccio, A. et al., 2013. Attention and P300-based BCI performance in people with amyotrophic lateral sclerosis. *Frontiers in Human Neuroscience*, 7:, pag.732.
- Ringholz, G.M. et al., 2005. Prevalence and patterns of cognitive impairment in sporadic ALS. *Neurology*, 65(4), pagg.586–590.
- Vogel, E.K., McCollough, A.W. & Machizawa, M.G., 2005. Neural measures reveal individual differences in controlling access to working memory. *Nature*, 438(7067), pagg.500–503.
- Wolpaw, J. & Wolpaw, E.W. a c di., 2012. *Brain-Computer Interfaces: Principles and Practice* 1° ed., Oxford University Press, USA.

# Towards Implementation of Motor Imagery using Brain Connectivity Features

Huaijian Zhang, Ricardo Chavarriaga and José del R. Millán

CNBI, Ecole Polytechnique Fédérale de Lausanne, Switzerland  
huaijian.zhang@epfl.ch

## Abstract

This study aims to explore modulation of the connectivity pattern when people perform left hand versus right hand motor imagery and probe the feasibility of adopting connectivity information to discriminate these tasks. Nine subjects were recorded with 16-channel EEG system, covering sensorimotor cortex. Non-normalized directed transfer function (DTF) is used to obtain the brain connectivity between EEG electrodes. The results demonstrate that the modulations of intrahemispheric and interhemispheric information flows are not identical during left and right hand motor imageries. Particularly, the mu rhythm is highly modulated in intrahemispheric brain interactions, whereas the high frequency bands are more related with distant interhemispheric brain interactions. Furthermore, classification results suggest that the DTF features bring additional informative features for the classification between two tasks.

## 1 Introduction

Brain-computer interface (BCI) has been developed as a possible solution for enabling communication capabilities to people who lose motor functions [9]. The use of motor imagery (MI) is a common approach for BCI systems to send mental command by imagining the movement of limbs [9]. Currently, most motor imagery BCI systems adopt power spectral density (PSD) or common spatial filter as discriminative features between tasks [7]. Recently, some studies have explored the possibility of using the brain interaction patterns between EEG channels as classification features, e.g. phase difference or directed transfer function (DTF) [8] [2]. The objective of the present work is to study the use of DTF features for BCI, as well as the modulations of brain connectivity during MI. We compare the performance of features extracted by DTF and the PSD features, as well as their combination for decoding BCI commands. Moreover, we assess the modulations of brain connectivity to evaluate the interhemispheric and intrahemispheric interaction patterns.

## 2 Methods

### 2.1 Experimental protocol

In this study, we focus on a two-class motor imagery task, left hand versus right hand [6]. Subjects were instructed to move a horizontal bar by imaging the movement of their hands. We analyzed offline experiments, involving nine subjects that perform two runs of BCI training. Each run comprised 30 trials of motor imagery for each class. Before each trial, a visual cue was presented to show the target, and the bar started moving after one second, as shown in Figure 1.A. During each trial a feedback bar that moved to the target continuously for 4 seconds was shown to the subject. EEG signals were recorded using a 16-channel g.USBamp amplifier (g.tec medical engineering, Schiedelberg) with a sampling rate of 512 Hz. The signal was band-pass

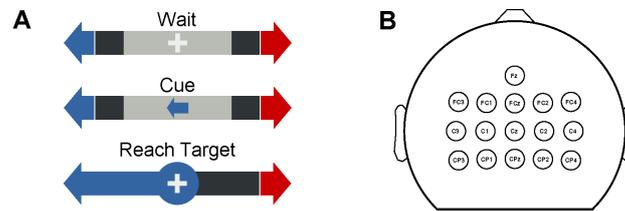


Figure 1: Experimental protocol of motor imagery and the montage of the EEG cap. Each trial includes fixation (1000 ms), cue (1000 ms) before the bar starts moving.

filtered between 0.1 Hz and 100 Hz and notch filtered at 50 Hz. The 16 EEG channels were placed over the sensorimotor cortex as shown in Figure 1.B.

## 2.2 Signal processing and classification

To compare the performance between PSD and DTF, features were computed 16 times per second. The PSD of each channel was computed using the Welch method for the past one second, with the window (Hanning) size 500 ms and overlapping 50%. PSD features were collected between 2-50 Hz with the step size 2 Hz. DTF was also computed in the same period of one second, by estimating the coefficients of multivariate autoregressive (MVAR) model based on Yule-Walker equation, which included all 16 EEG channels with order 8. The system transfer matrix in frequency domain was obtained by applying fast Fourier transform on the MVAR coefficients, as the so-called non-normalized directed transfer function [5]. The step size of DTF was 2 Hz, resulting in a three dimensional matrix  $16 \times 16 \times 25$  (channel  $\times$  channel  $\times$  frequency) for each time step of  $1/16$  s. The log values of either PSD or DTFS were then used for classification.

We compare the performance with three types of features: (1) PSD features; (2) DTF features; (3) Combination of both PSD and DTF. We computed the Fisher score on the offline runs to select the most informative features, defined by  $fs = |m_1 - m_2| / (s_1^2 + s_2^2)$ , where  $m_k$  and  $s_k^2$  represent the mean and variance of class  $k$ . In the combined case we used the features that are selected in PSD and DTF separately. Linear discriminant analysis (LDA) was used for classification (left vs right MI), assuming Gaussian distribution of the data samples. The classification performance was evaluated by 10-fold cross validation and the area under the curve (AUC) of the testing sets.

## 3 Results

### 3.1 Brain connectivity patterns

We compute the interaction *between* and *within* hemispheres to explore the modulation during motor imagery. The left hemisphere is defined as the channels FC3, FC1, C3, C1, CP3 and CP1, and the right hemisphere corresponds to FC2, FC4, C2, C4, CP2 and CP4. The brain interaction from left to right hemisphere is defined as the mean DTF value from the channels in the left hemisphere to the channels in the right, and vice versa for right to left interactions.

Figure 2.A shows the difference of DTF between left and right hand motor imagery ( $DTF_{left} - DTF_{right}$ ) for four brain interaction patterns: Interhemispheric (left to right and right to left) and intra hemispheric (left and right) connectivity. The curves illustrate the mean

value across all subjects. As shown in the figure, the connectivity levels within hemispheres are highly modulated in the mu rhythm (blue and red curves), around 10-14 Hz, indicating evident disassociation between the two motor imagery tasks. In particular, the connectivity within left hemisphere is much higher in the condition of left hand motor imagery, cf., the positive peak in red curve, and vice versa (negative peak in the blue curve). This might be due to the higher mu rhythm desynchronization in contralateral than in ipsilateral areas, which causes the lower interaction level in the left hemisphere when people are perform right hand motor imagery. The modulation of intrahemispheric connectivity disappear in the high frequency bands, i.e. above 35 Hz.

The green and black curves represent the interhemispheric connectivity patterns, which are almost zero in the low frequency bands, indicating no difference between two motor imagery tasks. Differences are noticeable in high frequency bands, particularly 26-32 Hz, in which the information flow from left to right hemisphere is higher when subjects are performing left hand motor imagery (green curve), and vice versa (black curve).

Figure 2.B and C represent the mean values of the curve in Figure 2.A in two selected frequency bands, 10-14Hz and 26-32 Hz. Significant differences (Wilcoxon signed-rank test) could be found between within left and within right hemisphere interactions in the mu band. The modulation in higher frequency bands is not as obvious as the low band, however. It is also significant between two interhemispheric interactions. These results indicate that the interhemispheric, or distant, brain interactions are more modulated in high frequency bands than local connections, and motor related brain signals are originated from the contralateral brain regions.

### 3.2 Classification performance

10-fold cross validation was used to evaluate the offline training performance. We tested the effect of feature quantity by varying the number of selected features, from 1 to 50. Results of each subject were averaged for all testing folds. Rapid increase of AUC can be observed when the feature number is very low, i.e. below 10 features for all the feature sets, after which the increasing trend stops and keep constant with more features, shown in Figure 2.D. This is caused by the level of redundancy between features, i.e. no more novel information is contributed by new features. In average, the PSD features show higher AUC (0.58) than the DTF method (0.57). However, the combined feature set (BOTH) yields the highest performance, which is around 0.6 after using more than 12 features. Wilcoxon signed-rank test (uncorrected) was applied and found that below 20 features, the performance of combination is significantly better ( $p < 0.05$ ) than PSD, except using 8 features, Figure 2.D.

## 4 Discussion & Future works

The present study verified the potential application of using DTF to decode motor imagery tasks, as well as the the modulation patterns of brain connectivity between the two tasks. Results of offline analysis of 9 subjects, show modulation of EEG rhythms that are consistent with previously reported results. In particular activity in higher frequency bands related to the interhemispheric interactions [4], as well as the reduced mu band in intrahemispheric patterns with contralateral sites. Further experiments will be performed to confirm the current results including online evaluation.

One should notice that an essential issue of the online DTF method is the computational consumption, since heavy computation is required to obtain the coefficient matrix of the MVAR

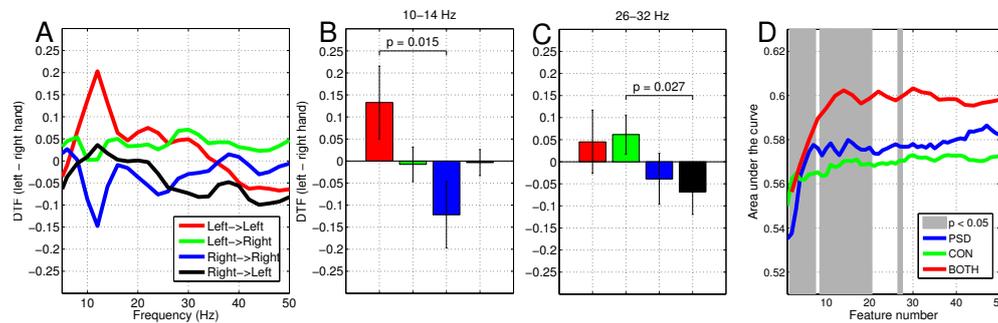


Figure 2: A: The modulation of DTF during left and right hand motor imagery. B and C: Mean and variance of brain interaction patterns in two frequency ranges across subjects (B: 10-14 Hz; C: 26-32 Hz). D: Offline performance of classifying two motor imagery tasks with three types of features. Gray area show the results of statistical tests between PSD and BOTH.

model [3], for our case  $16 \times 16 \times 8$  (8 is the order of MVAR model, determined by AIC criteria) parameters are estimated with 1 s EEG data, and the computation should be shorter than 1/16 second, since we update the DTF feature with 16 Hz to obtain BCI control. The use of higher order models or more EEG channels will increase the time consumption. In our experiment, it took around 0.01 s to do the computation of DTF and PSD, short than 1/16 s. Furthermore, future work may try to improve the stability of DTF features by averaging MVAR coefficients in a sliding window within past 1 s EEG signal with certain overlap, as the Hanning window in Welch's method. For online implementation, the other option could be applying dynamic updating of MVAR coefficients other than re-compute for each 1 s, as adaptive MVAR model, which might be helpful to decrease the computation [1].

## References

- [1] M. Arnold, X.H.R. Milner, H. Witte, R. Bauer, and C. Braun. Adaptive AR modeling of nonstationary time series by means of Kalman filtering. *IEEE Trans Biomed Eng*, 45(5):553–562, 1998.
- [2] M. Billinger, C. Brunner, and G.R. Müller-Putz. Single-trial connectivity estimation for classification of motor imagery data. *J Neural Eng*, 10(4):046006, 2013.
- [3] I. Daly, S.J. Nasuto, and K. Warwick. Brain computer interface control via functional connectivity dynamics. *Pattern Recogn*, 45(6):2123 – 2136, 2012.
- [4] M. Grosse-Wentrup. Understanding brain connectivity patterns during motor imagery for brain-computer interfacing. In *NIPS*, pages 561–568, 2008.
- [5] M.J. Kaminski and K.J. Blinowska. A new method of the description of the information flow in the brain structures. *Biol Cybern*, 65(3):203–210, 1991.
- [6] R. Leeb, S. Perdikis, L. Tonin, A. Biasucci, M. Tavella, M. Creatura, A. Molina, A. Al-Khodairy, T. Carlson, and J.d.R. Millán. Transferring brain-computer interfaces beyond the laboratory: Successful application control for motor-disabled users. *Artif Intell Med*, 59(2):121 – 132, 2013.
- [7] H. Ramoser, J. Müller-Gerking, and G. Pfurtscheller. Optimal spatial filtering of single trial EEG during imagined hand movement. *IEEE Trans Rehabil Eng*, 8(4):441–446, 2000.
- [8] Y. Wang, B. Hong, X. Gao, and S. Gao. Phase synchrony measurement in motor cortex for classifying single-trial EEG during motor imagery. *EMBC*, 1:75–78, 2006.
- [9] J.R. Wolpaw, N. Birbaumer, D.J. McFarland, G. Pfurtscheller, and T.M. Vaughan. Brain-computer interfaces for communication and control. *Clin Neurophysiol*, 113(6):767–791, 2002.

# Multi-way Decoding of Wirelessly Transmitted ECoG Signals from WIMAGINE<sup>®</sup> Implant for Self-Paced Brain Computer Interface

Tetiana Aksenova<sup>1\*</sup>, Andrey Eliseyev<sup>1</sup>, Louis Korczowski<sup>1</sup>,  
Thomas Costecalde<sup>1</sup>, Michael Foerster<sup>1</sup>, Nana Arizumi<sup>1</sup>, Christiane Poeschl<sup>1</sup>,  
Nicolas Tarrin<sup>1</sup>, Napoleon Torres-Martinez<sup>1</sup>, Corinne Mestais<sup>1</sup> and Alim-  
Louis Benabid<sup>1</sup>

<sup>1</sup>CEA, LETI, CLIMATEC; Grenoble, France

tetiana.aksenova@cea.fr, eliseyev.andrey@gmail.com,  
louis.korczowski@gmail.com, thomas.costecalde@cea.fr,  
michael.foerster@cea.fr, arizumi@gmail.com,  
Christiane.Poeschl@gmx.at, napoleon.torres-martinez@cea.fr,  
nicolas.tarrin@cea.fr, corinne.mestais@cea.fr, alimlouis@sfr.fr

## Abstract

The multi-way decoding algorithms are adapted to incomplete wirelessly transmitted data and integrated to CLIMATEC<sup>®</sup> BCI platform. The platform includes wireless 64-channels ElectroCorticoGram (ECoG) recording implant WIMAGINE<sup>®</sup> and BCI software environment associated to a 4-limbs exoskeleton EMY.

## 1 Introduction

Multi-way (tensor-based) analysis was recently reported as an effective tool for neuronal signal processing. It was applied for Brain Computer Interface (BCI), e.g., (Li & Zhang, 2010). Movement-related BCI aims to provide an alternative non-muscular communication pathway for individuals with severe motor disability to send commands to the external world using measurements of brain activity. Generally, a common approach for brain signal analysis consists in the extraction of the information from spatial, frequency and/or temporal domains. The commonly used time-frequency decomposition of the signals leads to a matrix valued process. Introducing delays in time leads to the observations stored in a 4th order tensor. For the decoding model identification, a projection of the high dimensional tensor of observation to low dimensional space is generally applied using unfolding procedure or various tensor decomposition technics. The multi-way decoding was chosen for BCI project in CLIMATEC<sup>®</sup>, CEA, Grenoble (Eliseyev, et al., 2011; Eliseyev & Aksenova, 2013). The goal of the project is to allow a tetraplegic subject to control external effectors, such as an exoskeleton. Despite encouraging success, movement related BCIs have yielded limited application outside of the laboratory, mainly due to unsolved problems of efficiency and long-term stability. The key criterion for clinical applications is reliable BCI devices and their stable and robust functioning. Moreover, the self-paced BCI system should be activated by users whenever they want. The important requirement is fast and easy calibration of the BCI system. In addition, to be comfortable for users, the high decision rate and low latency are desirable. The latency comprises algorithm delay, as well as

---

\* Corresponding author

data transmission and computation time. To approach the requirements of the clinical applications, a set of tensor-based algorithms were developed in CLINATEC<sup>®</sup>, CEA. The Recursive block-wise N-way PLS (Eliseyev & Aksenova, 2013) provides the progressive adaptive learning of the BCI system. L1-penalized N-way PLS algorithm performs the slice-oriented tensor decomposition that allows the selection of the groups of informative features (Eliseyev, et al., 2012). Sparse models reduce the computation time, increase the decision rate and, thus, improve the latency of the BCI system. The algorithms were tested in self-paced mode in series of preclinical experiments in animals, namely, the brain switch in freely moving rats (Eliseyev, et al., 2011), as well as brain switch and upper limb trajectory reconstruction in minimally restricted nonhuman primates (Eliseyev, et al., 2012; Eliseyev & Aksenova, 2013). In parallel, to estimate the performance, the publically evaluable data (Shimoda, Nagasaka, Chao, & Fujii, 2012) were analyzed (Eliseyev & Aksenova, 2013). The algorithms' delays were evaluated in series of experiments in rats and nonhuman primates. The negative delay of the decoder was observed in the brain-switch BCIs (Eliseyev, et al., 2011; Eliseyev, et al., 2012). The next step is integration and testing of the decoding algorithm at clinical BCI platform.

The CLINATEC<sup>®</sup> BCI platform includes ECoG recording implants WIMAGINE<sup>®</sup> which wirelessly transfer neural activity of the brain to a PC and the software environment associated to an effector. The wireless connection may introduce temporal loss of the signals. Moreover, the important point of the real time functioning of BCI system is the delay of the signal processing: signal acquisition, transmission, and the commands generation. To ensure stable system functioning, a set of gaps' filling algorithms was investigated. Then, preclinical experiments were carried out to study the latency and the performance of the system in a realistic context.

## 2 Methods

### 2.1 CLINATEC<sup>®</sup> BCI Platform

For real-life applications, the decoder is to be integrated to BCI environment. The CLINATEC<sup>®</sup> clinical movement related BCI platform is based on a wireless 64-channels ECoG recording implant WIMAGINE<sup>®</sup>, designed for the long-term clinical application (Charvet, et al., 2013). Two implants will record simultaneously the neural activity of the brain for wireless transfer to the base station, then a PC. The BCI software environment is associated to a 4-limbs exoskeleton EMY (Enhancing Mobility) dedicated to medical purposes (Perrot, Verney, Morinière, & Garrec, 2013).

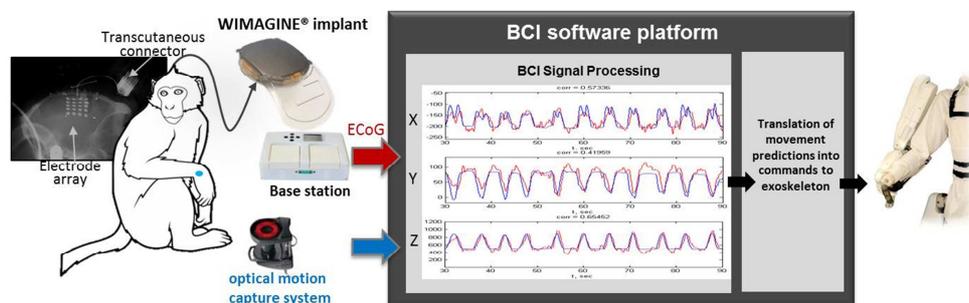
### 2.2 Integration of Decoder to BCI Platform

As the physical conditions are changing (quality of the wireless link, perturbations, interference) the bitrate of a radio frequency transmission inevitably varies. Thus, the wireless connection could introduce temporal loss of the data. To define the best strategy of decoding in the case of incomplete data, a set of computational experiments were carried out. The goal of the experiments was a comparison of different strategies of signal recovering and an evaluation of decoder robustness. For this study, the problem of reconstruction of hand trajectory was considered. The publically available ECoG recordings (Shimoda, Nagasaka, Chao, & Fujii, 2012) were corrupted with artificial sequences of the gaps. They were generated according to the gaps' distribution in the wirelessly transmitted signals in CLINATEC<sup>®</sup> BCI Platform. The sequences were imposed on the gaps-free data. The medians of the gaps segment were about 20 ms and 30% of the gaps do not exceed 5 ms. The number of the gaps was increased 100 times to obtain an upper estimate of the prediction biases. A set of algorithms for gaps' filling was studied: Zero-Order Hold, First-Order Hold, Autoregressive Model (8<sup>th</sup> Order), Spline Cubic Interpolation, Piecewise Polynomial Interpolation, Sinusoidal Amplitude L1 Estimation, and Sinusoidal Amplitude and Phase L1 Estimation (Cowpertwait & Metcalfe, 2009). In

all the cases, the algorithm has demonstrated significant robustness. The biases of the prediction (relative root mean squares error) were, 1.4%, 1.4%, 1.4%, 3.6%, 1.4%, 1.8%, 1.3%, and 1.3% respectively. Thus, even in the simplest case of the filling, namely, Zero-Order Hold, the influence of the data loss on prediction was negligible.

### 2.3 Preclinical Test

Integrated to the BCI platform, the decoder was tested in real-time in preclinical experiments carried out in non-human primates in CLINATEC<sup>®</sup>. Ethical approval was obtained from ComEth in accordance with the European Communities Council Directive of 1986 (86/609/EEC). Designed to be compatible with human's skull, WIMAGINE<sup>®</sup> cannot be implanted directly to primates. Thus, a silicone/platinum-iridium cortical electrode array was implanted in the region of monkey's left motor cortex and connected to the recording electrodes of the WIMAGINE<sup>®</sup> implant by means of a connector. Then, it was integrated to the BCI platform. The setup of the experiment is shown in Figure 1. The monkey was trained to reach an exposed target using the right hand. The hand movements were recorded by an optical motion capture system Vicon (Motion Systems, Oxford, UK). During the calibration stage, the monkey's ECoG data were used together with information about the hand position to identify a prediction model. For the BCI system, the calibration training tensor was formed using 1-second epochs of training recording. Normalized absolute values of complex wavelet coefficients (Morley) in logarithmic scale were considered as features. Feature tensor (6000 epochs) is split into training tensor to identify coefficients (4000 epochs) and validation tensor (2000 epochs). L1-penalised tensor decomposition was applied for features selection (frequencies and channels). Sparse model provided reducing of computation time with slight improvement of decoding performance (about 6% of improvement for the validation tensor). The correlation of the observed and the decoded trajectories was varying in the range of 0.4-0.8.



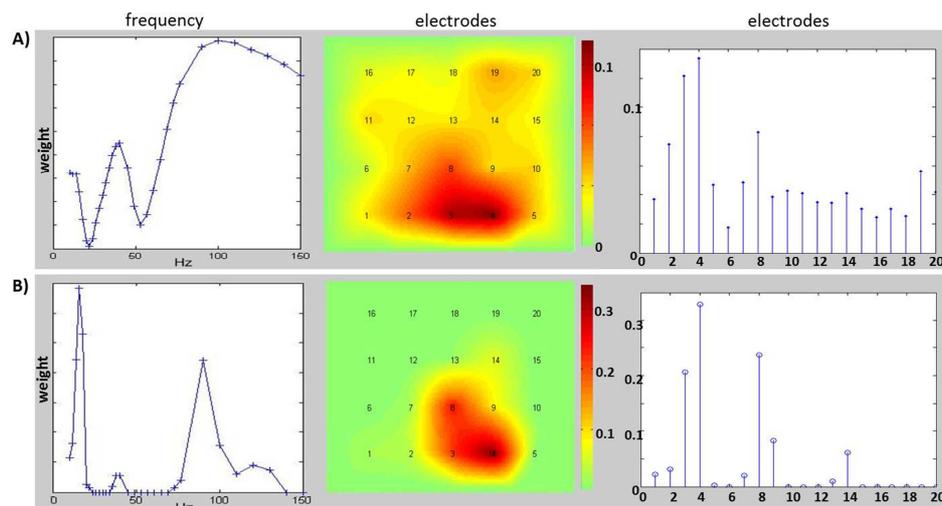
**Figure 1:** Preclinical experiments using CLINATEC<sup>®</sup> BCI Platform. Recording of ECoG activity of the brain is transmitted wirelessly by the WIMAGINE<sup>®</sup> implant. ECoG data are analyzed in real-time to send the commands to the external effector.

Figure 2 illustrates the relative weights of the linear model coefficients in the spatial and frequency domains for both full and sparse models. In the frequency domain, high frequencies 80-130 Hz have the highest contribution. At the same time, the frequencies 35-40 Hz as well as around 7 Hz are informative.

During the online experiments, the control commands were generated with decision rate DR=10 Hz. The estimated delay of the signal processing, including the signal acquisition, transmission, as well as the commands generation, was approximately equal to 300 ms. The next step of the CLINATEC<sup>®</sup> project is optimization and adjusting of all the components of the BCI system. Additional tests in animals are in progress to better estimate system performance (more animals, various experiment protocols etc.).

### 3 Acknowledgment

Research is supported by French National Research Agency, Fondation Motrice, Fondation Nanosciences, Fondation de l'Avenir, Région Rhône-Alpes, Philanthropic Foundation Edmond J. Safra.



**Figure 2:** Relative weights of the linear model coefficients in the spatial and frequency domains for both full (A) and sparse (B) models.

### References

- Charvet, G., Sauter-Starace, F., Foerster, M., Ratel, D., Chabrol, C., Porcherot, J., et al. (2013). WIMAGINE®: 64-channel ECoG recording implant for human applications. *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*, (pp. 2756-2759).
- Cowperrwait, P. S., & Metcalfe, A. V. (2009). *Introductory time series with R*. Springer.
- Eliseyev, A., & Aksenova, T. (2013). Recursive N-way partial least squares for brain-computer interface. *PLoS One*, 8(7), e69962.
- Eliseyev, A., Moro, C., Costecalde, T., Torres, N., Gharbi, S., Mestais, C., et al. (2011). Iterative N-way partial least squares for a binary self-paced brain-computer interface in freely moving animals. *J Neural Eng*, 8(4), 046012.
- Eliseyev, A., Moro, C., Faber, J., Wyss, A., Torres, N., Mestais, C., et al. (2012). L1-penalized N-way PLS for subset of electrodes selection in BCI experiments. *J Neural Eng*, 9(4), 045010.
- Li, J., & Zhang, L. (2010). Regularized tensor discriminant analysis for single trial EEG classification in BCI. *Pattern Recognition Letters*, 31(7), 619-628.
- Perrot, Y., Verney, A., Morinière, B., & Garrec, P. (2013). EMY: Full-body Exoskeleton. *ACM SIGGRAPH Emerging Technologies*. Anaheim, US.
- Shimoda, K., Nagasaka, Y., Chao, Z., & Fujii, N. (2012). Decoding continuous three-dimensional hand trajectories from epidural electrocorticographic signals in Japanese macaques. *J Neural Eng*, 9(3), 036015.

# Detecting Brain Network Changes Induced by a Neurofeedback-based Training for Memory Function Rehabilitation After Stroke

J. Toppi<sup>1,2</sup>, L. Astolfi<sup>1,2</sup>, M. Riseti<sup>1</sup>, S. E. Kober<sup>3</sup>, A. Anzolin<sup>1</sup>, F. Cincotti<sup>1,2</sup>, G. Wood<sup>3</sup> and D. Mattia<sup>1</sup>

<sup>1</sup>IRCCS Fondazione Santa Lucia, Rome, Italy. <sup>2</sup>Dept. of Computer, Control, and Management Engineering, Univ. of Rome "Sapienza", Italy. <sup>3</sup>Dept. of Psychology, Univ. of Graz, Austria.  
jlenia.toppi@uniroma1.it

## Abstract

The efficacy of rehabilitative interventions in stroke patients is routinely assessed by means of a neuropsychological test battery. Nowadays, more evidences indicate that the neuroplasticity which occurs after stroke can be better understood by investigating changes in brain networks. In this pilot study we applied advanced methodologies for effective connectivity estimation in combination with graph theory approach, to define EEG derived descriptors of brain networks underlying memory tasks. In particular, we proposed such descriptors to identify substrates of efficacy of a Brain-Computer Interface (BCI) controlled neurofeedback-based intervention to improve cognitive function after stroke. EEG data were collected from two stroke patients before and after a neurofeedback-based training for working memory deficits. We show that the estimated brain connectivity indices were sensitive to different training intervention outcomes, thus suggesting an effective support to the neuropsychological assessment in the evaluation of the changes induced by the BCI-based rehabilitative intervention.

## 1 Introduction

In the post-stroke subacute time frame, the estimated proportion of patients having cognitive impairment ranges from below 50% to over 90%. Currently, the diagnosis of cognitive impairments after stroke and their treatment efficacy relies upon a neuropsychological assessment battery. Nowadays, evidences from neuroimaging studies indicate that the neuroplasticity which occurs after stroke might be better understood by investigating changes in brain networks (Cramer et al., 2011).

In this study we proposed the application of advanced methodologies for effective connectivity estimation (Milde et al., 2010), combined with a graph theoretical approach (Astolfi et al., 2013) to extract indices describing the topology of the brain networks as derived from EEG signals. The aim was to seek for neurophysiological descriptors which could serve as a sensitive outcome measures and thus, support the neuropsychological assessment in evaluating the efficacy of a Brain-Computer Interface controlled neurofeedback training to promote recovery of memory function after stroke. As such the BCI-controlled neurofeedback training module has been implemented within the EU funded project CONTRAST ([www.contrast-project.eu](http://www.contrast-project.eu)) which is currently deploying a Brain Neural Computer Interface based technology to provide cognitive training modules to improve cognitive rehabilitation outcomes in institutionalized and at home stroke patients.

## 2 Material and Methods

### 2.1 Experimental Design

Two representative stroke patients (Patient A, female, 70 years old; right hemisphere stroke lesion and Patient B, male, 20 years old, left hemisphere stroke) were selected among those currently enrolled in a neurofeedback-based intervention protocol implemented in BCI close loop to target post-stroke memory disorders. According to the pilot study purpose, the selection of 2 “representative” patients was solely based on their response to the training experimental intervention. The adopted protocol consisted of 10 training sessions in which the patients were instructed to voluntarily increase their sensorymotor rhythm (SMRs; 12-15 Hz; Kober et al., 2014) amplitude over an established threshold. Each time the SMR amplitude exceeded the threshold for  $\geq 250$  ms, the participant was rewarded by gaining points. The threshold was automatically adapted after each run on the basis of all previous runs. Cz was used as feedback electrode; each training session lasted 25' (3 min baseline; 6 feedback runs, 3-min each). The acquisition and the real time feedback were implemented in Biotrace software for Nexus10 (Mind Media).

Before (PRE) and after (POST) the neurofeedback based intervention, EEG scalp signals were recorded (64 channels; Brain products, 200Hz sampling frequency) while patients were performing the Sternberg memory task. The Sternberg paradigm task consists of 3 phases Encoding, Storage and Retrieval during which a series of numbers have to be memorized, retained and retrieved within a short time interval (Sternberg, 1966). Patient's declarative memory and the visuo-spatial short-term memory deficits were assessed before and after the training by means of the Rey Auditory Verbal Learning Test (RAVLT) and the Corsi Block Tapping Test (CBTT), respectively.

### 2.2 Effective Connectivity and Graph Theory

#### *Partial Directed Coherence*

The Partial Directed Coherence (PDC) is a full multivariate spectral measure used to determine the directed influences between any given pair of signals in a multivariate data set (Baccalá and Sameshima, 2001). In this work we applied an adaptive formulation of PDC based on a time varying multivariate autoregressive (MVAR) model, whose coefficients are estimated by means of Kalman filter (Milde et al., 2010).

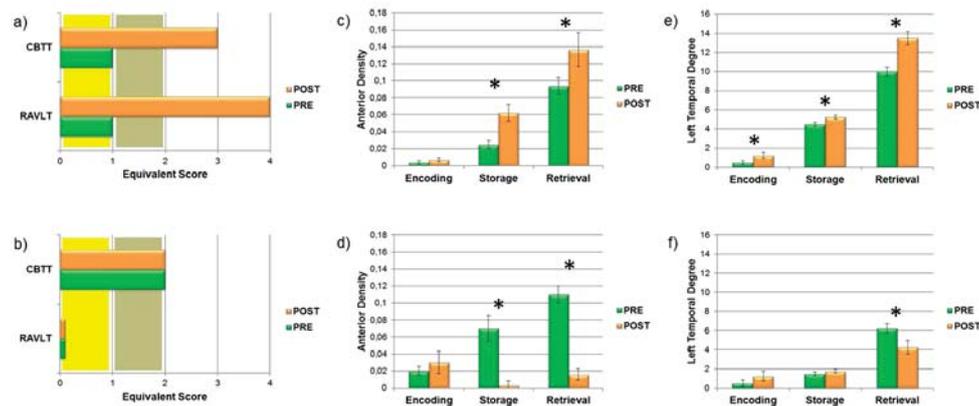
#### *Graph Theory Approach*

A graph is a mathematical object consisting in a set of nodes linked by connections which represent the existence of interactions between the nodes. The structure of a graph is described by means of an adjacency matrix  $G$  whose entries are  $G_{ij} = 1$  if the link exists, otherwise  $G_{ij} = 0$ . We considered two indices to describe the properties of the EEG derived patterns of effective connectivity: i) Degree, consisting in the number of links connected directly to a node and ii) Anterior Density defined as the number of connections exchanged between the electrodes located in the anterior part of the scalp. The sensitivity/specificity of these 2 indices in accounting for brain network characteristics during memory tasks was already validated in a previous study on healthy age-matched volunteers (Astolfi et al.2013).

#### *Connectivity Analysis*

PRE and POST EEG data were pre-processed (1-45Hz band-pass filter) and analysed at single subject level. Time-varying effective connectivity networks (Milde et al., 2010) were estimated for each memory phase and frequency band (defined according to Individual Alpha frequency) and the corresponding salient properties were derived by means of graph theory approach (Sporns et al., 2004). PRE and POST comparisons were performed to capture, at single subject level, the differences

in connectivity networks and in graph indexes associated to them. In particular, an independent sample t-test was applied for a significance level of 5% corrected by means of False Discovery Rate for preventing type I errors.



**Figure 1:** a-b) Bar diagrams reporting the equivalent scores achieved for RAVLT and CBTT tests administered to patient A (panel a) and B (panel b) before (PRE, green bars) and after (POST, orange bars) the SMR training. Equivalent scores below 2 (in yellow) highlight a pathological condition. c-f) Anterior Density and Left Temporal Degree indexes estimated in alpha band in PRE (green bars) and POST (orange bars) sessions relative to the two stroke patients A (panels c and e) and B (panels d and f). The symbol (\*) denotes a significance difference between PRE and POST session scores (unpaired t-test;  $p < 0.05$ ).

### 3 Results

#### 3.1 Patient A

The Patient A was able to learn the modulation of her SMR as indicated by the increase of SMR amplitude from  $7.7 \mu V^2$  to  $8.4 \mu V^2$  across the 10 training sessions.

**Memory Assessment.** As reported in Fig.1a, the neuropsychological tests revealed a significant improvement of the tested memory function after the neurofeedback-based training. Equivalent scores for both CBTT and RAVLT tests increased from 1 to 3 and 4 respectively, thus indicating a transition from a pathological (PRE) to a physiological (POST) condition.

**Behavioral Data.** Analysis of the behavioral performance obtained at the Sternberg task revealed an increase of correct answers and a significant decrease ( $df=68$ ,  $t=2.16$ ,  $p=0.034$ ) of the reaction time after training (PRE-POST comparison, unpaired t-test).

**EEG derived Brain Network.** Analysis of the connectivity patterns revealed a significant POST training increase of Anterior Density index (Fig.1c) estimated in the alpha band only for Storage ( $df=198$ ,  $t=2.87$ ,  $p=0.0045$ ) and Retrieval ( $df=198$ ,  $t=3.97$ ,  $p=0.0001$ ) phases of the Sternberg task associated with an increase of Left Temporal Degree index (Fig.1e) in alpha band for all the three memory phases (Encoding ( $df=198$ ,  $t=1.99$ ,  $p=0.048$ ), Storage ( $df=198$ ,  $t=2.08$ ,  $p=0.039$ ) and Retrieval ( $df=198$ ,  $t=3.05$ ,  $p=0.0026$ )).

#### 3.2 Patient B

The Patient B did not show changes in the amplitude of his SMR across the 10 training sessions. The amplitude value remained stable around  $2.3 \mu V^2$ .

**Memory Assessment.** In this Patient (Fig.1b) we did not find significant changes in the memory functions as evaluated by means of neuropsychological assessment. Equivalent scores for both RAVLT and CBTT tests remained around 1 and 2 respectively, indicating a persistency of the pathological profile.

**Behavioral Data.** Similar negative outcome was found for the behavioral assessment. Data analysis revealed a decrease of the percentage of correct answers and no significant difference ( $df=68$ ,  $t=0.76$ ,  $p=0.46$ ) in reaction time between PRE and POST sessions of Sternberg task (unpaired t-test).

**EEG derived Brain Network.** Connectivity pattern analysis revealed in Patient B an opposite profile of changes in the POST training analysis with respect to what observed in Patient A. In fact, a significant decrease in the Anterior Density index (Fig.1d) for Storage ( $df=198$ ,  $t=3.07$ ,  $p=0.0024$ ) and Retrieval ( $df=198$ ,  $t=2.11$ ,  $p=0.036$ ) phases and of Left Temporal Degree index (Fig.1f) for the Retrieval ( $df=198$ ,  $t=2.01$ ,  $p=0.046$ ) memory phase both estimated in the alpha band, were observed.

## 4 Discussion and Conclusion

The preliminary findings obtained in this pilot study indicated that the estimated brain connectivity indices were sensitive to different response (outcome) to the BCI-controlled neurofeedback training. In particular, for both *representative* patients the changes observed in the Anterior Density index and the Left Temporal Degree indices between PRE and POST training assessment were in agreement with Sternberg behavioral changes (Sternberg, 1966; Astolfi et al., 2013) and even more interesting, with the outcome of neuropsychological tests of memory function.

Validation in a proper sampled study is currently in progress within the EU CONTRAST project. If these preliminary findings will be confirmed, estimation of the brain networks derived from a non-invasive, cost/effective EEG technique would provide a solid evidence for clinical application of such methodology to empower a quantitative outcome measurement of novel post-stroke rehabilitation strategy aiming at promoting cognitive functional improvement after stroke.

## 5 Acknowledgment

Research supported by the European ICT Program FP7-ICT-2009-4 Grant Agreement 287320 CONTRAST.

## References

- Astolfi, L., Toppi, J., Wood, G., Kober, S., Riseti et al., (2013). *Advanced methods for time-varying effective connectivity estimation in memory processes*. Conf. IEEE Eng. Med. Biol., 2936–2939.
- Baccalá, L. A., and Sameshima K., (2001). *Partial Directed Coherence: A New Concept in Neural Structure Determination*. Biological Cybernetics 84: 463–74.
- Cramer, S.C., Sur, M., Dobkin, B.H., O'Brien, C., Sanger, T.D., Trojanowski, J.Q., Rumsey, et al., (2011). *Harnessing neuroplasticity for clinical applications*. Brain. 134, 1591–1609.
- Kober, S. E., Witte M., Stangl M., Våljamäe A., Neuper C., and Wood G. (2014). *Shutting down Sensorimotor Interference Unblocks the Networks for Stimulus Processing: An SMR Neurofeedback Training Study.* Clinical Neurophysiology, S1388-2457(14)00186-2.
- Milde, T., Leistriz, L., Astolfi, L., Miltner, W., Weiss, T., Babiloni, F., Witte, H., (2010). *A new Kalman filter approach for the estimation of high-dimensional time-variant multivariate AR models and its application in analysis of laser-evoked brain potentials*. Neuroimage 50, 960–969.
- Sternberg, S., (1966). *High-speed scanning in human memory* 53, 652–654.

# Automatic Pause Detection during P300 Web Browsing

A. Pinegger<sup>1</sup>, L. Deckert<sup>1</sup>, S. Halder<sup>2</sup>, J. Faller<sup>1</sup>,  
I. Käthner<sup>2</sup>, S. C. Wriessnegger<sup>1</sup>, A. Kübler<sup>2</sup> and G. R. Müller-Putz<sup>1</sup>

<sup>1</sup> Institute for Knowledge Discovery, Graz University of Technology, Graz, Austria

<sup>2</sup> Institute of Psychology, University of Würzburg, Würzburg, Germany

a.pinegger@tugraz.at

## Abstract

Brain-computer interfaces (BCI) have been investigated for more than 40 years. P300-based BCIs can nowadays control very complex applications such as spelling applications and even web browsers. To do this in a practical way, it is essential to have the possibility to pause the command selection of the BCI. We implemented and evaluated an automatic pause detection system for P300-based BCIs based on artifact detection with an inverse filtering method. Experimental results of an offline study (9 healthy participants) demonstrate the feasibility of the proposed approach and its high performance.

## 1 Introduction

The electroencephalogram (EEG) can be used to establish a noninvasive communication/control channel between the human brain and a computer, a so-called brain-computer interface (BCI). A very prominent BCI application is the P300 speller [1]. This system enables healthy as well as severely impaired users to communicate [2, 4]. However, a standard P300 speller is designed to work in synchronous mode, i.e., after defined stimulation sequences one item of all selectable items will be selected. This is not an issue as long as the user just wants to write a text without making a pause. However, if the user wants to make a pause during spelling a text or because she/he wants to look at the content of a web page it becomes a substantial problem. A very simple approach to avoid unintended selections is to include a pause element into the spelling matrix. This approach has two main disadvantages: First, you have to select two correct elements to go into and leave the pause mode and second, there is a certain probability that the pause-end element is selected by chance.

In this study, we introduce an automatic pause detection method on the basis of artifact detection with inverse filtering. Originally, the inverse filtering method was introduced in [5] to detect muscle and movement artifacts in the EEG of a sensorimotor rhythm (SMR)-based BCI. The main idea behind our approach is that a user produces more EEG artifacts during a pause than when she/he is actively engaged with the BCI. We use this difference to distinguish between the pause and the control state, i.e., when the user wants to select something.

## 2 Methods

### 2.1 Participants, Data Acquisition, and Experimental Design

Ten volunteers participated in this study. All participants stated that they have no history of neurological or psychiatric disorders. Due to a technical problem the data of one participant was not useable for this study. The final sample comprised 9 participants (3 female; mean age

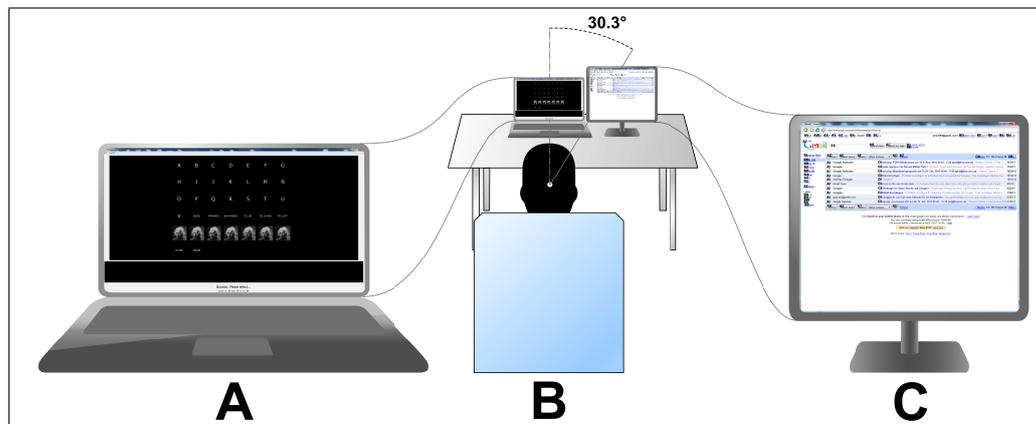


Figure 1: (A) Laptop displaying the user interface for feedback and P300 stimulation. (B) Sketch of the experimental design. The angle between the participant, the laptop, and the monitor was  $30.3^\circ$ . (C) Monitor for the web browser.

$23.9 \pm 1.3$  years).

EEG was acquired with a wireless EEG amplifier with dry electrodes (g.Nautilus, Guger Technologies OG, Graz, Austria). Signals from Fz, Cz, Pz, PO7, PO8, and Oz were used in this study with a sampling rate of 250 Hz. The channels were referenced to the right mastoid and grounded at the left mastoid. The raw signal of the Wi-Fi headset was filtered with a 0.5–30 Hz bandpass.

The participants were seated in a comfortable chair approximately 65 cm away from two screens (39.5 cm and 43 cm diameter), see Figure 1 (B). One screen was centered in front of the participants. At this screen a P300 matrix was displayed to control a special web browser (see Halder et al., in preparation), which was shown on a second screen placed right beside the first one, see Figure 1 (A) and (C).

The P300 user interface and signal processing in Matlab (MathWorks, Natick, USA) was presented in [3].

Calibration was performed with fifteen highlightings per row and column and ten letters as described in [3]. The online task for the participants was to write an email to a given address and reply to an automatically generated email. The whole task needed a minimum of 52 selections and was aborted if the goal was not reached within 78 regular selections.

## 2.2 Manual Pause

A “PAUSE/RUN” element was selectable with the matrix. If the user selected this element, no further selections were sent to the web browser until the same element was selected again. The participants had to select the “PAUSE/RUN” element in the study when they were waiting for the reply of the first email and they could select it whenever they needed a pause.

## 2.3 Automatic Pause Detection

The automatic pause detection was performed on the offline data by detecting artifacts in the EEG during the flashing time periods. The principle of inverse filtering was applied to detect the artifacts, cf. [5]. For this method autoregressive filter model parameters have to be estimated

out of clean (i.e., artifact free) EEG data. Our assumption was that the participants generated few artifacts during the P300 calibration period. Consequently, we used the data of the P300 calibration to estimate autoregressive filter model parameters (model order  $p = 10$ ) by using the Burg method.

The created filter model was applied inversely to the EEG data of every online task selection. An artifact detection threshold was set to three times the standard deviation of values calculated with this inverse filter from the calibration EEG data. If in more than 1 percent of the online task data artifacts were detected, the selection was marked as pause state related.

### 3 Results

The participants needed on average 11.8 (SD 2.9) highlighting sequences to select a command. Eight participants completed the task within the maximum allowed value of 78 selections. Only participant S7 did not complete the task. They had an average selection accuracy of 88.7% (SD 9.4) and needed an average time of 63.1 (SD 16.4) minutes including pauses to complete the whole task.

#### 3.1 Manual Pause

Two selections were necessary to go manually into pause mode and leave the pause mode. The time the participants needed to perform these two selections was between 58 and 100 seconds depending on the number of highlighting sequences and the actual number of rows and columns. Seven participants had no problem to switch between pause and control mode. Two participants (S6, S7) needed more than one attempt to leave the pause mode. The probability that the pause mode was left by chance was  $1/N$  with  $N$  being the actual number of matrix elements. This occurred once (S7) in this study.

#### 3.2 Automatic Pause Detection

Participant	All	Pause Detection						Cohen's
	Selections <sup>a</sup>	TP	TN	FP	FN	TPR	TNR	$\kappa$
S1	61	5 ( 6)	52	4 (3)	0	100.0%	92.9%	0.68
S2	62	7 ( 7)	46	9 (9)	0	100.0%	83.6%	0.54
S3	61	3 ( 5)	56	2 (0)	0	100.0%	96.6%	0.73
S4	66	2 ( 2)	63	1 (1)	0	100.0%	98.4%	0.79
S5	93	26 (29)	63	4 (1)	0	100.0%	94.0%	0.90
S6	101	8 ( 9)	92	1 (0)	0	100.0%	98.9%	0.94
S7	94	9 (12)	74	6 (3)	5	64.3%	92.5%	0.55
S8	73	10 (12)	60	3 (1)	0	100.0%	95.2%	0.85
S9	64	6 ( 6)	64	0 (0)	0	100.0%	100.0%	1.00

<sup>a</sup> incl. selections during pause.

Table 1: Automatic pause state detection results. The true positive (TP), true negative (TN), false positive (FP), and false negative (FN) detections as well as the true positive rate (TPR) and the true negative rate (TNR) are presented for every subject. Values in parentheses indicate prevented wrong item selections during the control state. Cohen's Kappa was calculated to give a measure of agreement.

In Table 1 the offline simulation results of the automatic pause detection are shown. The number of selections of some participants was higher than 78 because selections during the pause were also counted for evaluation. The sensitivity (TPR) of the automatic pause detection was 100 % for eight participants and 64.3 % for one participant. Consequently, the overall mean TPR was 96.0% (SD: 11.9). The specificity (TNR) was between 83.6 % and 100% with a mean of 94.7 % (SD: 4.9). At eight participants no false negative (FN) detections were made, see 6<sup>th</sup> column in Table 1. This is very important because false negative detections would result in unintended, random selections. The measured Cohen's Kappas for our results were between 0.54 and 1.00, indicating moderate to strong agreement.

## 4 Discussion

In this study we provide evidence that our suggested automatic pause detection method works comparable to the manual pause selection method without its disadvantages.

The introduced automatic pause detection method detected the pause state with 100% accuracy at eight participants and the control state with an accuracy between 83.64% and 100%. Unintended selections in the pause state are almost non-existent and the number of prevented selections in the control state is low and acceptable. Considering the prevented wrong selections during the voluntary control periods by the automatic pause detection the number of wrong classifications would be even lower (numbers in parentheses beside the TPs and FPs in Table 1). In conclusion, this study shows that detecting artifacts in a P300-based BCI can be used as a very reliable and effective automatic P300 pause state detection method.

## Acknowledgments

This work is supported by the FP7 research project BackHome (No. 288566). This paper only reflects the authors' views and funding agencies are not liable for any use that may be made of the information contained herein.

## References

- [1] 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*, 70:510–523, 1988.
- [2] F. Piccione, F. Giorgi, P. Tonin, K. Priftis, S. Giove, S. Silvoni, G. Palmas, and F. Beverina. P300-based brain computer interface: reliability and performance in healthy and paralysed participants. *Clinical Neurophysiology*, 117(3):531–537, 2006.
- [3] A. Pinegger, S. Wriessnegger, and G. Müller-Putz. Introduction of a universal P300 brain-computer interface communication system. In *Biomedical Engineering*, volume 58, 2013.
- [4] C. Pokorny, D. S. Klobassa, G. Pichler, H. Erlbeck, R. G. L. Real, A. Kübler, D. Lesenfants, D. Habbal, Q. Noirhomme, M. Riseti, D. Mattia, and G. R. Müller-Putz. The auditory P300-based single-switch BCI: Paradigm transition from healthy subjects to minimally conscious patients. *Artif Intell Med*, 59(2):81–90, 2013.
- [5] R. Scherer, A. Schlögl, F. Lee, H. Bischof, J. Jansa, and G. Pfurtscheller. The self-paced Graz brain-computer interface: methods and applications. *Computational Intelligence and Neuroscience*, 2007.

# Eye-blink related changes in EEG during an auditory working-memory task performance

Petar Horki<sup>1</sup> and Gernot R. Müller-Putz<sup>1</sup>

<sup>1</sup>Institute for Knowledge Discovery, BCI Lab, Graz University of Technology, Graz, Austria  
petar.horki@tugraz.at, gernot.mueller@tugraz.at

## Abstract

We investigated eye-blink related changes in EEG during an auditory working memory task. To that end, we used the onset of eye blinks following the presentation of the target letter as triggers for data analysis. No evidence for an active role of spontaneous eye-blinks in the working memory task could be found.

## 1 Introduction

In (Horki and Müller-Putz, 2013) we reported frontal EEG theta band oscillations for a mental task related to working memory, but found no such event-related changes for a control task. Based on the current literature, we assumed these EEG changes were not due to eye-movement artifact, such as eye-blinks. However, having performed a more detailed analysis, precluded by a rigorous artifact rejection and removal procedure, we were surprised to find these EEG changes absent in the cleaned data. Thus, it seems that some users exhibited spontaneous eye-blinks correlated to the presentation of the target stimuli, even after being instructed differently.

A closer visual inspection of the original EEG data revealed that, while performing the working memory tasks, eight out of eleven participants blinked with their eyes more often than every second time. Was this merely a coincidence, or where these eye-blinks somehow task-related? The latter case could lead to a classifier adapted to eye-movements, and thus to a biased evaluation.

Spontaneous eye-blinks have been observed at breakpoints of attention during reading, listening to speech, and while viewing videos (Nakano et al., 2013). In a recent work (Nakano et al. 2013), it was suggested that spontaneous eye-blinks play an active role in the release of attention from external stimuli while attentively engaging in a cognitive task. Based on these findings, we investigated whether the eye-blinks observed during our experiment are somehow related to the working memory task. To that end, we have redone the analysis from (Horki and Müller-Putz, 2013), this time using the onset of eye blinks following the presentation of the target letter as triggers.

## 2 Methods

### 2.1 Subjects

Eleven healthy subjects (5 male, 6 female; 22 to 29 year old, mean age 26) participated in this experiment. They were recruited through university public notice boards (i.e. newsgroup, forum). Participants gave informed consent prior to the beginning of the experiments and received monetary compensation afterwards. Half of the participants had no previous exposure to EEG experiments. The experiment was undertaken in accordance with the Declaration of Helsinki.

## 2.2 Recording

The EEG was recorded with 29 active electrodes (g.tec, Guger Technologies, Graz, Austria) overlying the frontal, central, and parietal scalp areas. In detail, the electrodes were placed at positions AFz, F3, F1, Fz, F2, F4, FC2, FC1, FCz, FC4, C5, C3, C1, Cz, C2, C4, C6, CP3, CP1, CPz, CP2, CP4, P3, P1, Pz, P2, P4 and POz according to the international 10 / 20 electrode system. The EEG electrodes were referenced to the left ear lobe with the ground electrode placed on the right ear lobe. The electrodes were integrated into a standard EEG cap (Easycap GmbH, Herrsching, Germany) with an inter-electrode distance of 2.5 cm and connected to EEG amplifiers (g.tec, Graz, Austria).

The electrooculogram (EOG) was recorded with three active electrodes (g.tec, Guget Technologies, Graz, Austria), positioned above the nasion, and below the outer canthi of the eyes. The electromyogram (EMG) was recorded with four electrodes from both legs (musculustibialis anterior). The EEG amplifiers were set up with a bandpass filter between 0.5 and 100 Hz, and a notch filter at 50 Hz. The EEG and EOG were sampled with 512 Hz, the EMG with 2000 Hz. Participants were seated in an electrically shielded room.

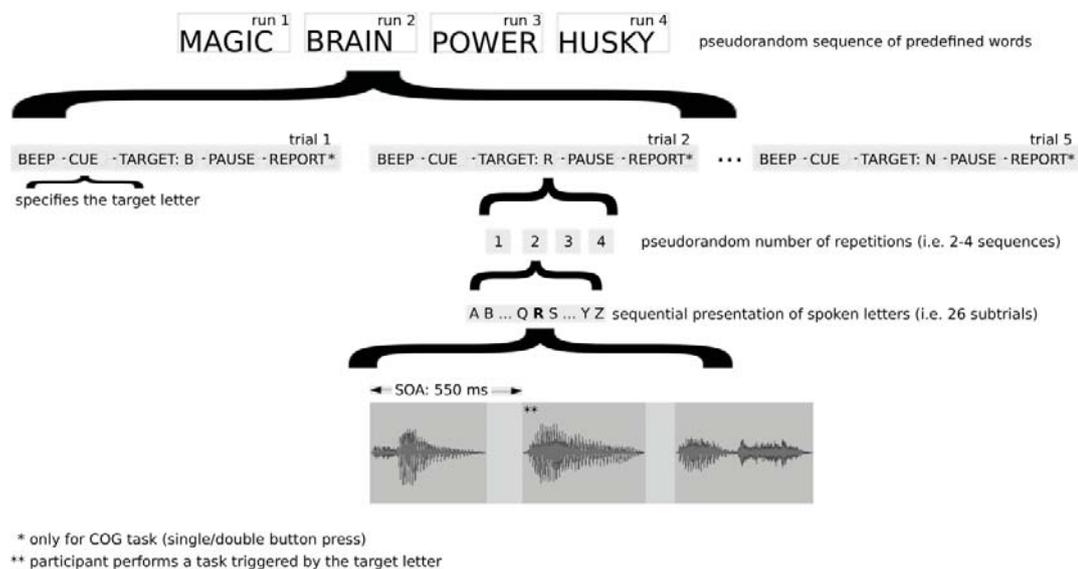
## 2.3 Stimuli

Spoken letters of the English alphabet, generated by a text-to-speech program (AT&T Natural Voices, AT&T, USA), were presented sequentially in alphabetical order through a right head phone for one of several predefined words. Presenting acoustic cues through one ear only, keeps the other ear free for incoming communication from surroundings. The task irrelevant acoustic cues were presented in either male or female voice, and were balanced across all the subjects. Stimulus onset asynchrony was set to 550 ms, including a 50 ms pause. Thus, it took 14.3 s for a single presentation of the whole alphabet. For each target letter, indicated through a verbal cue, the alphabet was repeated one to three times, for a total of two to four alphabet presentations, followed by a short break of random length (i.e. four to six seconds).

## 2.4 Experimental paradigm

The experimental paradigm is depicted in Figure 1. For the investigation the predefined words “brain”, “power”, “husky” and “magic” – had to be spelled in copy spelling mode. They were chosen because their letters are distributed across the whole alphabet range. Each word was spelled letter by letter within a single run. Runs were separated by short break of 1-2 min to avoid fatigue.

The participants were instructed verbally to perform either a motor or a non-motor mental task whenever a target letter was presented. For the purpose of this abstract, we focus only on the following non-motor mental tasks: (i) discrimination of the target voice’s gender and comparison to the following repetition (i.e. whether the target voice’s gender has changed or it remained the same; reporting through single / double button press with index finger of the right hand in a dedicated time window) as a cognitive task (COG); and (ii) mental repetition of the target letter as a control condition (AEP). The COG and AEP conditions were pseudo randomized. We randomized the order of words, and balanced the voice of presentation (male / female). The participants were also verbally instructed to avoid any movements, and received no feedback.



## 2.5 Data analysis

EEG analysis was performed for the COG tasks using MATLAB 2009a (MathWorks, USA) and EEGLAB version 11.

The data was high-pass filtered (3rd order butterworth filter) with cut-off frequency at 1 Hz, and segmented into consecutive epochs of 0.5 s. Bad channels and prominent artifacts (i.e. swallowing, electrode cable movements, etc.) were identified by visual inspection and removed. The data was triggered by the onset of eye blinks. To that end, only the first eye blink during the 2s following the onset of the target letter presentation was used to determine the trigger. The eye-blinks before the target letter presentations were ignored. The same triggers were used to calculate grand average EEG event-related potentials (ERPs).

To analyze the percentage of power decrease (ERD) or power increase (ERS) relative to a reference interval (0.5 s preceding the stimulus onset), a time-frequency map for frequency bands between 4 and 40 Hz (35 overlapping bands using a band width of 2 Hz) was calculated on downsampled (128 Hz) data. Logarithmic band power features, calculated by band-pass filtering, squaring and subsequently averaging over the trials, were used to assess changes in the frequency domain. To determine the statistical significance of the ERD/ERS values a t-percentile bootstrap algorithm with a significance level of  $\alpha=0.01$  was applied. The ERDS analysis was conducted for the COG task on a single bipolar derivation AFz-Fz.

## 3 Results

In Table 1 the number of epochs containing eye-blinks during the 2 s following the target letter presentation, the most frequent eye-blink trigger latency relative to the target letter presentation onset, and the accuracy of reporting, for eight participants, are shown for the COG condition. The maximum number of epochs per participant possible was 60, and only the participants with 30 eye-blink epochs or more were included in the analysis. The distribution of the aforementioned eye-blink trigger latency

could be approximated by a normal distribution peaking at around 0.8 s relative to the target letter presentation onset.

	S1	S2	S3	S4	S5	S6	S7	S8	$\mu \pm \sigma$
#Eye-blink epochs	44	40	47	30	38	41	43	35	40 $\pm$ 5
Trigger latency [s]	0.7	0.8	0.5	0.8	0.6	0.8	0.6	0.8	0.7 $\pm$ 0.1
Reporting acc [%]	90	100	100	90	100	100	40	90	89 $\pm$ 20

**Table 1:** Number of epochs containing eye-blinks during the 2s following the target letter presentation (COG condition), the most frequent eye-blink trigger latency relative to the target letter presentation onset, and the accuracy of reporting, for eight participants. The maximum number of epochs per participant possible

Reanalysis of eye-blink triggered data revealed no significant ( $p=0.01$ ) frontal oscillatory EEG changes. The only time oscillatory EEG changes were observed, was when the original triggers (i.e. onset of the target letter presentation) were employed instead, indicating that these EEG changes were due to eye-blink artifacts. The grand average EEG event-related potentials were, as expected, dominated by the eye-blinks used to trigger the data.

## 4 Discussion and Conclusion

Using the methods from (Horki and Müller-Putz, 2013), no evidence for an active role of spontaneous eye-blinks in the working memory (COG) task could be found. However, methods for analysis of connectivity may provide further insight. Removing the EOG artifacts in COG condition revealed task-related modulation of several ERP components (Horki et al. 2014).

Notable is that in (Horki and Müller-Putz, 2013) a separate control condition (AEP) yielded different results in the presence of eye-blink artifacts. Also, for the AEP condition there were only two out of eight participants with 30 eye-blink epochs or more ( $\mu \pm \sigma$ : 22 $\pm$ 13, range: 8 to 38), with neither an apparent distribution nor mode of eye-blinks.

## 5 Acknowledgements

Contribution of Daniela S. Klobassa and C. Pokorny is gratefully acknowledged. The authors would like to thank C. Breitwieser and M. Billinger for advice and assistance in the development of aspects of the software used in this work. This research was supported by the European ICT Programme Project FP7-247919 (DECODER) and the "Land Steiermark" Project ABT08-22-T-7/2013-17 (EEG@Wachkoma). This paper only reflects the authors' views and funding agencies are not liable.

## References

- P. Horki, and G. Müller-Putz. (2013) *A Novel Approach to Auditory EEG-Based Spelling*. Proceedings of the Fifth international BCI meeting, pp. 316 - 317.
- P. Horki, D. S. Klobassa, C. Pokorny, and G. Müller-Putz. (2014) *Evaluation of Healthy EEG Responses for Spelling Through Listener-Assisted Scanning*, IEEE Journal of Biomedical and Health Informatics, in press.
- T. Nakano, M. Kato, Y. Morito, S. Itoi, and S. Kitazawa. (2013) *Blink-related momentary activation of the default mode network while viewing videos*. PNAS 110 (2), pp. 702-706.

# Augmenting communication, emotion expression and interaction capabilities of individuals with cerebral palsy

Reinhold Scherer<sup>1</sup>, Johanna Wagner<sup>1</sup>, Martin Billinger<sup>1</sup>, Gernot Müller-Putz<sup>1</sup>, Rafael Raya<sup>2</sup>, Eduardo Rocon de Lima<sup>2</sup>, Yehya Mohamad<sup>3</sup>, Dirk Tassilo Hettich<sup>4</sup>, Elaina Bolinger<sup>4</sup>, Marco Iosa<sup>5</sup>, Febo Cincotti<sup>5</sup>, Ana Rita Londral<sup>6</sup>, Joana Mesquita<sup>6</sup>, Mariano Lloria Garcia<sup>7</sup> and Juanma Belda<sup>8</sup>

<sup>1</sup> Institute for Knowledge Discovery, Graz University of Technology, Graz, Austria  
{reinhold.scherer, johanna.wagner, martin.billinger, gernot.mueller}@tugraz.at

<sup>2</sup> Bioengineering Group, Spanish National Council for Science Research (CSIC), Madrid Spain  
{rafael.raya, e.rocon}@csic.es

<sup>3</sup> Fraunhofer FIT, Sankt Augustin, Germany  
yehya.mohamad@fit.fraunhofer.de

<sup>4</sup> Erberhard-Karls-University of Tübingen, Tübingen, Germany  
dirk-tassilo.hettich@uni-tuebingen.de, elaina.bolinger@med.uni-tuebingen.de

<sup>5</sup> Neuroelectrical Imaging and BCI Lab, IRCCS Fondazione Santa Lucia, Rome, Italy  
{m.iosa, f.cincotti}@hsantalucia.it

<sup>6</sup> PLUX Wireless Biosignals, Lisbon, Portugal  
{alondral, jmesquita}@plux.info

<sup>7</sup> Avapace, Asociacion Valenciana de ayuda a la paralisis cerebral, Valencia, Spain  
coordinacion@avapace.org

<sup>8</sup> Insituto de Biomecanica de Valencia, Universitat Politecnica de Valencia, Valencia, Spain  
juanma.belda@ibv.upv.es

## Abstract

Providing individuals with cerebral palsy (CP) tools to communicate and interact with the environment independently and reliably since childhood would allow for a more active participation in education and social life. We outline first steps towards the development of such a hybrid brain-computer interface-based (BCI) communication tool.

## 1 Introduction

Cerebral palsy (CP) is a non-progressive condition caused by damage to the brain during early developmental stages (annual incidence of 2 per 1000 live births in Europe). Individuals with CP may have a range of problems related to motor control, speech, comprehension, or mental retardation. The majority of children with CP have normal intelligence, however, due to lack of appropriate communication means they are classified as educational abnormal. Communication solutions are available; however, they strongly depend on motor activity and assistance of others.

The aim of the ABC project (<http://abc-project.eu>) is to develop an interface for individuals with CP that improves independent interaction, enhances non-verbal communication and allows expressing and managing emotions. Both motor activity and (electro)physiological signals will serve as input for the interface. In fact, one key component of the ABC system is a non-invasive electroencephalogram-based (EEG) brain-computer interface (BCI). In this paper, we outline the basic structure of the ABC system and briefly report on the functionality of individual modules, which were developed following user-centered design principles, i.e., in accordance with user needs and requirements defined in terms of usability and functionality.

## 2 Functional ABC Prototype Modules

The ABC system has three main functions: (i) Intentional communication (by use of EEG, electromyography (EMG) and inertial measurement units (IMUs)), (ii) emotion expression and management (by EEG, and electro-dermal activity (EDA) and blood volume pulse (BVP) sensors placed at the wrist), and (iii) health monitoring (by use of a sensing chest band with electrocardiogram (ECG), respiration and accelerometer sensors). The Communicator application is the graphical user interface (GUI) for accessing all features of the ABC system

### 2.1 Communicator

The Communicator application provides different pointing and scanning techniques (dependent on input signal) for selecting available options. Menu options (pictograms) describing the actions to be executed (e.g. access to video, music or social network) are arranged in a grid (matrix). Audio and visual feedback informs the user about selected choices and the action being performed. Changes in the affective state of the user and vital health parameters are shown in a status window. As impairments and capabilities in CP users vary considerably between individuals, the communication with the system can be tailored to the specific needs and capabilities of each individual user.

### 2.2 Communication

#### 2.2.1 EEG-based interface

Developing BCIs for CP user is challenging for several reasons. Firstly, more or less frequent spasms and involuntary movements negatively impact on the EEG signal quality. Secondly, BCI paradigms have mostly been developed for able-bodied individuals. Hence, experimental paradigms need to be adapted to the end-users capabilities following user-centered design principles. Thirdly, data collection is time consuming and tedious for the user. Hence, besides the need to develop novel motivating training paradigms, machine learning methods need to be improved in a way to handle the lack of training data and infer robust models that predict the users intend from EEG signals [1, 2].

An imagery-based BCI (brain-switch) was developed to give CP users on-demand access to the Communicator by row-column scanning protocols. Spontaneous EEG oscillations were selected instead of evoked-potentials to encode messages to avoid the risk of epileptic seizures in younger CP users. Note, that the prototype will be most useful to children. To enhance robustness and reliability of mental imagery detection during early training, i.e., when activity patterns are usually not well established and correct detection may be challenging, selection redundancies were integrated. Currently user have to confirm a selection  $n = 3$  out of  $m = 5$  times. Test in healthy users confirm that, at the cost of an increased selection time, the developed strategy enables users to make robust selections even when binary classification performance was around 80%. Experiments in CP users are currently in progress.

One core component of the BCI system is a novel fully-automatic artifact reduction method [3]. The method, based on online Wavelet decomposition, independent component analysis and thresholding, automatically removes a number of different artifacts while preserving information that is useful for BCI. Experiments in healthy users suggest that BCI operation is still possible while users for example walk and hold objects with their hand [4].

### 2.2.2 Head mounted interface

A head-mounted interface based on inertial technology has been developed. This interface translates head movements and posture of the user with CP into mouse pointer positions following an absolute control (based on angular orientation). The inertial interface consists of a headset with an inertial measurement unit (IMU). The IMU integrates a three-axis gyroscope, accelerometer and magnetometer. The first study was focused on characterizing the motor capabilities and abnormalities of users with CP. The task consisted in reaching targets on the screen using the head motion. We selected some outcome metrics to evaluate the performance of the users. Three healthy subjects participated as control group. 1) Throughput (TP), a metric proposed by the ISO9241 that measures the usability of the interface, 2) Frequency (f) of the head motion, aiming to identify some abnormal movements such as tremor or spasms, 3) Range of motion (ROM) of the head, aiming to identify difficulty to maintain the posture to control the interface. Results showed that there are no significant differences between users with CP and healthy control participant in the frequency domain of the head movement. However, there are significant differences for posture control between healthy users and subjects with CP. This result suggests that posture control leads to a lack of usability more than high frequency movements. The results and conclusions of this work are considered an objective description of user needs, which should be used to optimize the inertial interface. Based on these results, a new control mode based on angular velocity (relative) instead of orientation (absolute) has been implemented. Additionally, the EMG signal is being studied to generate the click in a natural way instead of dwell click that resulted complicated for users with CP.

## 2.3 Emotion detection and communication

Auditory affect induction paradigms based on short emotionally-laden sounds from the International Affective Digitized Sounds (2nd edition) database were developed to investigate electro-dermal activity (EDA) and EEG activity of emotional processing in individuals with CP in the absence of the ability to visually fixate on traditional emotional imagery.

An EDA-based emotion detection module was developed that discriminates between five distinct affective states: high arousal positive, low arousal positive, neutral, low arousal negative and high arousal negative. An average 5-class classification performance of 80% was computed over 12 CP users. To implement a reaction of the system to the detected emotional state e.g. play music, we have furthermore developed an emotion management system (EMS). EMS allows the caregiver of CP users to define how the system reacts to the current detected emotional state. EMS was evaluated with 15 CP users. All users appreciated the EMS and gave many suggestions for improvements, the majority of these suggestions has been implemented.

EEG-based emotion detection is work in progress. Time and frequency domain EEG features of emotion, such as the late positive potential (LPP), as well as valence-dependent inter-hemispheric alpha and event-related de/synchronization effects were analyzed. Promising preliminary LPP and inter-hemispheric alpha activity results are comparable to literature from a healthy population. Movement artifact-contaminated EEG, however, resulted in low accuracies in CP users. The custom developed artifact rejection method is currently being adapted to clean the data.

## 2.4 Health Monitoring

The main health-related problem of individuals with CP is dyspnea. In fact, 59% of the immediate causes of death in this population are diseases related to respiratory system; other

21% are related to infections and inflammations of the lungs (pneumonia). Accordingly, a health monitoring module was integrated into the ABC prototype. Conceptually, it is divided into clinical health monitoring for severely affected people and physical activity monitoring for mildly affected ones. Movement artifacts make clinical health-state monitoring challenging. The system, however, provides helpful information on vital parameters during sleep or periods of reduced movement activity. The accelerometer embedded in the chest band provides important information on the quantity of locomotion performed by the individual during the day. This information can be therapeutically important for mildly to moderately affected individuals.

### 3 Conclusion and Future Work

The ABC prototype is a modular and flexible system that can be adapted according to the users need. This is especially important for CP users as impairments and capabilities vary considerably between individuals. Pilot studies with the head mounted IMU interface and emotion detection by EDA activity show that these methods can be applied reasonably well and are accepted by CP users. Besides improving robustness and performance of the developed functions, in a next step individuals modules will be tested in daily use by CP users.

Movement artifacts are one major issue for the sensor network and EEG signals are most severely affected. To ensure that correlates of cortical activity are used for BCI control, focus was put on the characterization of "clean" EEG and the development of artifact reduction and detection methods. A novel robust protocol for imagery-BCI based interaction has been tested in healthy users and is currently being evaluated in CP users. Note, results in healthy users cannot directly be transferred to real life situation in users with CP. In a next step the prototype will be updated according to the CP users comments and training paradigms will be adapted. The updated ABC prototype system will provide an excellent basis to study advantages and limitations of imagery-BCI control in CP. We expect that, the system will allow us to identify scenarios and applications where CP users truly benefit from the use of BCI technology.

### Acknowledgments

This work was supported by the FP7 Framework EU Research Project ABC (No. 287774). This paper only reflects the authors views and funding agencies are not liable for any use that may be made of the information contained herein.

### References

- [1] I. Daly, M. Billinger, J. Laparra-Hernández, F. Aloise, M. Lloria Garcia, J. Faller, R. Scherer, and G. Müller-Putz. On the control of brain-computer interfaces by users with cerebral palsy. *Clinical Neurophysiology*, 124(9):1787–1797, 2013.
- [2] I. Daly, J. Faller, R. Scherer, C.M. Sweeney-Reed, S.J. Nasuto, M. Billinger, and G.R. Müller-Putz. Exploration of the neural correlates of cerebral palsy for sensorimotor bci control. *Front Neuroeng*, 7:20, 2014.
- [3] I. Daly, R. Scherer, M. Billinger, and G. Müller-Putz. Force: Fully online and automated artifact removal for brain-computer interfacing. *IEEE Trans Neural Systems Rehab Eng*, accepted.
- [4] R. Scherer, J. Wagner, M. Billinger, and G. Müller-Putz. Online artifact reduction and sequential evidence accumulation enhances robustness of thought-based interaction. In *Proceedings of the 6th International BCI Conference, Sep. 16-19, 2014, Graz, Austria*, 2014.

# A virtuous BCI loop: adaptive decision making improves P300-spelling in two ways

M. Perrin<sup>1,2</sup>, E. Maby<sup>1,2</sup>, O. Bertrand<sup>1,2</sup> and J. Mattout<sup>1,2</sup>

<sup>1</sup>INSERM U1028, CNRS UMR5292, Lyon Neuroscience Research Center, Brain Dynamics and Cognition Team, Lyon, F-69000, France, <sup>2</sup>University Lyon 1, Lyon, F-69000, France  
Jeremie.mattout@inserm.fr

## Abstract

A major challenge in Brain-Computer Interfaces (BCI) is the optimization of performance, regardless of the within and between user variability. A promising option is to move towards adaptive BCI that can accommodate fluctuations over time and subjects. An obvious criterion to be optimized is the speed-accuracy trade-off. We implemented a BCI whose decision speed or reaction time depends on the reliability of accumulated evidence. We instantiated a probabilistic classifier whose outcome can be up-dated online based on new incoming information. An entropic measure then enables us to derive an optimal stopping strategy, in a user-dependent fashion. We evaluated the proposed approach on 11 participants, during online P300-Spelling. We thus quantified the beneficial effect of this method. Importantly, we show that the adaptive mode creates a virtuous circle such that: the higher the spelling accuracy, the more the participant engages in the task and, in return, the higher the motivation the higher the BCI performance.

## 1 Introduction

A central question in BCI is how fast the system can produce a reliable command. Although BCI performances are often evaluated and reported using measures of bit rate, this one is rarely explicitly maximized online by the system. The vast majority of P300-Speller studies do fix the time for spelling a letter in a way that is supposed to optimize speed given some expected level of accuracy. Yet it is clear that the same strategy might not be optimal for every individual, at any time. In a given subject, adaptation would consist of varying the number of flashes per trial, given some measures of the user's level of engagement. Such a system should stop earlier whenever the user is very well focused on the task, and it would keep acquiring data whenever the user is unfocused and produces ambiguous signals. Hence for a given averaged trial duration, we expect an increase in accuracy with the adaptive approach compared to the traditional one. Moreover, in the case of BCI, such an adaptive approach might also trigger up the user's motivation. If so, we expect the online results to reflect not only the improvement due to the adaptive method but also some further positive effect due to the ensuing boost in motivation. This issue of optimizing the stopping criterion has already been addressed in a couple of studies relying on different measures of how robust is the decision about to be made (Serby et al. 2005; Jin et al. 2011). Remarkably though, the authors used a still fairly rigid approach instead of a fully flexible one. Their approaches were restricted to stop acquiring new observations after some varying amount of repetitions (blocks), rather than considering the possibility of deciding after an arbitrary amount of flashes (trials). As a consequence, the machine's reaction time can only take a few discrete values. Our approach proposes an information theoretic and probabilistic criterion, which enables us to generalize this optimal strategy, by allowing the machine to stop at any time during the evidence accumulation process.

## 2 Method

### 2.1 Data acquisition

Eleven healthy subjects took part in this study (4 men, mean age =  $26.9 \pm 6.4$  (SD), range 19-40). They all signed an informed consent approved by the local Ethical Committee. We used a traditional 6x6 matrix made of letters (A-Z), digits (1-9) and an additional symbol for blanks ( $\square$ ). Pseudo-random groups of letters (adapted from Townsend et al, 2010) were flashed alternatively for 80ms, while the SOA was set to 150ms. The whole experiment included one training followed by three test sessions. The former consisted of 15 characters spelled with 10 repetitions and each test session was made of 20 5-letter words. Target letters were defined prior to the experiment and indicated to the subject by a green circle. Subjects were instructed to visually fixate the target and count how many times it was flashed.

We here report the online and offline comparison between two experimental conditions, which differed in the way the decision was made: one used a time-based decision (called the fixed condition) and the other involved an accuracy-based decision (called the adaptive condition). In the fixed condition, each trial consisted of 60 flashes (5 repetitions). In the adaptive condition, there was a maximum of 180 flashes by item (15 repetitions), but the actual number of flashes varied from one trial to the next, depending on the pre-determined threshold and the entropy of the current posterior probability distribution. The threshold was determined from each individual training set so that the average number of flashes equalled roughly 60 (5 repetitions). Participants spelled 4 blocks of five 5-letter words per condition and conditions were presented pseudo-randomly over time. The experiment lasted about an hour in total for each subject.

EEG data were recorded from 9 electrodes (Pz, P7, P8, P3, P4, PO9, PO10, O1, O2) following the extended 10-20 system and referenced to the nose. Data were digitized at 1000 Hz, band-pass filtered between 1 and 20 Hz and down-sampled at 100 Hz.

### 2.2 Feature extraction, classification and decision

Feature extraction consisted in a linear spatial filtering named xDAWN (Rivet et al., 2009). For classification, we used a simple probabilistic generative model of the data, based on a two multivariate-Gaussian mixture, further assuming conditional independence between features, over time and space (Naïve Bayes hypothesis). Importantly, we here extended our model to compute the posterior probability associated with each item of the matrix in a Markovian fashion; that is by applying Bayes rule after each new flash and considering the posterior belief as the prior for the next observation. After each flash, this method enables us to compute and update each letter's probability of being the target. Based on these up-dates, the adaptive decision relies on a natural information theoretic measure of uncertainty, the Shannon's entropy of the posterior distribution. Entropy decreases as information is accumulated, *i.e.* as the posterior distribution gets closer to an ideal distribution with full probability mass associated with a single item, meaning that the machine is sure about the target location. In the adaptive condition, a decision is made as soon as the entropy falls below the individually chosen threshold. By default, if the threshold is never met, a decision is made after fifteen repetitions.

### 2.3 Evaluation metrics and statistical tests

In order to compare the above described methods and conditions, we used two well-known measures of performance: spelling accuracy and bit rate (in bits/minute) as defined in (Wolpaw et al., 2000). To assess the statistical significance of differences in performance, we compared spelling accuracies, averaged numbers of flashes and bit rates using Wilcoxon tests.

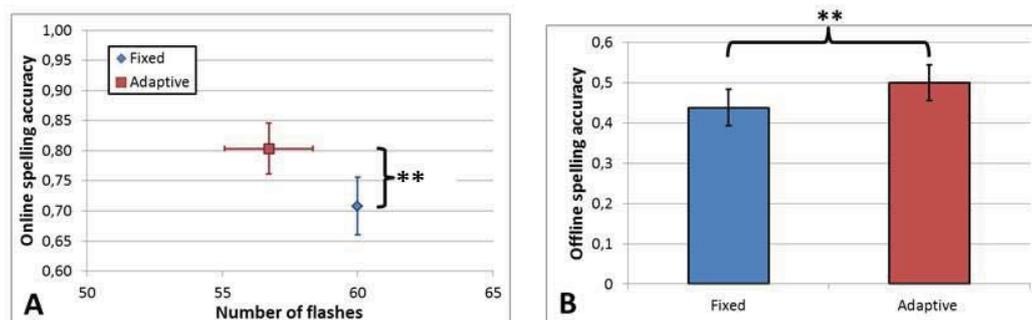
In order to evaluate a putative additional effect of motivation in the adaptive condition, we reprocessed some of the data offline in order to compare the two conditions based on the same fixed number of observations (24 flashes, i.e. 2 repetitions).

### 3 Results

In the fixed condition, the online spelling accuracy was  $71\% \pm 16$  (SD), which corresponds to 18.8 bits/minute. In the adaptive condition, it was  $80\% \pm 14$  (SD), for an average of  $57 \pm 4$  (SD) flashes, which corresponds to 24.1 bits/minute. (Figure 1.A).

Wilcoxon tests revealed that both the spelling accuracy and the bit rate are significantly higher in the adaptive condition compared to the fixed condition ( $p < 0.01$  for both tests). Importantly, the number of flashes was not significantly different between the two conditions ( $p = 0.1$ ), it was even slightly lower in the adaptive condition.

To evaluate the effect of motivation, data from both conditions were reanalyzed offline, using the same time-based stopping criterion: a decision was made after the 24 first flashes. The obtained spelling accuracy proved significantly higher in the adaptive than in the fixed condition ( $p < 0.01$ ) (Figure 1.B). Since the number of observations was the same for both conditions, the ensuing bit rate proved also significantly higher in that same condition ( $p < 0.01$ ).



**Figure 1.** A/ Online spelling accuracy as a function of the averaged number of flashes for each condition: fixed (blue diamond), adaptive (red square). Error bars indicate the standard error of the mean. B/ Offline spelling accuracies obtained with the same datasets reanalyzed using a time-based criterion (decision made after 24 flashes). P-value: \*\*  $p < 0.01$ .

### 4 Discussion

In this P300-Speller study, we first developed a new classification approach which up-dates the belief of the machine about target location, after each single electrophysiological observation. This single-trial based classification enabled us to propose and evaluate an adaptive decision making, which consist of implementing an optimal reaction time strategy in the machine, allowing for short spelling when the first few incoming pieces of evidence are strong enough and vice versa.

Adaptive decision making was proposed to overcome the limitation of the traditional time-based decision criterion used in the P300-speller and BCI in general. Indeed, a machine's adaptive decision, based on some information or accuracy criterion, allows for an optimal stopping strategy. In other words, the reaction time of the machine can be optimized in a way that mimics the reaction time of human beings, which highly relies on the amount and quality of accumulated evidence from incoming sensory information. What is expected from such a strategy is to produce a short reaction time, whenever the accumulated evidence in favor of a given single choice is strong.

Conversely, reaction time should be longer, whenever evidence is noisy and ambiguous, since more data will be needed to make a reliable decision. Compared to a time-based criterion, this can accommodate the slow intrinsic fluctuations of the electrophysiological signals, which might be due to fluctuations in attention.

In the P300-Speller, this is particularly relevant, since sustained attention is what is required from the subject to keep performing the task efficiently. To implement adaptive decision making, we used a classical entropic measure, which efficiently summarizes and quantifies the uncertainty about our belief, the latter being represented by a probabilistic distribution.

The first significant effect we indeed observed with this new criterion is that, for the same spelling duration, the user is able to spell letters more accurately. The time saved by stopping the flashes earlier, whenever possible, was efficiently reallocated to letters that required longer stimulation time in order to be accurately identified. Equivalently, given an objective in terms of accuracy, fewer flashes should be required with adaptive decision making, on average.

Secondly, a very interesting and significant effect of motivation could be observed online. Indeed, spelling accuracy was found higher for adaptive sessions than for fixed ones when these datasets were reprocessed offline with the same stopping criterion. This suggests that the subjects were on average more engaged into the task during the adaptive session, thus producing electrophysiological responses with a larger signal-to-noise ratio, which resulted in higher spelling accuracies. Indeed, the N1 and P300 responses, which are the electrophysiological responses used to identify the target, are known to reflect the participant's involvement in the task (Treder and Blankertz, 2010). The P300 has also been shown to increase with motivation in a BCI context (Kleih et al., 2010). The fact that spelling accuracy is optimized by continuously and explicitly adapting the stimulation to the user's need appears to create a virtuous cycle by boosting the user's motivation.

## 5 Acknowledgements

This work was carried out as part of the CoAdapt project, funded via ANR-09-EMER-002-01.

## References

- Jin J., Allison B.Z., Sellers E.W., Brunner C., Horki P., Wang X. and Neuper C. 2011 An adaptive P300-based control system. *J Neural Eng* **8**, 3:036006.
- Kleih S.C., Nijboer F., Halder S. and Kübler A. 2010 Motivation modulates the P300 amplitude during brain-computer interface use. *Clin Neurophysiol* **121**, 7:1023-31.
- Rivet B., Souloumiac A., Attina V. and Gibert G. 2009 xDAWN algorithm to enhance evoked potentials: Application to brain-computer interface. *IEEE Trans Biomed Eng* **56**, 8:2035-43.
- Serby H., Yom-Tov E. and Inbar G.F. 2005 An improved P300-based brain-computer interface. *IEEE Trans Neural Syst Rehabil Eng* **13**, 1:89-98.
- Townsend G., LaPallo B.K., Boulay C.B., Krusienski D.J., Frye G.E., Hauser C.K., Schwartz N.E., Vaughan T.M., Wolpaw J.R. and Sellers E.W. 2010 A novel P300-based brain-computer interface stimulus presentation paradigm: Moving beyond rows and columns. *Clin Neurophysiol* **121**, 7:1109-20.
- Treder M.S. and Blankertz B. 2010 (C)overt attention and visual speller design in an ERP-based brain-computer interface. *Behav Brain Funct* **6**, 1:28.
- Wolpaw J.R., Birbaumer N., Heetderks W.J., McFarland D.J., Peckham P.H., Schalk G., Donchin E., Quatrano L.A., Robinson C.J. and Vaughan T.M. 2000 Brain-computer interface technology: A review of the first international meeting. *IEEE Trans Rehabil Eng* **8**, 2:164-73.

# Extending Language Modeling to Improve Dynamic Data Collection in ERP-based Spellers

B. O. Mainsah, L. M. Collins, K. D. Morton and C. S. Throckmorton

Department of Electrical and Computer Engineering, Duke University, Durham, NC, U.S.A.  
leslie.collins@duke.edu

## Abstract

In this study, we extend the bigram language model to a higher order model in our dynamic stopping algorithm for the ERP-based P300 speller, with additional consideration to minimize erroneous character revisions. Prefix alternatives are generated to initialize the language model based on the likelihoods of being the target character. On-line results indicate there is potential to improve ERP speller performance with our proposed method, as statistically significant improvements were observed in participant communication rates.

## 1 Introduction

Language models can optimize ERP-based BCI speller performance by incorporating the probabilistic information about how letters are ordered in words when making character selection decisions [1–3]. We developed a dynamic stopping algorithm for the ERP-based P300 speller that uses a Bayesian approach to determine the amount of data collection based on a probabilistic level of confidence that a character is the target. The dynamic stopping algorithm with uniform initialization priors resulted in significant increase in bit rate from static data collection [4], with additional improvements observed with the inclusion of a bigram language model in the character initialization process [5]. In this study, we expand to a higher order model that uses all of the user’s spelling history, not just the previous character selection, with the hypothesis that it will improve the predictive capacity of the algorithm due to an increasing number of selected characters reducing the set of possible intended characters. However, the utility of a higher order model is dependent on the number of preceding characters that are selected correctly as these are used to obtain the correct target word prefix for initializing subsequent character probabilities. We consider a solution for misspellings by using information about character likelihoods post-data collection [6] to weight possible prefixes.

## 2 Methods

Participants ( $n = 20$ ) were recruited from the student and work population at Duke University, who gave informed consent prior to participating in the study. Participants performed word copy-spelling tasks online with a  $9 \times 8$  matrix speller grid using the Bayesian dynamic stopping algorithm (detailed in [5]) with different language models, with order counter-balanced.

Prior to the Bayesian update process, the language model is used to initialize character probabilities given a sequence of previously selected characters,  $\mathcal{A}_{T-n+1}^{T-1} = a_{T-n+1}, \dots, a_{T-1}$ :

$$P(C_{i,T} = C_T^*) = \alpha P(C_{i,T} | \mathcal{A}_{T-n+1}^{T-1}) \left( 1 - \sum_{NAC} \frac{1}{N} \right) + (1 - \alpha) \frac{1}{N} \quad (1)$$

where  $P(C_{i,T} = C_T^*)$  is the initialization probability of character  $C_i$  being the  $T^{th}$  character in the target word,  $C_T^*$ ;  $P(C_{i,T} | \mathcal{A}_{T-n+1}^{T-1})$  is the probability that the next character is  $C_i$

given the previously selected characters  $\mathcal{A}_{T-n+1}^{T-1}$ , which is based on the order of the language model;  $\alpha$  denotes the weight of the language model;  $1 - \alpha$  denotes the weight of a uniform distribution, which is an error factor to account for possible misspellings;  $\sum_{NAC} \frac{1}{N}$  is the sum of the non-alphabetic character (NAC) probabilities, which is subtracted from 1 to normalize the probabilities. The language model probabilities were derived from a corpus compiled by Norvig [7].

## 2.1 Bigram model

The conditional probability,  $P(C_{i,T}|\mathcal{A}_{T-n+1}^{T-1})$ , depends on the previously selected character,  $\mathcal{A}_{T-n+1}^{T-1} = a_{T-1}$  [5].

## 2.2 $n$ -gram model with dictionary-assisted prefix search (DAPS)

In the  $n$ -gram model, the conditional probability,  $P(C_{i,T}|\mathcal{A}_{T-n+1}^{T-1})$ , depends on all the previously spelled characters,  $\mathcal{A}_{T-n+1}^{T-1} = a_1, \dots, a_{T-1}$ . However, when using the  $n$ -gram model, there is the possibility of an erroneously selected character generating an invalid prefix (e.g. **VIS2** for the word **VISUAL**), or an incorrect valid prefix which can lead to the wrong initialization probabilities (e.g. **ANC** for the word **INCOME**), unless the erroneous character is revised. ERP-based P300 classifier confidences post-data collection can provide some information about the likelihood of being the target character [6]. We denote the P300 classifier confidences with a  $\mathbf{Q}^T = [Q_1, Q_2, \dots, Q_T]$  matrix, where  $Q_t = [q_t^1, \dots, q_t^N]^\top$  is a column vector of P300 classifier confidences post-data collection for all  $N$  grid characters for the  $t^{th}$  selected character. An example  $\mathbf{Q}$  matrix is shown in Figure 1. The user intended to spell the word **BEHIND** but the simulation yielded the word **BLGIND**. While the target character may not have the highest probability post-data collection, one of the next most probable characters usually is the target.

The  $\mathbf{Q}$  matrix can thus be used to select the  $k$  most likely prefixes from a dictionary to calculate initialization probabilities prior to spelling the  $T^{th}$  character. The set  $\mathcal{D}_T^k$  consists of valid prefixes in the dictionary with the top  $k$  values of the product of their character likelihoods ( $\prod_1^{T-1} q_t^{l(D_t^j)}$ ), and the prefixes are retained from one character to the next to generate the next  $k$  most likely prefixes. The conditional probability in the initialization step to select the  $T^{th}$  character thus also depends on the P300 classifier confidences via the  $\mathbf{Q}^{T-1}$  matrix:

$$w(D^j) = \frac{\prod_1^{T-1} q_t^{l(D_t^j)}}{\sum_{D^j \in \mathcal{D}_T^k} \left( \prod_1^{T-1} q_t^{l(D_t^j)} \right)} \quad (2)$$

$$P(C_{i,T}|\mathcal{A}_1^{T-1}, \mathbf{Q}^{T-1}) = \sum_{D^j \in \mathcal{D}_T^k} w(D^j) P(C_{i,T}|D^j) \quad (3)$$

where  $w(D^j)$  is the weight of the prefix  $D^j$ ;  $l(D_t^j)$  is the label for the  $t^{th}$  character in prefix  $D^j$ ;  $P(C_{i,T}|D^j)$  is the initialization probability that the  $T^{th}$  character is  $C_i$ , given prefix  $D^j$ .

## 3 Results

Figure 2 shows each participant's selection accuracy. For most participants, the accuracy from the bigram to  $n$ -gram model was similar or noticeably improved, ( $p < 0.09$ ). The  $\mathbf{Q}$  matrix tends to be sparse, meaning that only a few characters are considered likely to be the target.

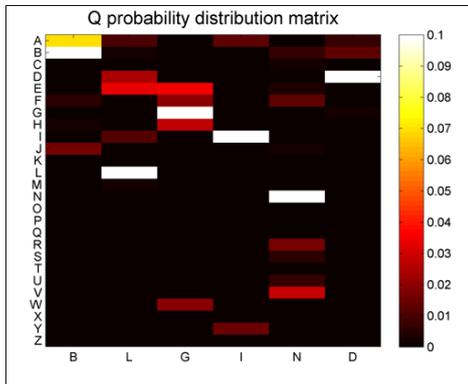


Figure 1: Example  $Q$  matrix post-data collection for the word **BEHIND**. The x-axis labels show the characters selected by the ERP-based P300 speller simulation, **BLGIND**, with the corresponding probabilities of alphabet characters at each character position,  $Q_t$ . Probability values are clipped for visualization purposes ( $P_{max} \geq 0.9$ )

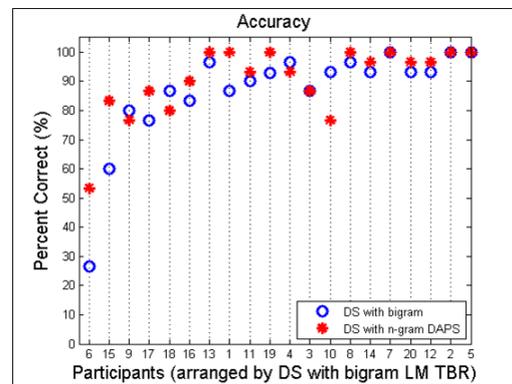


Figure 2: Participant performance comparison of online dynamic stopping between different language models showing accuracy of character selections

These characters typically have at some point been in the same flash group as the target, and this is often the source of most character selection errors. We hypothesize that the strong priors introduced by a higher order language model give the target character an added advantage to prevent an erroneous character selection, thereby sometimes improving accuracy.

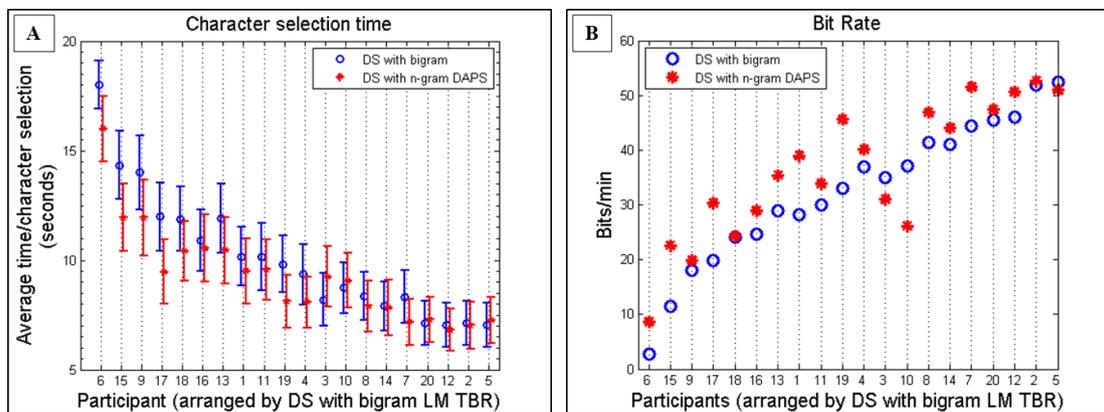


Figure 3: Participant performance comparison between on-line dynamic stopping with different language models. (A) Average character selection time, with standard error bars, (B) Bit Rate.

Figure 3A shows the average amount of time per character selection for each participant. A significant decrease was observed in character selection time with the  $n$ -gram DAPS model ( $p < 0.002$ ). In Mainsah *et al.* [5], off-line analysis of participant EEG data revealed that the rate of convergence to the threshold probability in dynamic data collection increased with the inclusion of the bigram language model. We hypothesize that the stronger priors introduced

by the higher order model further causes the character probabilities to converge faster.

The accuracy and average character selection time were used to calculate bit rate, including the time pauses between character selections [8]. Figure 3B shows participant bit rates with both algorithms. Due to similar accuracy levels and a significant reduction in character selection time, most participants observed significant improvement in their performance ( $p < 0.007$ ), with on average 26% increase in bit rate. Performance improvements with then-gram model are consistent with off-line analysis performed on EEG data from [5].

## 4 Discussion

The relatively slow communication rates of ERP-based BCI speller systems can be improved by exploiting the predictability of language. However, sometimes the manner of integration of language information in the ERP speller can lead to a decrease in performance due to increased task difficulty e.g. selecting from a drop-down menu in a predictive speller as in [2]. Our online results indicate there is potential to improve performance with a higher order language model in dynamic data collection, with additional consideration to minimize erroneous character revisions when used in combination with a dictionary. Further development includes adapting the algorithm for sentence spelling tasks, where word-space boundaries are important. There is the potential to further enhance performance using natural language processing tools such as word prediction and/or dictionary-based spelling correction. For example, the algorithm can be adapted to include likely word alternatives generated from the prefixes which can be displayed directly in the speller matrix [3], as this has been shown to not negatively affect performance.

## Acknowledgments

This research was supported by NIH/NIDCD grant number R33 DC010470-03.

## References

- [1] U. Orhan, D. Erdogmus, B. Roark, B. Oken, and M. Fried-Oken. Offline analysis of context contribution to erp-based typing bci performance. *Journal of Neural Engineering*, 10(6), 2013.
- [2] D. Ryan, G. Frye, G. Townsend, D. Berry, G. Mesa, N. Gates, and E. Sellers. Predictive spelling with a p300-based brain-computer interface: Increasing the rate of communication. *Int J Hum Comput Interact*, 27(1):69–84, 2011.
- [3] T. Kaufmann, S. Volker, L. Gunesch, and A. Kübler. Spelling is just a click away - a user-centered brain-computer interface including auto-calibration and predictive text entry. *Front Neurosci*, 6(72), 2012.
- [4] C. Throckmorton, K. Colwell, D. Ryan, E. Sellers, and L. Collins. Bayesian approach to dynamically controlling data collection in p300 spellers. *IEEE Trans Neural Syst Rehabil Eng*, 21(3):508–17, 2013.
- [5] B. Mainsah, K. Colwell, L. Collins, and C. Throckmorton. Utilizing a language model to improve online dynamic data collection in p300 spellers. *IEEE Trans Neural Syst Rehabil Eng*, 22(4), 2014.
- [6] B. Mainsah, K. Morton, L. Collins, and C. Throckmorton. Achieving spelling correction in p300 spellers without error-related potentials. *Society for Neuroscience Conference*, 2013.
- [7] Peter Norvig. How to write a spelling corrector. <http://norvig.com/spell-correct.html>, 2007.
- [8] D.J. McFarland, W.A. Sarnacki, and J.R. Wolpaw. Brain-computer interface (bci) operation: Optimizing information transfer rates. *Biological Psychology*, 63(3):237–251, 2003.

# Novel single trial movement classification based on temporal dynamics of EEG

M. Wairagkar<sup>1</sup>, I. Daly<sup>1</sup>, Y. Hayashi<sup>1</sup> and S.J. Nasuto<sup>1</sup>

<sup>1</sup> Brain Embodiment Lab, School of Systems Engineering, University of Reading, Reading, U.K.  
m.n.wairagkar@student.reading.ac.uk

## Abstract

Various complex oscillatory processes are involved in the generation of the motor command. The temporal dynamics of these processes were studied for movement detection from single trial electroencephalogram (EEG). Autocorrelation analysis was performed on the EEG signals to find robust markers of movement detection. The evolution of the autocorrelation function was characterised via the relaxation time of the autocorrelation by exponential curve fitting. It was observed that the decay constant of the exponential curve increased during movement, indicating that the autocorrelation function decays slowly during motor execution. Significant differences were observed between movement and no movement tasks. Additionally, a linear discriminant analysis (LDA) classifier was used to identify movement trials with a peak accuracy of 74%.

## 1 Introduction

Neural correlates of movement have been increasingly explored for applications in brain-computer interfacing (BCI) as they enable very intuitive control [1]. Previous studies have suggested the possibility of the involvement of various complex oscillatory processes in motor command generation [2]. Most research focuses on the spectral domain of the EEG for detecting movement [1]. This project takes a different approach on understanding the motor commands by studying the temporal dynamics of the EEG using novel features.

The principle of Event Related (De)synchronization (ERD/S) corresponding, respectively, to attenuation and increase predominantly in mu power and beta power [3], is widely used for detecting movements. Single trial analysis is important for online BCI implementation. Although these spectral features are able to reliably detect the motor command, they may not completely describe all aspects of motor command generation available in the EEG and do not indicate how one part of the EEG depends on another. Moreover, it is challenging to compute accurate instantaneous frequency distributions without compromising the temporal resolution and inducing delays in the motor command detection. Utilizing EEG signals' high temporal resolution, we have developed a novel method of detecting motor commands on a single trial basis by performing time domain analysis. Continuous autocorrelation analysis has been used for extracting temporal features from EEG.

In this study, different correlation based time domain analysis methods were explored for understanding the neural basis of motor command generation. Previous studies report that the first zero-crossing time, the time at which the autocorrelation function crosses 0, increases before and during voluntary movement [2]. Autocorrelation analysis was motivated by looking at temporal dependencies in the EEG. This approach was expanded by considering the evolution of the autocorrelation function over time and studying changes in relaxation time of the decay of the autocorrelation.

## 2 Methods

### 2.1 Experimental Paradigm

EEG was recorded from three participants. All the participants were males (2 right handed and 1 left handed) with ages 25, 23 and 29 years. An experimental paradigm was developed for recording self-paced index finger tapping of the right and left hand using tools from the BioSig toolbox [4]. A fixation cross was displayed on the screen placed at eye level for 2 sec at the beginning of each trial and followed by a textual cue for right or left hand finger tapping or resting. Participants were asked to perform a self-paced single finger tap at a random time of their choice within the 10 sec window following the cue. Each trial was followed by a random break of 1 to 1.5 sec. The experiment was broken down into separate runs of 12 trials with 4 cues per class displayed in random order to avoid pattern learning by the participants. The experimental setup is illustrated in Figure 1.

A bespoke tapping device was developed using a programmable microcontroller to record the tapping signals from both the fingers. In order to mark the exact onset of the movement in EEG, both EEG and finger tapping signals were recorded simultaneously and co-registered using tools developed as part of the TOBI framework [5]. EEG from 19 electrodes (impedances kept below 8k $\Omega$ ) was recorded using a Deymed TruScan amplifier with a sampling frequency of 1024 Hz. Forty trials for each of the three conditions were recorded for each participant.

### 2.2 EEG pre-processing and Artifacts removal

Signal pre-processing was done using a fourth-order Butterworth filter. DC offset in the signal was removed using a high-pass filter with a cut-off frequency of 0.5Hz. Power line noise was filtered using a notch filter at 50Hz. Finally, high frequency noise was eliminated using a low-pass filter with a cut-off frequency of 60Hz.

Independent Component Analysis [6] was used to remove artefacts from the recorded signals. Independent components (ICs) with artefacts were identified manually. Artefact-free EEG was reconstructed by eliminating these ICs. EEG was then segmented into individual trials. Trials of length 6 sec were obtained by extracting 3 sec before and 3 sec after the onset of movement.

### 2.3 Autocorrelation analysis based on exponential decay

In order to examine the time development of the relaxation time of brain activity before, during, and after the movement, the autocorrelation function was calculated to extract the relaxation. The autocorrelation function shows the degrees of un-correlation as a function of time from initial state.

For a given signal  $A(t)$ , the auto-correlation is defined by  $C(\Delta t) = \langle A(t)A(t-\Delta t) \rangle$ , where  $\langle \dots \rangle$  represents the average over time. At the initial time,  $C(0) = \langle A^2 \rangle$ , and after infinite time, the signal is completely uncorrelated, giving  $C(\infty) = \langle A \rangle^2$ . How the signal becomes uncorrelated as a function of time may be described by  $C(t) = \langle A^2 \rangle e^{(-t/\tau)}$  to describe the general trend of the relaxation process when the average of the signal  $\langle A \rangle = 0$ . If the auto-correlation is normalized,  $C(t) = e^{(-t/\tau)}$  where  $\tau$  represents the relaxation time of the signal and is an indicator of the relaxation process.

Autocorrelation functions were derived for the 30Hz low-pass filtered EEG. A windowing approach was used for determining instantaneous autocorrelation. Windows of length 1sec and shifted by 100ms were extracted. Normalised continuous autocorrelation was performed on each window at all lags with non-zero values.

The exponential curve  $y=K.e^{(-t/\tau)}$  was fitted to the local maxima of the positive lags of the autocorrelation function obtained from each window of the trial and the decay constant  $\tau$  of the fitted curve was extracted as a feature (see Figure 2). The constant  $K$  was set to 1. The  $\tau$  values for all the windows for each trial were plotted (see Figure 3).



Figure 1: Experimental Setup

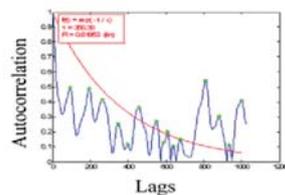


Figure 2: Exponential Curve  
The fitting represents autocorrelation relaxation for right tap trial on C3

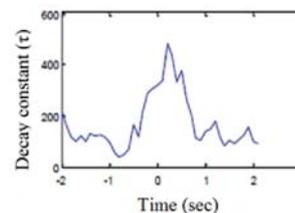


Figure 3: Plot of changes in  $\tau$  in a single right tap trial on C3. Time is reported relative to movement onset.

### 2.4 EEG analysis and classification

The 9 EEG channels around the motor cortex (F3, Fz, F4, C3, Cz, C4, P3, Pz, and P4) were analysed. Before beginning further analysis and classification with the novel time domain method, to assess the quality of data, EEG was validated for the presence of ERD using event-related spectral perturbation. To observe ERD, average spectrograms of resting state trials were subtracted from average spectrograms of movement trials. Figure 4 shows the decrease in mu power around movement onset.

To analyse the results obtained by plotting  $\tau$  for single trials, element-wise 2 sample *t*-test were performed to identify statistically significant differences between right tap vs. rest and left tap vs. rest on 9 EEG channels. A Linear discriminant analysis (LDA) classifier was used for classification. LDA was applied in a sliding window (length 1s, step size 0.1s). A 10x10 cross-fold validation scheme was used with binary classification of right/left tap vs. no tap and the best channel was selected manually.

## 3 Results

Increases in the value of  $\tau$  around the onset of movement were clearly observed in most trials. The  $\tau$  values of the resting state trials appeared stable throughout the trial. Features around the onset of the movement showed statistically significant differences between tap vs. rest conditions (see Figure 5). The most responsive channels for right and left hand tap differed between participants. Using the autocorrelation function decay constant movement could be detected from single trials.

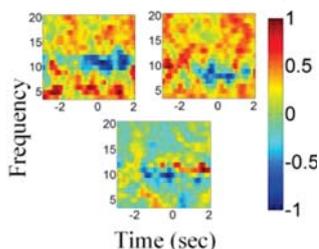


Figure 4: ERD on channel C3 for participant 1, 2 and 3

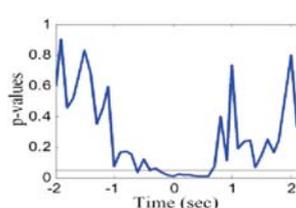


Figure 5: *t*-Test for left hand trials, participant 3 in Fz. The horizontal line indicates statistical significance ( $p < 0.05$ ).

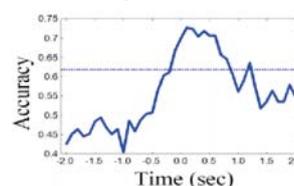


Figure 6: Classification accuracy for left hand, participant 3 in Fz. The horizontal line indicates statistical significance ( $p < 0.05$ ).

LDA classification accuracies for all the participants were plotted for the classification of movement of right vs. rest and left vs. rest. The accuracy obtained was considered statistically significant at  $p < 0.05$ . A peak accuracy of 74% was achieved for participant 3 (shown in Figure 6).

Table 1 shows classification accuracies for all the participants. Significant accuracies were obtained for all the participants, except for participant 1 right hand tapping condition. It's interesting to note that participant 3, who exhibited the smallest ERD response, gave the best results using this method.

Participant	Right hand tapping classification accuracy (%)	Left hand tapping classification accuracy (%)
1	58.0	<b>68.4</b>
2	<b>66.0</b>	<b>69.0</b>
3	<b>68.0</b>	<b>74.0</b>

**Table 1:** Movement classification accuracies. Statistical significance ( $p < 0.05$ ) is indicated in bold.

## 4 Conclusions and future work

A novel approach to extract features from the temporal dynamics of brain oscillations on a single trial basis was used to study the neural mechanisms of movement. This time domain single trial analysis has great potential for online BCI. Oscillations of a wide frequency range were taken into account without limiting the feature search into pre-determined frequency bands. This has led to the novel discovery of the behaviour of the autocorrelation function during voluntary movement. The autocorrelation function decays slower during movement as compared to rest. When there is no movement, decreases in the autocorrelation function are sharp. This suggests that during rest, the oscillatory processes and relaxation process of the autocorrelation function are distinct. However, during movement, coupling occurs between relaxation and oscillatory processes. Thus, the relaxation time of autocorrelation is a measure of temporal dependency in EEG.

Since the study performed was very novel, initial analysis was done on only three participants to validate the proposed hypothesis. There is large scope for further work. To validate and confirm the robustness of this method, EEG analysis will be done on more participants and a comparison will be made to ERD based classification of movement. The system will be adapted for use in online BCI.

## References

- [1] H. Hwang, et al., "EEG-based Brain-Computer Interfaces (BCIs): A Thorough Literature Survey," *International Journal of Human-Computer Interaction*, 29(12), 2013.
- [2] Y. Hayashi, et al., "Analysis of EEG Signal to Detect Motor Command Generation Towards Stroke Rehabilitation," in *Converging Clinical and Engineering Research on Neurorehabilitation*, Berlin Heidelberg, 2013.
- [3] G. Pfurtscheller and F. Lopes da Silva, "Event-related EEG/MEG synchronization and desynchronization: basic principles," *Clin. Neurophysiol.*, 110(11) pp. 1842-1857, 1999.
- [4] "BioSig for Matlab and Octave – an open source software library for biomedical signal processing," in *Available online: <http://biosig.sf.net>*, 2005.
- [5] C. Breitwieser, I. Daly, C. Neuper and G. Muller-Putz, "Proposing a Standardized Protocol for Raw Biosignal Transmission," *{IEEE} Trans. Biomed. Eng.*, 59(3), pp. 852-859, 2012.
- [6] T. Jung, S. Makeig, C. Humphries and T. Lee, "Removing electroencephalographic artifacts by blind source separation," *Psychophysiology*, 37, p. 163–178, 2000.

# A BNCI-based technology for cognitive rehabilitation after stroke: survey on usability

Monica Risetti<sup>1</sup>, Sonja C. Kleih<sup>2</sup>, Silvia Erika Kober<sup>3</sup>, Cristina Campillo<sup>4</sup>, Jlenia Toppi<sup>1,5</sup>, Laura Astolfi<sup>1,5</sup>, Claus Vogele<sup>4</sup>, Guilherme Wood<sup>3</sup>, Andrea Kübler<sup>2</sup> and Donatella Mattia<sup>1,2</sup>

<sup>1</sup>Fondazione Santa Lucia, IRCCS, Rome, Italy

<sup>2</sup>University of Würzburg, Germany

<sup>3</sup>University of Graz, Austria

<sup>4</sup>University of Luxembourg, Luxembourg

<sup>5</sup>Dept. of Computer, Control and Management Engineering - "La Sapienza" University of Rome  
m.risettisantalucia.it

## Abstract

The treatment of cognitive impairment and prevention of further decline are essential aspects of stroke rehabilitation.

A new Brain Neural Computer Interface (BNCI) system was developed adopting user-centered design (UCD) principles, to operate bio- and neuro-feedback based cognitive training modules with the aim of improving post-stroke cognitive rehabilitation outcomes. The usability of such BNCI-based system for cognitive training was tested both with hospitalized patients (first testing phase) and those at home (second testing phase). Questionnaires were administered to 15 experts in rehabilitation, 7 stroke patients and 3 caregivers. Users' needs and requirements have been, and will further be collected, during all the experimental phases to integrate and refine the system according to the iterative cycling of the UCD.

The preliminary findings of this survey indicated that the newly proposed BNCI-assisted training for cognitive rehabilitation was well accepted by the majority of stroke patients (86%) and professionals (70%). The level of usability and acceptability revealed by the survey is encouraging for the system translation in clinical routine usage upon the training proves effective for improving cognition.

## 1 Introduction

Cognitive deficits occur in the majority of stroke patients (Haring, 2002) and have a high impact on their quality of life and that of their families (Carod-Artal et al., 2009). Currently, neuropsychological rehabilitation lacks in intensity and duration of specific rehabilitation strategies. As yet, patients do not have the opportunity to continue the rehabilitation endeavor after discharge, and they are left with no options for a professional training and monitoring of outcomes at home. The

---

1 Masterminded EasyChair and created the first stable version of this document

2 Created the first draft of this document

EU funded project CONTRAST ([www.contrast-project.eu](http://www.contrast-project.eu)) is currently deploying a Brain Neural Computer Interface (BNCI)-based technology to provide cognitive training modules to improve cognitive rehabilitation outcomes in institutionalized patients, and also to support patient's training at home by remote controlled supervision. The term BNCI denotes for the implementation within the CONTRAST system of several training modules for attention, declarative memory, working memory and executive functions that are operated by biometric signals such as heartbeat (HRV), Electrooculogram (EOG) and Electromyography (EMG) and the Electroencephalogram (EEG). Based on the UCD principles (Maguire, 1998), we iteratively involved different users such as professionals, patients (in the hospital and at home) and caregivers to evaluate the system *usability*. Here we report the preliminary results of the survey based on questionnaire data collected in the rehabilitation facilities.

## 2 Materials and Methods

### 2.1 Assessment of system usability

At the current stage of the CONTRAST project, 15 professionals (psychologists, neurologists, neuropsychologists and therapists from several EU countries), 7 stroke patients (admitted to the neurorehabilitation department for standard rehabilitation treatment and the experimental intervention) and 3 caregivers have been involved to evaluate the usability of the training module for declarative memory. Such module relies on EEG sensorimotor rhythm (SMR) self-modulation. The training consists of a 3-minute baseline trial during which the participants are instructed to relax followed by a total of six 3-minute feedback runs where patients instructions are to increase their SMR amplitude (fed back by means of a bar placed in the middle of a screen whose size has to be kept above a predefined threshold) while reducing EOG and EMG artifacts (fed back by means of 2 bars on the left and the right side of the screen whose size has to be kept below predefined thresholds), respectively. Successful trials are also rewarded by means of visual and auditory feedback (point score and a Midi Tone). The EEG signals are recorded for then Cz electrode (2 mastoid electrodes for reference and ground). The Blood Volume Pulse (BVP) and the heart rate (HR) are monitored by means of a finger clip sensor. All modules have been implemented within the NEXUS technology (NeXus-10 MKII, Mind Media BV). Six out of 7 involved patients underwent all the training protocol (10 SMR neurofeedback training sessions) and 4 patients were also able to linearly increase their SMR amplitude over the neurofeedback training runs.

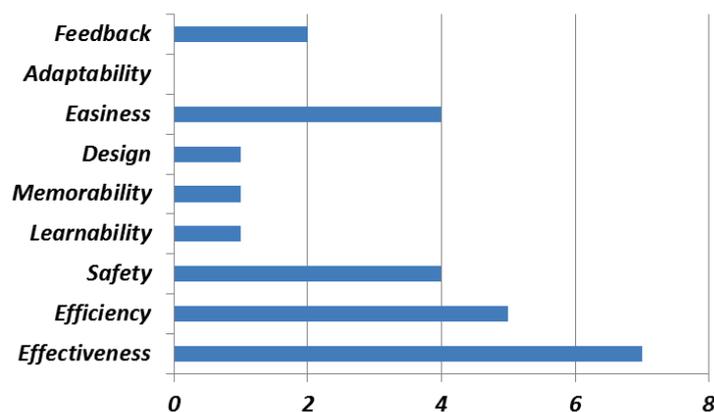
The adopted questionnaire was derived from the results of open interviews previously collected from a group of experts in the rehabilitation fields from different EU countries, and adapted to each category of users. The following main domains (a total of 70 questions related to the two different settings) were investigated: product use, intensity of the training, feedback for professional and patients, market feasibility, individualization, design and easiness to use, safety, and virtual reality. Finally, the professionals were asked to rank the aspects of the technology they considered most important. A shorter form of this questionnaire was then administered to patients and caregivers and it did not deal with cognitive functions, market feasibility, individualization and virtual reality fields whereas it provided questions related to standard rehabilitation program and the needs for interaction with the therapist. All items of the questionnaire required a qualitative response, thus the presented results are only descriptive.

### 3 Results

#### 3.1 Questionnaires for professional users

With regard to the BNCI-based system usage, the majority of the 15 interviewed experts (73%) reported it might be used in the facilities only after an initial period of conventional cognitive therapy. For the home-based setting, 70% of the professionals rated the independent system usage (by the patient and/or the caregiver) as the most relevant aspect. The training session duration of 30 minutes was indicated as preferred by 90% of those interviewed. The auditory feedback, the performance score and the smiley face for positive reinforcement were ranked by professionals as the most appropriate type of feedback for patients, while they wished to have fed back the performance and learning curves to estimate patients' progress.

The need for an easy-access was indicated by 45% of those interviewed who indicated that several solutions to access the devices are necessary and should be provided to patients in order to facilitate the basic operation of the system (i.e., select and initiation of the required training module; save data...). With regards to the safety of the BNCI-based technology, the system needed to be certified as medical product for 73% of the professionals. Finally, the professionals were asked to rank the several aspects of system with the indication, if they had considered one issue as important as another, to indicate by giving them the same number. Four professionals had the opinion that all areas were equally important. As illustrated in Figure 1, the most relevant aspects were effectiveness, efficiency, safety and easiness of use.



**Figure:** The histogram illustrates the BNCI-based training system related items (x axis) as ranked by the professional users (y axis = the number of preferences obtained for each item).

#### 3.2 Questionnaires for patients end caregivers

Despite 24% of the patients not being satisfied with the usual cognitive rehabilitative care, most of them (86%) would like to use the system in both hospital and home setting. Patients would be willing to spend money for training modules. As for the home-based usage, 57% of the interviewed patients felt confident to independently use the system at home only after using it during their stay in hospital

whereas 43% of them thought they would need help from family members (60%) or therapists (40%). Most of the patients requested a therapist to be present at least once a week (70%) or 2 times per month (30%). Intensity and duration of the intervention was rated as similar to professionals (average 3 times per week with each session duration of 30 minutes; 57%). The maintenance of a high level of patient's motivation would require to be in contact with the therapist (43%) and/or to provide user-friendly feedbacks (29%). Effectiveness and safety were ranked as the most important areas, followed by efficiency, easiness of use and feedback. Most family members (3 families were interviewed) would like to continue their relatives' rehabilitation at home. They could help patients in their rehabilitation for about 2 hours a day if necessary, but they would like to be trained up to 2 hours to learn the system use. They also would like to have contact with the therapist 2-3 times (67%) or 4-5 times (33%) per month. All family members indicated as potential limits the costs and the difficulties to use the system.

## 4 Conclusion and outlook

The preliminary findings of this survey indicated that the proposed BNCI-based approach to cognitive rehabilitation was well accepted by the majority of the stroke patients and professionals. The feedbacks obtained from both classes of users are being currently incorporated in the design of the interface between patients and therapists, as more intuitive icons provide a direct access to basic operations such as selecting and running the desired training module, automatic data saving, the possibility for the therapist to control the session (whenever is needed) as soon as the patients switch on the system. The level of usability and acceptability revealed by the survey is the basis for the system application in clinical routine, provided the training proves effective to enhance cognition. The approach is also timely as remote patient's supervision will become more relevant in the future, due to increasing population mean age (i.e. elderly) and the life expectancy of those who have to live with the burden of chronic disease.

### ACKNOWLEDGMENTS

This work was supported by the European ICT Programme Project FP7-2011-7 (CONTRAST). This paper only reflects the authors' views and funding agencies are not liable for any use that may be made of the information contained herein.

## References

- Carod-Artad, FJ and Egidio, JA. (2009). Quality of life after stroke: the importance of a good recovery. *Cerebrovasc Dis*; 27(1):204-214. doi: 10.1159/000200461.
- Cicerone, KD, Langenbahn, DM, Braden, C, Malec, JF, Kalmar, K, Fraas, M, Felicetti, T, Laatsch, L, Harley, JP, Bergquist, T, Azulay, J, Cantor, J and Ashman, T. (2011). Evidence-based cognitive rehabilitation: updated review of the literature from 2003 through 2008. *Arch Phys Med Rehabil*; 92(4): 519-530.
- Haring, HP (2002). Cognitive impairment after stroke. *Curr Opin Neurol*; 15: 79-84. Maguire, MC (1998). User-centered Requirements Handbook. H.R. Institute, Editor. *RESPECT Consortium*.

# How Well Can We Learn With Standard BCI Training Approaches? A Pilot Study.

Camille Jeunet<sup>1</sup>, Alison Cellard<sup>2</sup>, Sriram Subramanian<sup>3</sup>, Martin Hachet<sup>2</sup>,  
Bernard N’Kaoua<sup>1</sup> and Fabien Lotte<sup>2</sup>

<sup>1</sup> University of Bordeaux, Bordeaux, France

<sup>2</sup> Inria Bordeaux Sud-Ouest, Talence, France

<sup>3</sup> University of Bristol, Bristol, England

camille.jeunet@inria.fr

## Abstract

While being very promising, brain-computer interfaces (BCI) remain barely used outside laboratories because they are not reliable enough. It has been suggested that current training approaches may be partly responsible for the poor reliability of BCIs as they do not satisfy recommendations from psychology and are thus inadequate [3]. To determine to which extent such BCI training approaches (i.e., feedback and training tasks) are suitable to learn a skill, we used them in another context (without a BCI) to train 20 users to perform simple motor tasks. While such approaches enabled learning for most subjects, results also showed that 15% of them were unable to learn these simple motor tasks, which is close to the BCI illiteracy rate [1]. This further suggests that current BCI training approaches may be an important factor of illiteracy, thus deserving more attention.

## 1 Introduction

Brain-computer interfaces (BCIs) are communication systems allowing users to interact with the environment, using only their brain activity [6]. BCIs, although very promising, remain barely used outside laboratories because they are not reliable enough [6]. Two main reasons have been identified. The first one, extensively investigated, concerns brain signal processing, with current classification algorithms being still imperfect [1]. The second one concerns the users themselves. Indeed, many users seem unable to acquire good BCI skills (i.e. the capacity to generate specific and stable brain activity patterns): around 20% cannot control a BCI at all (the so-called “BCI illiteracy”), while most of the remaining 80% have relatively modest performances [1]. An appropriate training is needed to acquire these skills, especially for Mental Imagery-based BCI (MI-BCI). It has been suggested that currently used training and feedback protocols, which do not take into account recommendations from psychology to optimise human learning, might be partly responsible for BCI illiteracy and poor user performance [3]. For instance, it has been shown that, for efficient learning, training protocols have to fit the user learning style and propose an increasing and adaptive difficulty [3]. Yet standard BCI training protocols are the same for all users [3]. While instructive, these studies only provide theoretical considerations about training approaches. It is therefore necessary to concretely assess whether training approaches used in BCI are appropriate to train a skill. Moreover, it is necessary to perform this evaluation independently of BCI, to rule out possible biases due to BCI complexity, non-stationarity and poor signal-to-noise ratio. Thus in this work, we propose to study these BCI training approaches without using a BCI: participants were asked to learn specific and simple motor tasks using the same feedback and training tasks used for MI-BCI. We then studied how well they could learn such motor tasks to assess the quality of the training approaches, independently of BCI use. We studied here two different approaches: 1) the training approach

used in “standard” MI-BCI [5] and 2) a variant of it which provides some autonomy to the user. Indeed, with the “standard” approach, no autonomy is given to the user, who always has to perform the tasks required by the protocol. Yet, autonomy is known to increase motivation and learning efficiency in general [3]. Interestingly enough, the study described in [4] obtained promising results when providing more autonomy to a single BCI user.

## 2 Methods

Participants were asked to learn to perform two motor tasks: drawing triangles and circles with a pen on a graphic tablet (see Figure 1(b)), using standard MI-BCI training approaches [5]. Indeed, as with MI-BCI, in which users have to learn a suitable movement imagination strategy, the participants here had to learn the strategy which allows the computer to correctly recognise their drawing, e.g., they had to identify the suitable shape size, angles or speed of drawing.

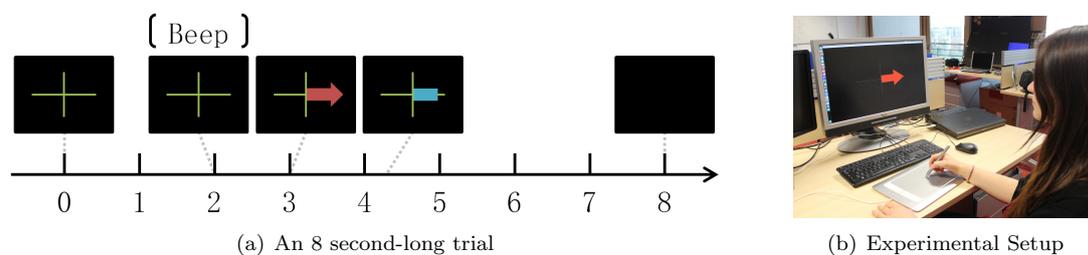


Figure 1: (a) Outline of a trial from a standard run; (b) Experimental setup.

### 2.1 Experimental protocol

Participants had to learn to draw circles and triangles that can be recognised by the computer during different runs, which were either standard (s) or self-paced (sp). S-runs were composed of 20 trials per task. As shown in Figure 1(a)), at the beginning of each trial a green cross was displayed. After 2s, an auditory cue (a beep) announced the beginning of the task. Then, after 3s, a red arrow was displayed, indicating which task the participant had to perform: continuously drawing circles or triangles upon appearance of a left or right arrow, respectively. After 4.25s, a blue feedback bar appeared and was updated continuously for 4s. Its direction indicated the shape recognized by the classifier (left: circle, right: triangle) and its length was proportional to the classifier output (i.e., the distance to the classifier separating hyperplane), as with MI-BCI. During sp-runs, no instructions were given: the participants were asked to do the motor tasks in an autonomous and free way. Half of the participants were asked to learn using a Standard (S) training approach: they did 4 seven-minute-long s-runs. The other half learned using a training approach with increased autonomy, denoted Partially Self-Paced (PSP) approach: the 1<sup>st</sup> and 4<sup>th</sup> runs were s-runs, while the 2<sup>nd</sup> run was replaced by a 3.5 minute long sp-run followed by a shortened s-run (10 trials per task, 3.5 minutes), and the 3<sup>rd</sup> run was replaced by a shortened s-run followed by a 3.5 minute long sp-run. The training duration was the same in both conditions. We studied the impact of the condition, S vs. PSP, on the recognition accuracy of triangles and circles over runs (i.e., learning effects) and on subjective experience (using a questionnaire). 20 participants (10 per group) took part in our experiment.

## 2.2 Signal Processing

In order to discriminate triangle from circle pen gestures, we used a pattern recognition approach as in BCI. From the past 1s-long time window (in a sliding window scheme, 937.5ms overlap) of the 2D pen position (16Hz sampling rate), a histogram of angles was computed. More precisely, the angles between each consecutive segment of the time window were first computed. Then the number of angles falling in the ranges 0-30°, 30-75°, 75-105°, 105-150° and 150-180° were counted, and these 5 count values were used as input features for a Linear Discriminant Analysis (LDA) classifier. The (subject-independent) LDA classifier was trained on 60 trials from each gesture, from 2 persons (1 left-handed, 1 right-handed). The resulting classifier could discriminate triangles from circles with 73.8% classification accuracy (10-fold cross-validation on the training set), which is an accuracy equivalent to the average accuracy of a MI-BCI [2]. Classification accuracy was measured as the average number of 1s-long time windows correctly classified during the feedback period from each trial (see Figure 1(a)).

## 2.3 Analyses

In order to analyse the interaction between the “Condition” (2 modalities: S and PSP; independent measures) and the performance obtained at each “Run” (4 modalities: run1, run2, run3 and run4; repeated measures), we performed a 2-way ANOVA. Moreover, we asked the participants to complete a Usability Questionnaire (UQ) which measured 4 dimensions: learnability/memorability (LM), efficiency/effectiveness (EE), safety (Saf.) and satisfaction (Sat.). Thus, we did a two-way ANOVA to analyse the interaction between the “Condition” and the “Evaluated Dimension” (4 modalities: LM, EE, saf. and sat.; repeated measures).

## 3 Results & Discussion

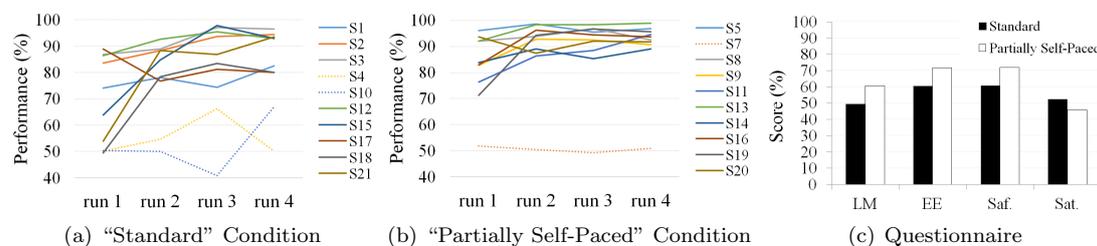


Figure 2: (a) Performance (pen gesture classification accuracy) over runs in the S Condition; (b) Performance over runs in the PSP Condition; (c) Average UQ scores.

Irrespectively of the condition, classification accuracy results (see Figures 2(a) and 2(b)) showed that 17 out of 20 participants managed to learn the task (best run accuracy  $\geq 82.6\%$ ), with accuracies increasing with the number of runs on average, as can be observed in BCI. However, 3 participants did not manage to learn the tasks (best run accuracy  $< 70\%$ ). This is particularly interesting considering the simplicity of the motor tasks which ensured the users could technically perform them. Such training approaches thus seem suboptimal. Moreover, this rate of 15% of people who did not manage to learn is close to the BCI-illiteracy rate (20% [1]). Overall, this suggests that BCI illiteracy may not be due to the user only, but also substantially to the training protocol. Then, results showed neither a main effect of the

Condition [ $F(1,18)=2.33$ ;  $p=0.15$ ] nor a Condition X Run interaction [ $F(3,45)=1.35$ ;  $p=0.27$ ] when considering all the participants. However, when the 3 illiterates are excluded, the 2-way ANOVA showed a main effect of the Condition [ $F(1,15)=7.48$ ;  $p=0.01$ ]: the PSP group seemed to perform better than the S group (meanPSP= $91.00 \pm 3.82$ , meanS= $83.29 \pm 6.95$ ). However, random sampling of the participants led by chance to a PSP group with classification accuracies for the first run that are higher than that of the S group, which prevents us from drawing any relevant conclusion on comparative learning effects.

UQ results (see Figure 2(c)) showed no main effect of the Condition [ $F(1,18)=0.98$ ;  $p=0.33$ ]. However, they showed a trend towards a Condition X Evaluated Dimension interaction [ $F(3,54)=2.40$ ;  $p=0.077$ ], which is due to the better evaluation of LM, EE and Saf. in the PSP condition than in the S condition, which is not the case for the Sat. dimension. These results suggested that while the PSP approach is not more pleasant to learn with than the S approach, it is easier. Interestingly enough, 8 subjects reported in an open-question of the questionnaire that the feedback was very uninformative, which made learning the tasks difficult.

## 4 Conclusion

This study aimed to concretely assess how well one could learn a given skill with BCI training approaches. To do so, we proposed to study BCI training approaches without using BCI, i.e., we used feedback and training tasks from MI-BCI to train participants to draw triangles and circles (i.e., simple motor tasks) so these can be recognized by the computer. Half of the participants did so using a S training approach while the other half used a PSP one. In terms of learning effects, results unfortunately showed no relevant differences between conditions (S vs. PSP), due to initial performances that differed between conditions, by chance. However, irrespectively of the condition, 15% of the participants (3 out of 20) seemed unable to learn the motor tasks, despite their simplicity. This suggests that such training approaches are not optimal for learning and thus that they may be an important factor of BCI illiteracy. Concerning user experience, UQ showed a tendency towards a better feeling of learnability/memorability, efficiency/effectiveness and safety with the PSP than with the S approach, while the satisfaction appeared similar for both. Overall, this study confirmed in practice the theoretical analyses from [3] suggesting that current BCI training approaches were suboptimal and need to be changed. In the future, we will increase the number of participants, explore the PSP approach with actual BCIs, and propose new training approaches that consider the user's cognitive style and motivational states to improve both the learning experience and performance.

## References

- [1] B.Z. Allison and C. Neuper. Could anyone use a BCI? In Desney S. Tan and Anton Nijholt, editors, *Brain-Computer Interfaces*, pages 35–54. Springer London, 2010.
- [2] B Blankertz, C Sannelli, S Halder, EM Hammer, A Kübler, K-R Müller, G Curio, and T Dickhaus. Neurophysiological predictor of SMR-based BCI performance. *NeuroImage*, 51(4):1303–1309, 2010.
- [3] F. Lotte, F. Larrue, and C. Mühl. Flaws in current human training protocols for spontaneous BCI: lessons learned from instructional design. *Frontiers in Human Neurosciences*, 7:568, 2013.
- [4] C. Neuper, G. Müller, A. Kübler, N. Birbaumer, and G. Pfurtscheller. Clinical application of an EEG-based BCI. *Clinical Neurophysiology*, 114:399409, 2003.
- [5] G. Pfurtscheller and C. Neuper. Motor imagery and direct brain-computer communication. *proceedings of the IEEE*, 89(7):1123–1134, 2001.
- [6] J. R. Wolpaw and E. W. Wolpaw. *BCI: Principles and Practice*. Oxford University Press, 2012.

# Analyzing EEG Source Connectivity with SCoT

Martin Billinger<sup>1</sup>, Clemens Brunner<sup>1</sup> and Gernot R. Müller-Putz<sup>1</sup>

Institute for Knowledge Discovery, Graz University of Technology, Graz, Austria  
 bci@tugraz.at

## Abstract

In this article we demonstrate how to use SCoT<sup>1</sup>, our source connectivity toolbox, to estimate connectivity on motor imagery data. We show both, multi- and single-trial analysis examples. The latter can be useful for feature extraction in brain-computer interfaces if reasonable regularization constraints are applied.

## 1 Introduction

Quantifying interactions in dynamic large-scale brain networks is an important and useful tool in neuroscience. The source connectivity toolbox (SCoT<sup>1</sup>) [3] is a Python package for estimating spectral effective connectivity between brain sources. SCoT extracts connectivity measures from vector autoregressive (VAR) models fitted to source signals. Typically, the sources are obtained by performing MVARICA [5] or CSPVARICA [3], which are based on independent component analysis (ICA) decomposition of VAR residuals.

Several brain-computer interface (BCI)-related studies have included connectivity features for classification [4, 2]. Although the tools in SCoT were originally designed for single-trial BCI feature extraction, they also support multi-trial data and are useful for functional and effective connectivity analysis of electroencephalogram (EEG) signals.

In this article, we demonstrate several ways to use SCoT on a motor imagery (MI) data set.

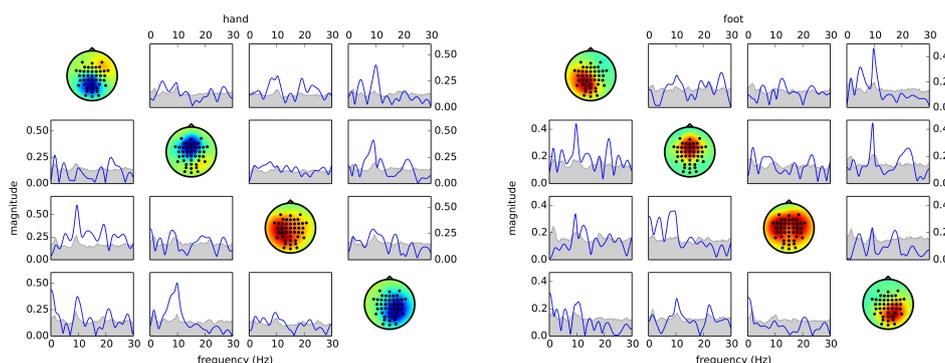


Figure 1: Multi-trial joint source/connectivity estimation. Left: hand MI, right: foot MI. Each plot shows the PDC spectrum of a source pair. The area shaded in gray is the one-sided 95 % confidence interval of the PDC under the null-hypothesis of no connectivity.

<sup>1</sup><https://github.com/SCoT-dev/SCoT>

## 2 Materials and Methods

**SCoT Workflow** SCoT provides routines for EEG source and connectivity estimation. Two estimation approaches are possible in SCoT. One approach is joint estimation, where sources and connectivity are estimated together. This is performed by applying MVARICA or CSPVARICA to data where sources can be assumed to be spatially and temporally stationary. Alternatively, a two-step approach estimates sources and connectivity separately, possibly on different data sets. MVARICA or CSPVARICA can be employed in the source decomposition step by discarding their VAR estimates. In the second step, the unmixing matrix is used to obtain source activations. Connectivity measures are estimated from VAR models fitted to these source activations.

**Example Data Set** An example data set is available with SCoT. This data set contains a recording of 45 EEG channels from one healthy subject performing hand and foot MI. A total of 180 trials (90 trials per MI task) were recorded. In each trial the subject was cued to perform either MI task by an arrow pointing up (hand) or down (foot). The beginning of a trial was indicated by a fixation symbol appearing on the screen, and the motor imagery period started with the cue after 2.5 s. The motor imagery period was 4.5 s long and was followed by a 2.5–3.5 s break.

**Usage Examples** In this article, we demonstrate how to use SCoT to estimate and visualize connectivity on the example data set. We show how to perform joint estimation, multi-trial two-stage estimation, single-trial connectivity, and circular plots. In each case we measure connectivity with the PDC [1] in a 1 s window starting 2 s after the cue. For clarity of demonstration, we only use four sources in the first three examples and seven sources for the circular plots. We manually removed sources that were clearly related to artifacts (such as eye movement or neck muscle activity).

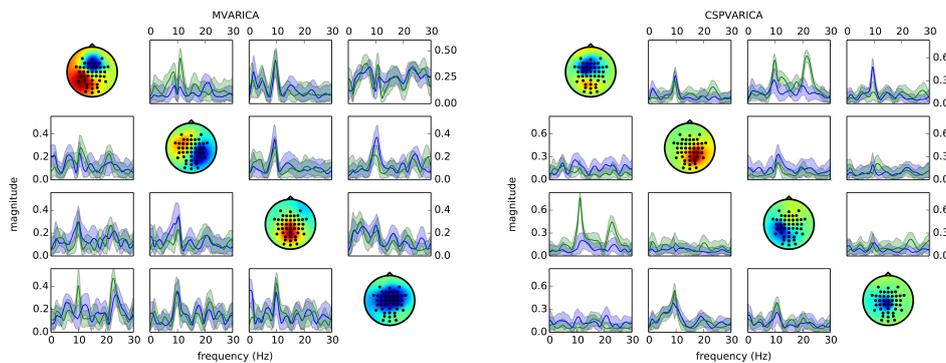


Figure 2: Multi-Trial Two-Stage Estimation. Left: MVARICA, right: CSPVARICA. Each plot shows PDC spectra of a source pair for hand MI (blue) and foot MI (green). The shaded areas correspond to the 95 % confidence intervals obtained by bootstrapping.

**Interpreting the Results** We show different spectral connectivity plots in Figures 1, 2, and 4, which are discussed in the results section below. These plots are arranged so that columns correspond to the origin, rows to the destination of connectivity, and source topographies are

located along the diagonal. A high connectivity value in a plot generally means that there is causal interaction from the source in the same column to the source in the same row, at a certain frequency.

### 3 Results

**Multi-Trial Joint Source/Connectivity Estimation** Here, we performed joint estimation of sources and connectivity on all trials of each class separately. Figure 1 shows that slightly different sources are obtained for each class, which makes it difficult to evaluate class differences in connectivity.

**Multi-Trial Two-Stage Estimation** The idea of two-stage estimation is to re-use the source decomposition on different data sets, which allows us to evaluate changes in connectivity. We applied MVARICA and CSPVARICA to all trials of both classes to obtain source decompositions. Subsequently, we used the same sources for estimating the PDC under each class separately, as shown in Figure 2.

**Single-Trial Estimation** Single-trial estimation suffers from the *curse of dimensionality*, because the amount of data available in one trial is limited. SCoT solves this problem by supporting regularized VAR model fitting. Figure 3 shows how regularization improves estimates for a model order of 20. In this example, we have four sources and estimate connectivity on 100 time samples, which results in a total of 400 available data samples. The number of free parameters in the VAR model is 320 ( $4 \cdot 4 \cdot 20$ ), which is a rather ill-posed fitting problem.

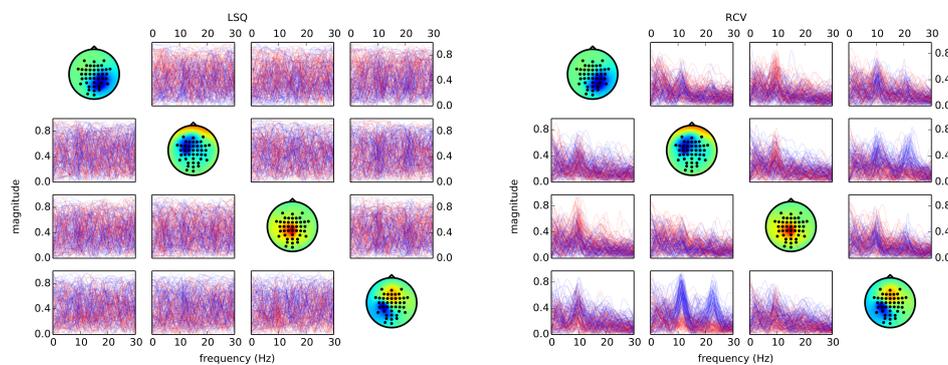


Figure 3: Single-trial estimation. Left: least squares fitting (LSQ), right: ridge regression (RCV). Lines correspond to individual trials, and the colors indicate the different classes (red: hand, blue: foot).

**Circular Plots** In this example, we show an alternative to spectral connectivity visualization. Instead of plotting the full connectivity spectrum, we show interaction only in selected bands. For this purpose, we averaged the PDC in the alpha (8–12 Hz) and beta (16–24 Hz) bands. If this average exceeds a threshold of 0.18 (alpha) or 0.25 (beta), we draw an arrow from the origin to the destination (Figure 4). Thus, these arrows indicate frequency dependent causal relations between sources.

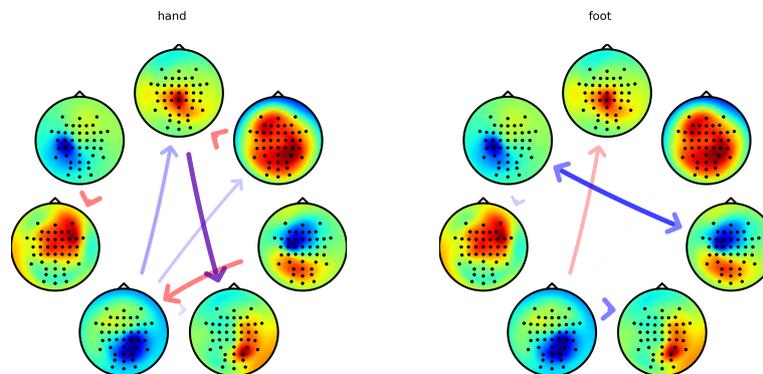


Figure 4: Circular plots. Left: hand MI, right: foot MI. These plots show the relative strength of connectivity (PDC in the alpha (8–12 Hz, red) and beta (16–24 Hz blue) bands). The width and intensity of the arrows indicates the strength of the connection.

## 4 Discussion and Conclusions

We demonstrated different approaches of source/connectivity estimation with SCoT. While joint estimation is the easiest approach, it is not suitable for evaluating changes in connectivity since sources change as well. Instead, the two-step approach allows us to evaluate changing connectivity between constant sources. Consequently, the two-step approach can be applied to single-trial estimation, which facilitates the use of connectivity features in BCIs [2, 3]. However, it is important to use regularization when fitting VAR models on low amounts of data.

**Funding** This work was supported by the FP7 Framework EU Research Project ABC (No. 287774) and the FWF Project Coupling Measures for BCIs (P20848-N15).

## References

- [1] L. A. Baccalá and K. Sameshima. Partial directed coherence: a new concept in neural structure determination. *Biological Cybernetics*, 84:463–474, 2001.
- [2] Martin Billinger, Clemens Brunner, and Gernot R Müller-Putz. Single-trial connectivity estimation for classification of motor imagery data. *Journal of Neural Engineering*, 10:046006, 2013.
- [3] Martin Billinger, Clemens Brunner, and Gernot R Müller-Putz. SCoT: A python toolbox for EEG source connectivity. *Front. Neuroinform.*, 8, 2014.
- [4] C. Brunner, R. Scherer, B. Graimann, G. Supp, and G. Pfurtscheller. Online control of a brain-computer interface using phase synchronization. *IEEE Transactions on Biomedical Engineering*, 53:2501–2506, 2006.
- [5] German Gómez-Herrero, Mercedes Atienza, Karen Egiazarian, and Jose L. Cantero. Measuring directional coupling between EEG sources. *NeuroImage*, 43:497–508, 2008.

# Active participation during walking reduces single trial connectivity in sensorimotor areas

Johanna Wagner<sup>1</sup>, Martin Billinger<sup>1</sup>, Teodoro Solis-Escalante<sup>2</sup>, Clemens Brunner<sup>1</sup>, Reinhold Scherer<sup>1</sup>, Christa Neuper<sup>3</sup> and Gernot Müller-Putz<sup>1</sup>

<sup>1</sup> Laboratory of Brain-Computer Interfaces, Institute for Knowledge Discovery, Graz University of Technology, Graz, Austria

johanna.wagner@tugraz.at, martin.billinger@tugraz.at, reinhold.scherer@tugraz.at, clemens.brunner@tugraz.at, gernot.mueller@tugraz.at

<sup>2</sup> Department of Biomechanical Engineering, Delft University of Technology, Delft, the Netherlands  
t.solisescalante@tudelft.nl

<sup>3</sup> Department of Psychology, University of Graz, Graz, Austria  
christa.neuper@uni-graz.at

## Abstract

Active contribution to a movement is crucial for motor learning. Previously, we showed that active participation is related to a suppression of mu and beta band activities over sensorimotor areas. In the current analysis, we aim at differentiating active and passive movements during robotic assisted gait training based on measures that quantify interactions between brain areas. Due to high artifact contamination of the EEG during walking, the data was pruned using independent component analysis (ICA). Single trial connectivity between brain sources was estimated using the fDFTF (full frequency directed transfer function). Three frequency bands were used for classification:  $\mu$  (7-12Hz),  $\beta$  (15-21Hz), and a subject specific frequency band ranging from 24-40 Hz. Based on the connectivity measures, we were able to separate active and passive movements with classification accuracies of  $81.0\% \pm 6.7$  on average. However a major challenge for the online application of these methods during gait rehabilitation remains automatic artifact correction.

## 1 Introduction

Extensive training in gait rehabilitation after stroke may be provided by using a robotic gait orthosis. Robotic rehabilitation requires little effort from the individual and can lead patients to move passively. However active contribution to a movement, has been shown to be crucial for motor learning [3].

Neural correlates of active participation during gait training have previously been shown by our group. We showed significant differences between active and passive walking in EEG  $\mu$  and  $\beta$  sensorimotor rhythms over the foot area of the sensory cortex [6]. The Authors showed that it is possible to distinguish between active and passive walking in single trials (mean accuracy: 68%) in 6 subjects using EEG band power features [5].

In the current manuscript, we show single trial connectivity analysis from the EEG recorded during active and passive walking in a gait robot. The goal of this analysis is twofold: First, we want to evaluate whether single trial connectivity measures allow to distinguish between active and passive walking. Second, we want to investigate the cortical networks related to active participation in gait training. This may help to better understand the underlying mechanisms and the optimal activation patterns for gait recovery.

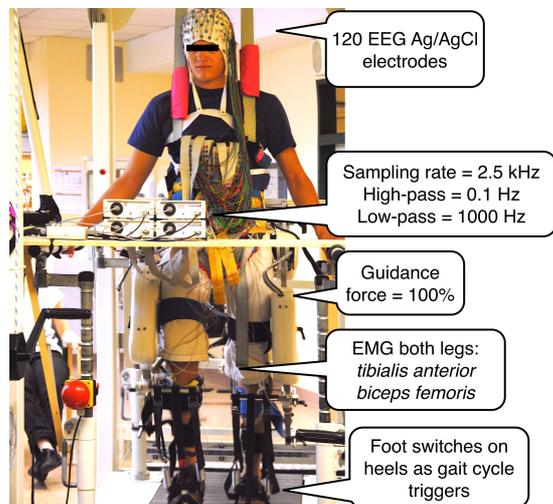


Figure 1: Experimental setup. Walking in the robotic gait orthosis. Speed (1.8-2.2 km/h) and body weight support ( $\sim 30\%$ ) were adjusted for each participant.

Participant	mean accuracy	std
1	76.4	6.0
2	83.7	7.9
3	75.7	4.7
4	82.3	4.5
5	75.6	5.9
6	92.8	3.8
mean	81.0	6.7

Table 1: Mean classification accuracy and standard deviation (std) for each participant in %.

## 2 Methods

We recorded the electroencephalogram (EEG) from 120 sites from 6 healthy volunteers ( $24 \pm 2$  years, 5 male) during active and passive walking (4 runs of 6 min each) with a robotic gait orthosis (Lokomat, Hocoma). In the active walking condition, participants were instructed to walk independently in the gait robot at the speed of the treadmill supporting their own weight. Passive walking demanded participants to let their legs be moved by the robot. Foot contact was measured by electrical foot switches placed on the heels of both feet. Figure 1 summarizes the experiments. For a more detailed description of the experiment see [6]

### 2.1 EEG Analysis

#### Preprocessing

Due to the high artifact contamination of the EEG during walking the data was pruned using Independent Component Analysis (ICA) prior to single trial analysis. Preprocessing for ICA included filtering from 1 to 200 Hz, resampling at 500 Hz, and manual rejection of non-stereotyped artifacts. Typical stereotyped artifacts (eye movements, muscle tension) were kept in the analysis as they are separated by ICA into only a few independent components (ICs). Infomax ICA [4] decomposed the EEG into ICs, representing brain, muscle, and artifact sources. The ICs were then categorized into cortical sources and artefact components considering scalp map, power spectrum, and event-locked time course. This procedure left on average 14 cortical sources per participant (range: 6-18 ICs). These cortical sources were then backprojected to the EEG data. To minimize computational effort, only 62 channels equally distributed over the scalp were selected for further analysis. PCA (Principal Component Analysis) was applied to the remaining 62 channels to reduce the dimensions of the remaining EEG channels to the number of backprojected ICs. The data was segmented from -0.5 to 2.5 seconds around the right heelstrike (time between contralateral steps  $\sim 1$  s). Around 250 trials were used for each class.

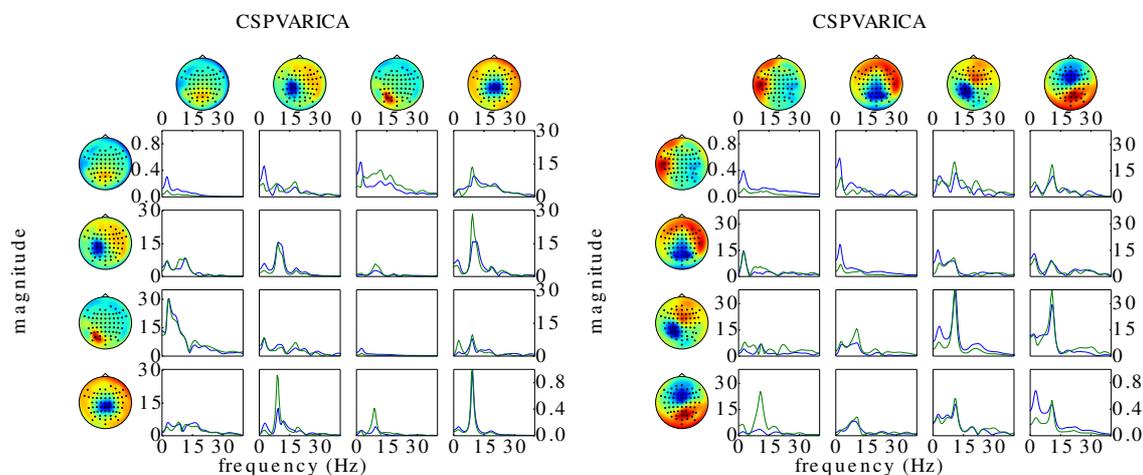


Figure 2: Directed connectivity with CSPVARICA for participant 1 (left) and 6 (right). The ffDTF between four components is plotted for active (blue) and passive (green) walking. The x-axis corresponds to the frequencies from 0 to 40 Hz, the y-axes to the magnitude of the ffDTF (in arbitrary units). Columns represent sources and rows sinks. The power spectral density of each source is plotted along the diagonal.

### Single trial connectivity analysis

For single trial connectivity analysis, the Python-based source connectivity toolbox SCoT [1] was used. Single trial analysis is performed in two steps. In the first step independent components are estimated with CSPVARICA [1]. The method transforms the EEG with Common Spatial Patterns (CSP) to find components that maximize the variance between conditions. Then a VAR model is fitted to the CSP components, and the residuals of the VAR model are decomposed by ICA to estimate the final unmixing matrix  $U$ . In the second step single trial component activations are obtained by multiplying the EEG with  $U$ . Then an autoregressive model is fitted for these component activations. The ffDTF (full frequency directed transfer function) [2] was calculated from the model to measure connectivity. This method estimates the direction of causal influences between sources. The ffDTF was averaged in the frequency bands  $\mu$  (7-12Hz),  $\beta$  (15-21Hz) and a subject specific frequency band ranging from 24 to 40 Hz. We have previously shown in [6] that these frequency bands account for differences between active and passive walking. The number of components was four and model order was set to 100. This resulted in 48 features for classification. Classification was performed with linear discriminant analysis. The whole procedure was applied in a 10 fold cross-validation to estimate the performance. For a more detailed description of the method see [1].

## 3 Results

Connectivity measures allowed differentiation between active and passive movements with an average classification accuracy over subjects of  $81.0\% \pm 6.7$ . Classification accuracies for single subjects are displayed in Table 1. Visual inspection revealed that connectivity magnitude in classification relevant frequency bands  $\mu$  and  $\beta$  was higher during passive compared to active walking. Most participants had at least one contributing source in the central midline area and in posterior areas. For the results of two exemplary subjects see Figure 2.

## 4 Discussion

Our results show that it is possible to separate active and passive walking in single trial EEG data based on connectivity measures. Classification accuracies were above chance in all subjects, and were on average 10% higher compared to our previous results using band power features [5]. However, it still has to be evaluated whether this improvement is due to the method used, the features or to the pruning of the EEG with ICA prior to feature selection. Furthermore, a major challenge for the online application of these methods during gait rehabilitation remains automatic artifact correction.

From a neurophysiological perspective, the analysis revealed that connectivity magnitude in  $\mu$  and  $\beta$  bands between sensorimotor sources is higher during passive compared to active walking. The magnitude of the fDFTF depends on spectral power, thus the results fit with our previous findings showing a suppression of  $\mu$  and  $\beta$  bands related to active participation. Suppression of  $\mu$  and  $\beta$  rhythms has previously been related to the activation of sensorimotor areas. Thus, our results suggest that active walking increases the activation of sensorimotor regions and reduces information flow between these areas. Further analysis should evaluate in more detail the sources contributing to differences in connectivity over subjects. This may reveal neural networks underlying active participation in gait training.

## 5 Acknowledgements

This work was partly supported by the European Union research project BETTER (ICT-2009.7.2-247935), the Land Steiermark project BCI4REHAB and the FWF project (P-20848). The authors thank G. Schaffhauser and P. Holper for assistance during the experiments.

## References

- [1] M. Billinger, C. Brunner, and G. R. Müller-Putz. SCoT: a Python toolbox for EEG source connectivity. *Frontiers in neuroinformatics*, 8, 2014.
- [2] A. Korzeniewska, M. Manczak, M. Kaminski, K. J. Blinowska, and S. Kasicki. Determination of information flow direction among brain structures by a modified directed transfer function (dDTF) method. *Journal of Neuroscience Methods*, 125:195207, 2003.
- [3] M. Lotze, C. Braun, N. Birbaumer, S. Anders, and L. G. Cohen. Motor learning elicited by voluntary drive. *Brain*, 126:866–872, 2003.
- [4] S. Makeig, A. J. Bell, T. P. Jung, and T. J. Sejnowski. *Adv Neural Inf Process Syst*, chapter Independent Component Analysis of Electroencephalographic Data, pages 145–151. MIT Press Cambridge, MA, 1996.
- [5] T. Solis-Escalante, J. Wagner, R. Scherer, P. Grieshofer, and G. Müller-Putz. Assessing participation during robotic assisted gait training based on EEG: feasibility study. *Proceedings of the 3rd TOBI Workshop*, pages 61–62, 2012.
- [6] J. Wagner, T. Solis-Escalante, P. Grieshofer, C. Neuper, G. Müller-Putz, and R. Scherer. Level of participation in robotic-assisted treadmill walking modulates midline sensorimotor EEG rhythms in able-bodied subjects. *NeuroImage*, 63:1203–1211, 2012.

# Proposal on Brain Wave Personal Authentication with Wireless Neuroheadset

K. Tsuru<sup>1</sup>, M. Nagaki<sup>2\*</sup>

<sup>1</sup> Department of Information Engineering, Oita National College of Technology, Oita, Japan

<sup>2</sup> Department of Computer and Control Engineering, Oita National College of Technology, Oita, Japan

tsuru@oita-ct.ac.jp

## Abstract

Brain wave biometric personal authentication is an emerging technology in information security. This form of biometrics is effective in preventing attacks by impostors because of the difficulties of obtaining and impersonating personal brain wave data. Previous studies of this type of biometrics have generally used a wired electrode measurement system, but setting up the system was time-consuming. Hence, we applied brain wave biometrics using a wireless measurement device. Our results showed the authentication rate was over 0.9 on the discrete cosine transform (DCT) feature extraction and application for practical purposes.

## 1 Introduction

We have studied biometrics on brain waves to develop a diverse and secure authentication system. Brain wave biometrics has two advantages over prevailing biometrics. One is the difficulty of eavesdropping on personal brain wave data. The second advantage is that brain waves can reflect individual mental activities. This property leads to many possibilities for diverse uses of biometrics. Traditional biometric methods used single fixed templates. In contrast, brain wave biometrics could identify people based on templates that reflect different brain activities, such as cognitive processes.

The biometric using the brain wave approach is used to assess  $\alpha$  waves. Poulos et al. [1] first tried to identify individuals based on the EEG. They analyzed  $\alpha$  waves of four subjects' EEGs, using a neural network classification method. Paranjape [2] also used brain wave data based on  $\alpha$  waves during eyes open/closed for biometric analysis. These data based on  $\alpha$  waves reported consistent classification results, and this work required only a few electrodes. However, it was necessary for subjects to sit quietly for a relatively long period. A newer method utilized an event-related potential from a cognitive human brain process. Palaniappan [3] investigated the  $\gamma$  wave band of the visual-evoked potential elicited during a mental task for personal identification, and Mercel [4] studied personal authentication based on motor images of left or right hand movement and word generation. In addition we previously studied the approach to the discrete cosine transform (DCT) of motor imageries for features, extracting the best individual features [5]. However, all these studies measured brain signals using the wired electrode system. A major drawback is that preparation for using this system takes time and care, and it is also difficult to realize a brain wave biometric system.

In this paper, we investigate brain wave biometrics using a wireless measurement device on brain signals. This device has not previously been explored for this task. We attempted to classify five healthy subjects based on four situations.

---

\* Graduated in March 2014

## 2 Methods

Five healthy subjects (4 males and 1 female, age:  $20 \pm 0.32$ ) provided their informed consent to participate in the experiment on several different days.

### 2.1 Experimental Paradigm

To show individual features, we prepared four experimental situations as shown below:

Case I:	Close eyes and relax
Case II:	Write consecutive numbers on paper
Case III:	Solve a 40-piece jigsaw puzzle
Case IV:	Read a book

The purpose of Case I is for the subject to measure the authentication rate in a relaxed state of mind. The purpose of Case II is for the subject to measure the authentication rate in a centered state. In Case III, the purpose is for the subject to measure the authentication rate in a centered state with visual information, while in Case IV the purpose is for the subject to measure the authentication rate in a relaxed state with visual information. Four cases were measured in series with a brief resting time after every activity on several different days. The task duration in each case was about 6 minutes. All measurement was carried out on sitting subjects in a room.

### 2.2 Measurement

We employed an Emotiv EEG neuroheadset [6] to measure brain waves. The Emotiv has 14 saline electrodes with 2 reference electrodes to wirelessly transmit brain wave signals to a computer. The sampling frequency of the brain wave signals was 128 Hz. The brain data was removed as artifacts using digital notch filters (50 Hz and 60 Hz) and a low-pass filter ( $\sim 43$ Hz). To avoid the effect of impedance difference, we prepared the low impedance ( $\sim 2$  k $\Omega$ ) of each electrode with an impedance-measuring program. We used C++ program using Emotiv API for storing the brain wave data in a computer. The brain wave data was divided into 8-second lengths, and the average was subtracted. The number of each subject's brain wave data was between 208 and 320 in every case. So we put the data from each case together for all subjects.

### 2.3 Feature Extraction

We applied feature extraction methods based on DCT feature extraction as follows [5]. First, we employed the spectrum data by fast Fourier transform (FFT) with a rectangular window. With the FFT data, we calculated the sum of every 2 Hz spectrum power band from 0 Hz to 40 Hz. This frequency band included  $\alpha$  wave (8–13 Hz),  $\beta$  wave (14–30 Hz), and  $\mu$  wave (12–18 Hz) activity. The second method was done by adding DCT after the previous method. This method reduced the spectrum data that was converted to DCT data. The DCT is a technique for converting a signal into elementary frequency components. We obtained most of the features from the lower range of the DCT data because the spectrum information concentrated the low portion of these data [5]. These features at 14 electrodes were put into one data set. Thus, we extracted 4 features per each electrode. An authentication rate was required by classifiers on 56 features.

### 2.4 Classification

We selected three classifiers: linear discriminant analysis (LDA), support vector machine (SVM), and neural network (NN). The probability of personal identification was called an authentication rate. We estimated the rate using a 10-fold cross-validation method.

### 3 Results

Table 1 shows the results of the average authentication rate that evaluated four features of the DCT data per electrode at 0–40 Hz. Also the authentication rate was calculated by several classification methods. The authentication rates are the average for the five subjects. As NN classification is dependent upon initial values, their authentication rate is calculated based on an average of 10 trials. Taking into account the characteristics of the small subject group, all authentication rates showed consistent results. Those results of the authentication rates were not different among the four cases, as we anticipated.

Case	LDA				SVM			NN
	linear	quadratic	Mahalanobis	rbf	polynomial	quadratic	linear	
I	0.95	0.96	0.96	0.81	0.98	0.99	-	0.92
II	0.99	0.99	0.98	0.80	0.99	0.99	0.99	0.94
III	0.98	0.98	0.96	0.81	0.96	0.99	0.99	0.83
IV	0.98	0.98	0.97	0.81	0.97	0.98	0.97	0.93

-: no convergence

**Table 1: Results of average authentication rate of subjects on each classification method. Authentication rate obtained for 4 features per electrode.**

Next, we estimated the average authentication rates for each electrode by LDA classifier for finding the area of the head that distinguished our subjects. Table 2 shows average authentication rates. Their rates ranged from 0.47 to 0.76; the average authentication rates of all electrodes in the cases were 0.58, 0.59, 0.60, and 0.60 respectively. The rate of the electrodes used on the occipital region gave consistent results on the whole. Meanwhile, the rate of the electrodes on the temporal cortex regions depends on the activity.

Case	AF3	F7	F3	FC5	T7	P7	O1	O2	P8	T8	FC6	F4	F8	AF4
I	0.56	0.58	0.50	0.55	0.51	0.56	0.67	0.70	0.61	0.55	0.55	0.58	0.58	0.60
II	0.53	0.58	0.55	0.56	0.52	0.62	0.69	0.71	0.65	0.57	0.56	0.61	0.58	0.50
III	0.53	0.64	0.55	0.55	0.64	0.62	0.65	0.62	0.59	0.60	0.64	0.54	0.59	0.60
IV	0.47	0.58	0.52	0.53	0.57	0.64	0.76	0.68	0.65	0.64	0.69	0.48	0.61	0.59

**Table 2: The average authentication rates for each electrode by LDA classifier. The light gray areas are over 0.6 and the dark gray areas are over 0.7**

### 4 Discussion and Conclusion

At first, the experiment preparation time for the wireless neuroheadset took 5 to 10 minutes under low-resistance contact by an impedance check program, but it took subjects one minute to get used to the headset. We think this is within an acceptable range to achieve practical use of personal authentication. In previous brain wave authentication studies, wired electrodes were employed for the measurement. The preparation time for setting electrodes on the skin of the scalp was over 30 minutes.

Furthermore, setting them requires help from a few other people. Consequently it is a more challenging process when compared to the existing authentication methods, such as fingerprint, iris, and facial recognition. By contrast, the preparation for using the wireless EEG neuroheadset took little time, and an examinee can prepare for authentication by him/herself. However, we must develop one's positioning method for the location compensation in the next step. Hence, this represents a potential first step towards the practical application of a biometrics system with brain waves.

The authentication rates by DCT feature extraction showed consistent results by all classification methods though five subjects. A previous study showed the authentication rate to be 0.79 for 23 subjects using the wired measurement system [7]. The wireless device indicated the performance of the wired measurement system.

The authentication rate on the occipital electrode was better than that on the temporal electrode. These results were the same for every situation. Under the centered state, such as solving a puzzle or reading a book, the authentication rate on the temporal electrode gave better results. Those results show that brain wave biometrics with several selected electrodes provides a consistent authentication rate.

We investigated the brain wave biometric approach using a wireless brain wave device. The results indicate that the authentication rate of brain wave biometrics was over 0.9. In addition, biometrics using a wireless brain wave device is suitable for a personal authentication system because of its high level of accuracy and short setup time. Moreover, future research should increase the number of subjects in authentication experiments to improve reliability.

## Acknowledgements

This study was supported in part by a Grant-in-Aid for Scientific Research (C) 25330163 from the Japan Society for the Promotion of Science (JSPA).

## References

- [1] M. Poulos, M. Rangoussi, V. Chrissikopoulos, and A. Evangelou, "Person identification based on parametric processing of the EEG Electronics, Circuits and Systems," *Proceedings of ICECS '99. The 6th IEEE International Conference*, vol. 1, pp. 283–586, 1999.
- [2] R. B. Paranjape, J. Mahovsky, L. Benedicenti, and Z. Koles, "The electroencephalogram as a biometric," *Proc. Canadian Conf. Elect. and Comp. Eng.*, vol. 2, p. 1363–1366, 2001.
- [3] R. Palaniappan and D. Mandic, "Pattern Analysis and Machine Intelligence," *IEEE Transactions*, vol. 29, p. 738–742, 2007.
- [4] S. Marcel and J. Millan, "Pattern Analysis and Machine Intelligence," *IEEE Transactions*, vol. 29, p. 743–752, 2007.
- [5] K. Tsuru and G. Pfurtscheller, "Brainwave Biometrics: A New Feature Extraction Approach with the Cepstral Analysis Method," *Trans. of Jap. Soc. for Med. and Biolog. Eng.*, vol. 50, no. 1, pp. 162–167, 2012.
- [6] Emotiv Epoch, <http://emotiv.com/epoch/features.php> (Accessed April 4, 2014).
- [7] C. Miyamoto, S. Baba, and I. Nakanishi, "Biometric person authentication using new spectral features of electroencephalogram (EEG)," *Proc. of ISPACS*, p. 1–4, 2008.

# ERP assessment and EEG/fNIRs communication in a patient with a disorder of consciousness

Sandra Vesper<sup>1\*</sup>; Boris Kotchoubey<sup>1</sup>, Bin Xia<sup>1</sup> and Niels Birbaumer<sup>1</sup>

<sup>1</sup>University of Tübingen, Germany  
sandra.veser@googlemail.com

## Abstract

To examine the presences of electrophysiological indicators, by measuring event-related potentials (ERPs), is one solution to detect residual cognitive functions in patients with disorders of consciousness (DOC). To investigate such a patient, different kinds of stimuli like tones, words, and sentences were used with and without instruction and were hierarchically ordered according to the needed processing steps. In addition, the patient performed a semantic computer-brain-interface (BCI) while electroencephalography (EEG) and near-infrared spectroscopy (NIRs) signals were recorded. We found evidence that the DOC patient was able to follow active instructions and to shift his attention, but had only limited control of the BCI performance.

## 1 Introduction

One way to detect the residual cognitive functions in brain-damaged patients, who has lost the ability for behavioural performance, is to measure their brain responses, for example using event-related brain potentials (ERPs). This approach has been successfully reported in a number of studies (e.g., Cruse, Chennu, Chattele, C, Bekinschtein, Fernández-Espejo, Pickard et al., 2011; Daltrozzo, Wioland, Mutschler, Lutun, Calon, Meyer et al., 2009, Kotchoubey, Lang, Mezger, Schmalohr, Schneck, Semmler et al., 2005; Schoenle & Witzke, 2004) and indicates that the brain of such patients might be able to process information at various levels of complexity including semantic information as well as understanding active instructions like silent counting or even attentional shifts (e.g., Boly, Garrido, Gosseries, Bruno, Boveroux, Schnakers et al. 2011; Monti ; Vanhau denhuys, Coleman, Bol, Pickard, Tshibanda, et al., 2010).

In addition, brain responses can be used to set-up a communication channel by a so called Brain Computer Interface (BCI) for patients who are severe paralysed and cannot communicate by any other means. The brain responses of those patients can be controlled by using electroencephalography (EEG; e.g., Birbaumer, Ghanayim, Hinterberger, Iversen, Kotchoubey, Kübler, et al., 1999) or functional-near infrared spectroscopy (fNIRs; e.g., Gallegos-Ayala, Furdea, Takano, Ruf, Flor, Birbaumer, in press). Here, we wanted to investigate whether a good outcome in the cognitive assessment using ERPs leads to good BCI performance.

---

\* Masterminded EasyChair and created the first stable version of this document

## 2 Methods

A 60 year old, male patient, diagnosed was minimally conscious state (MCS: Giacino, Ashwal, Childs, Cranford, Jennett, Katz, et al., 2002) and having a CRS-R score of 23 participated in four sessions. In the first and second sessions he was assessed by using several passive and active ERP assessment paradigms. In the other two sessions he performed a semantic BCI with two answer categories.

### 2.1 ERP Assessment

ERP paradigms. (1) A multifeature oddball paradigm principally adapted by Näätänen et al. (2004). The standard stimuli were harmonic tones (440+880+1760 Hz) with a duration of 75 ms and an intensity of 70 dB, presented binaurally via headphones. 50 % of all stimuli were standards. The other kinds of stimuli were presented with the frequency of 5% each. Two of them differed from the standards by their pitch (f; 220+440+660 Hz and 880+1760+2640 Hz), two by their duration (d; 50 ms and 100 ms), two by their location (l, left or right monaural presentation), two by intensity (i, 50 dB and 90 dB) and two by complexity (c; 440+1760 Hz and 440+660+880+1760 Hz). The SOA was 500 ms. (2) A frequency oddball paradigm. The paradigm used a frequent complex tone (Standard: 440+880+1760 Hz) and a rare complex tone (Deviant: 247+494+988 Hz) with an ISI of 850 ms. (3) A word-prime paradigm. The paradigm tested semantic processing at the word level. 200 pairs of words spoken by female voice were presented. 100 pairs contained semantically closely related words (e.g., cold - warm) and the other 100 words containing unrelated words (e.g., cold - green). The ISI within word pairs were 400 ms and between word-pairs were 900 ms. (4) A sentence understanding paradigm, to test semantic processing at the sentence level. 200 sentences were used and in 100 of them the last word was highly expected in the context (e.g., the eel is a fish), while in the remaining sentences the ending was semantically incorrect (e.g., the eel is a bird). ISI between sentences were 900 ms. (5) An oddball paradigm with the same stimulation as in (2). In addition, the patients received the instruction to count the deviants. (6) A dichotic listening paradigm. It used a word stream containing five semantic categories (animals, professions, tools, body parts and household objects). The words were presented alternating to the left and right side with a jittered ISI between 150-300 ms. The patients' task was to attend to one side and count the animals of the attended side.

The EEG was recorded according to the international 10-20 electrode system with 31 active electrodes (F3 Fz F4 FC5 FC1 FCz FC2 FC6 T7 C3 Cz C4 T8 CP5 CP1 CPz CP2 CP6 P7 P3 Pz P4 P8 PO3 POz PO4 O1 Oz O2). The signal was digitized at 512 Hz and filtered with a bandpass filter between 0.1 Hz and 100 Hz. The vertical and horizontal electrooculogram were recorded. Offline, the EEG was filtered using a Kaiser low-pass at 25 Hz with 1856 points and ocular artifacts were corrected. The continuous EEG data were split into epochs, respective of the paradigm. In addition, trials of each type of categories were averaged. Difference waves were obtained by subtracting the standard from the each of the deviants. Finally, we computed a running *t*-test to evaluate whether the difference waves were significantly different from the baseline in the time range of the respective components.

### 2.2 EEG/fNIRs BCI

Like in the sentence understanding paradigm of the previous section, the patient was presented with correct and incorrect sentences grouped in 20 sentences in each block. A total of four blocks per session was presented. After each sentence the patient's task was to think "ja" (yes), if the sentence was correct and "nein" (no) if it was incorrect (Figure 1A) while EEG (using the same electrode set-up as in the assessment) and fNIRs (Spectratch OEG-Sp02, Spectratch Inc. Japan) data were recorded simultaneously with a sample rate of 12.2 Hz. 16 optodes were placed at the forehead. The

EEG data were preprocessed similarly to the method used above and then t-CWT analysis for feature classification (Bostanov et al. 2004). For fNIRs data we used the mean amplitude for each channels of the oxyhemoglobin. For both ERP and fNIRs data, we used a support vector machine to classify the correct and incorrect answers using a grid searching to find the best parameter und using a five-fold cross-validation. We trained the model using the first 3 blocks and tested the last block. In addition, we calculated the coincidences of the correct detected classification results of EEG and fNIRs.

### 3 Results

#### 3.1 ERP Assessment

In Table 1 we show the summarized results of the all expected ERP components. In the (1) multifeature paradigm 3 deviant elicited the expected ERP component, the mismatch negativity (MMN). In the frequency oddball (2) we also find the expected P3 component as well in the counting oddball task (5). However, the P3 component in the active task was not larger than in the active condition, thus, it is not clear whether the P3 component of the active counting task includes really active counting contingence. In addition, we found a N400 component in the sentence understanding paradigm but not in the word-priming paradigm. Further, we obtained a side specific effect in the dichotic listening task, where the P3 component for the attended side was enhanced as compared with the unattended side.

(1) multifeature					(2) frequency	(3) word-prime	(4) sentence	(5) counting	(6) dichotic	
d	f	i	l	c	+	-	+	+(-)	left	right
-	+	+	+	-					+	+

Table 1: Results of the ERP assessment paradigms: + is expected ERP component was present; - expected ERP component was absent.

#### 3.2 EEG/fNIRs data

The results using EEG and fNIRs classification are depicted in Figure 2B. As it can be seen, the results of the EEG/fNIRs data for each method alone are above or equal the theoretical chance level (55 % with alpha = 0.5; Müller-Putz G et al. 2008). However, the coincidence (trial was correctly identify with EEG and also fNIRs) of correctly classified trials was below the chance level.

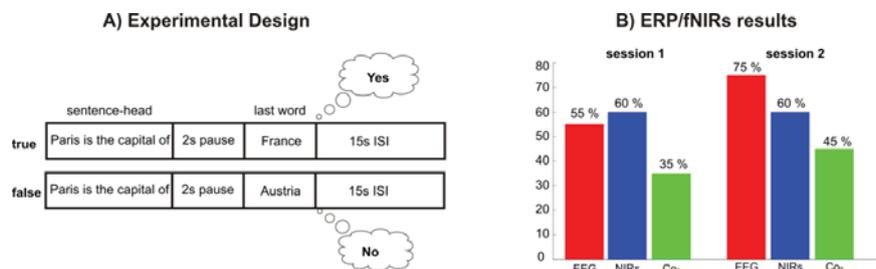


Figure 1: (A) Experimental design of the EEG/fNIRs paradigm; (B) ERP/fNIRs classification results

## 4 Discussion and Conclusion

We found evidence in the ERP assessment that the MCS patient was able to perceive and process various aspects of his environment, including speech perception. Moreover, he was able to modulate his attention. Even so, the technique still need to become more precise and is not yet capable of exact assessment in each individual paradigm (we found no N400 component in the word-prime paradigm but in the sentence paradigm). It is useful to assess the presence of higher cortical functions in patients who cannot express themselves behaviorally before applying a BCI.

BCI performance of this patient measured with EEG was slightly above the theoretical chance level in the first session but improved in the second session. Without the knowledge of the results of the ERP assessment, the BCI results of the first session would be quite discouraging for patients and their caregivers and might lead to aversion for training a BCI by itself because negative online results enhances the pressure on the patient and induces negative emotions. Furthermore, the data of the NIRs BCI were identical in both sessions. It should be noted that NIRs does not depend on the quality of the data which causes trouble in the EEG, like electricity flow of the surrounding but have other influences like the physiological noise or movement artifacts. Thus, combined EEG and fNIRs BCI seemed to be a promising tool for further BCI applications if it was shown that the patient is able to perceive and process complex stimulation.

## References

- Birbaumer, N., Ghanayim, N., Hinterberger, T., Iversen, I., Kotchoubey, B., Kübler, A., Perelmouter, J., Taub, E. and Flor, H. (1999). A spelling device for the paralyzed. *Nature*, 398, 297–298.
- Boly M, Garrido M.I, Gosseries O, Bruno MA, Boveroux P, Schnakers C, et al... Laureys S, Friston K. Preserved feedforward but impaired top-down processes in the vegetative state. *Science* 2011: Vol. 332 no.6031 pp 858-862.
- Bostanov V, Kotchoubey B. The t-CWT: A new ERP detection and quantification method based on the continuous wavelet transform and Student's t-statistics. *Clinical Neurophysiology*. 2006;117:2627-44.
- Cruse, D., Chennu, S., Chatelle, C., Bekinschtein, T. A., Fernández-Espejo, D., Pickard, J. D., . . . Owen, Adrian M. (2011). Bedside detection of awareness in the vegetative state: a cohort study. *The Lancet*, 378(9809), 2088-2094.
- Daltrozzo, J., Wioland, N., Mutschler, V., Lutun, P., Calon, B., Meyer, A., . . . Kotchoubey, B. (2009). Cortical information processing in coma. *Cognitive and behavioral neurology: official journal of the Society for Behavioral and Cognitive Neurology*, 22(1), 53-62..
- Gallegos-Ayala G, Furdea A, Takano K., Ruf CA, Flor H, Birbaumer N. Brain communication in a completely locked-in patient using bedside near-infrared spectroscopy. *Neurology* (in press).
- Giacino, J., Ashwal, S., Childs, N., Cranford, R., Jennett, B., Katz, D. I., . . . Zasler, N. D. (2002). The minimally conscious state: definition and diagnostic criteria. *Neurology*, 58(3), 349-353
- Monti M.M, Vanhauzenhuysse A, Coleman M.R, Boly M, Pickard J.D, Tshibanda L, Owen A.M, Laureys S. Willful modulation of brain activity in disorders of consciousness. *N Engl J Med* 2010;362:579-89.
- Kotchoubey, B., Lang, S., Mezger, G., Schmalohr, D., Schneck, M., Semmler, A., . . . Birbaumer, N. (2005). Information processing in severe disorders of consciousness: Vegetative state and minimally conscious state. *Clinical Neurophysiology*, 116(10), 2441-2453. doi: DOI 10.1016/j.clinph.2005.03.028
- Müller-Putz G, Scherer R, Brunner C, Leeb R, Pfurtscheller G. Better than random: a closer look on BCI results. *International Journal of Bioelectromagnetism*. 2008; 10(1): 52-55.
- Schoenle, P. W., & Witzke, W. (2004). How vegetative is the vegetative state? Preserved semantic processing in VS patients - Evidence from N400 event-related potentials. *Neurorehabilitation*, 19(4), 329-334.

# Improving ECoG-based P300 speller accuracy

Milena Korostenskaja<sup>1-3\*</sup>, Christoph Kapeller<sup>4</sup>, Robert Prueckl<sup>4</sup>, Rupert Ortner<sup>4</sup>, Po-Ching Chen<sup>1-3</sup>, Christoph Guger<sup>4</sup>, and Ki Hyeong Lee<sup>3</sup>

<sup>1</sup>Milena's Functional Brain Mapping and BCI Lab, Florida Hospital for Children, Orlando, FL, USA; <sup>2</sup>MEG Lab, Florida Hospital for Children, Orlando, FL, USA; <sup>3</sup>Comprehensive Pediatric Epilepsy Center, Florida Hospital for Children, Orlando, FL, USA; <sup>4</sup>g.tec Guger Technologies OG, Schiedlberg, Austria

\*Corresponding author: milena.korostenskaja@gmail.com

## Abstract

Brain-Computer Interfaces (BCIs) are expected to become a game-changing modality for the communication and control of both devices and virtual environments in humans whose direct neuromotor response needs to be augmented, bypassed, or replaced. Recent developments in the area of BCI research suggest that subdural electrocorticography (ECoG)-based BCIs might be more efficient than scalp electroencephalography (EEG)-based ones. In this study we compared the performance of an EEG-based P300 speller BCI (which allows typing without utilizing muscular activity) to an ECoG-based one. Three different approaches (i-iii) were used to select the best electrode sites from acquired ECoG signals. We found that the ECoG-based P300 speller performance accuracy varied widely between the different approaches. Approach iii, based on the choice of the longest duration of significant difference between TARGET and NON-TARGET stimuli, provided the most accurate ECoG-based results. Those results were comparable to and even surpassed the EEG-based results. However, the invasive nature of ECoG-based electrode selection has its own disadvantages. Therefore, as a future perspective, we propose to explore the potential contribution of non-invasive methods (for example, magnetoencephalography – MEG) to augment the subdural/depth electrode placement process.

## 1 Introduction

Brain-Computer Interfaces (BCIs) are powerful tools for enabling communication between people and the surrounding world using direct brain activity recorded non-invasively from the scalp (EEG), invasively from the brain surface (ECoG), or from deep within the brain itself using depth electrodes (Shih & Krusienski, 2012). The need for BCIs is immense as it can benefit a number of clinical populations, such as people with stroke, locked-in syndrome, amyotrophic lateral sclerosis (ALS), and other severe disorders involving deterioration or damage of muscular system (Silvoni, et al., 2013) (Birbaumer, Gallegos-Ayala, Wildgruber, Silvoni, & Soekadar, 2014). The “P300 speller” is currently the most popular BCI system enabling communication by spelling words or phrases with direct brain-controlled selection from menus presented on a computer screen (Farwell & Donchin, 1988). People achieve up to 91% accuracy with a speed of 2-3 characters per minute using a scalp EEG-based P300 speller (Guger, et al., 2009). However, some questions still remain: (1) Are ECoG-based BCIs superior to the EEG-based ones? and (2) How can the performance of ECoG-based BCIs be further improved? Recent, though limited, reports (Shih & Krusienski, 2012) suggest that invasive ECoG-based BCIs can be faster and more accurate than non-invasive EEG-based BCIs. Approaches to

improve the accuracy of ECoG-based P300 speller have been proposed (Speier, Fried, & Pouratian, 2013), however they still need further development. Therefore, in this study we sought (1) to compare ECoG- and EEG- based BCI performance; and (2) to maximize performance of the ECoG-based P300 speller by applying different approaches to the selection of electrodes chosen for ECoG signal detection.

## 2 Methods

One female and 3 male patients (mean age 23 +/-10.6) diagnosed with intractable epilepsy and undergoing evaluation for resective epilepsy surgery were recruited. Two patients (#1 and #2) underwent both EEG- and ECoG-based recordings with P300 speller; while two other patients (#3 and #4) underwent only an ECoG-based P300 speller session. Participants received both character- and face-based stimuli in the testing sessions.

### 2.1 Training

Each participant was presented with three 5 character words to train the linear classifier to distinguish the P300 response. The row/column speller that flashes an entire column or row of characters was utilized. Flashing rows and columns were each presented 15 times (15x15) and an individual classifier was calculated before the “free spelling” experiment began. All active electrodes (8 total) were used for free spelling based on EEG recordings, whereas the 8 “best” were chosen for each subject from all possible ECoG electrodes. The criteria for selecting the 8 “best” in each approach are detailed below.

### 2.2 Free spelling

The free spelling began with presenting flashing rows and columns 15 times each. The number of flashes was gradually decreased while the desired level of difficulty was obtained. The decrease was done in the following increments: 15, 8, 4, 2, and 1. A 5 character word was used for each level of difficulty.

### 2.3 EEG-based recording

The EEG data were acquired from 8 electrodes (Fz, Cz, P3, Pz, P4, PO7, POz, PO8) using a g.USBamp (24 Bit biosignal amplification unit, g.tec medical engineering GmbH, Austria) at a sampling frequency of 256 Hz. The ground electrode was located on the forehead while the reference was mounted on the right earlobe (Guger, et al., 2009).

### 2.4 ECoG-based recording

The total number of subdurally placed electrodes available to be used from synchronized g.USBamp devices varied for each of the individuals. The data was analyzed with g.BSanalyze Matlab-based program (g.tec Medical Engineering GmbH, Austria) and additional Matlab scripts created by engineers from g.tec medical engineering GmbH. These approaches were used to screen all available ECoG electrodes and identify 8 most informative ones for free spelling (number “8” is chosen to have a fair comparison with scalp EEG P300 speller performance; the same number of electrodes and the same algorithms are used): (i) choosing the signal with the highest amplitude; (ii) choosing the signal with the lowest correlation between standard and deviant responses; (iii) choosing channels based on the longest duration of significant difference ( $p < 0.05$ ) between TARGET and

NON-TARGET trials. The ECoG electrode locations for P300 speller varied between study participants. The main regions of overlap included frontal and central cortex (patients #1-3). In addition, patient's #3 electrodes were located in occipital cortex.

## 3 Results and Discussion

### 3.1 EEG-based recording

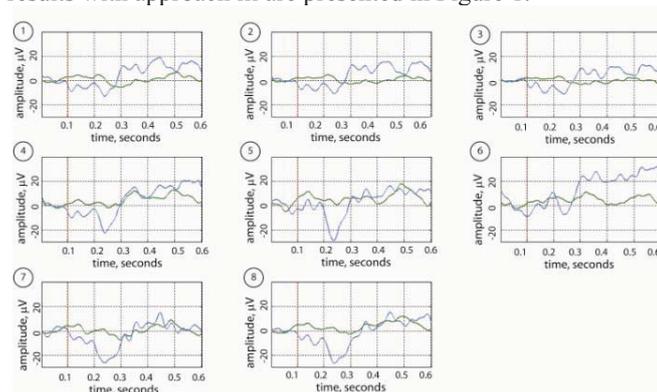
Accuracy in EEG-based recordings varied with different classifiers (ranging between 30% and 100%). It reached a maximum of 100% in patient #2 when three classifiers were combined for 15x15 presentation rate. Interestingly, accuracy was higher with faces of popular people (reaching 90% in most conditions) rather than simple flashing letters.

### 3.2 ECoG-based recording

Accuracy in ECoG-based recordings was significantly lower (ranging between 0% and 11%) in patients #1-#3 with flashing rows and columns each presented 15 times (15x15) than EEG-based recordings in patients #1 and #2 using approaches i and ii. Performance may be affected by the location on the brain surface of the grid coverage and a failure to identify the “best” channels. We aimed to investigate if the low accuracy results can be attributed to identification failure of the “best” channels. Therefore, approach (iii) was developed – see section New Analysis Approach.

### 3.3 New analysis approach

In order to improve accuracy of ECoG-based P300 speller performance, a new P300 analysis approach to identify the best 8 ECoG channels was developed and tested in patient #4. P300 responses were analyzed as follows: 1) A butterworth filter 4<sup>th</sup> order 0.1 Hz to 30Hz was applied and data was triggered into TARGET and NON-TARGET trials; 2) Artifacts were identified and removed; 3) A Mann-Whitney U-test was used to test if TARGET and NON-TARGET samples originate from the same distribution and led to a p-value for each sample and channel. The channels were selected according to the longest period of significant difference ( $p < 0.05$ ) between TARGET and NON-TARGET trials. The results with approach iii are presented in Figure 1.



**Figure 1.** Best eight ECoG P300 responses using approach iii in subject #4 Average signal amplitude over time for response to TARGET (blue) and NON-TARGET (green) stimuli.

Importantly, with approach (iii) for the selection of electrodes, the accuracy of the ECoG-based P300 speller surpassed results from approaches i-ii, as well as scalp EEG-based P300-speller by providing with accuracy ranging from 85% to 100% (again, for a session with flashing rows and columns each presented 15 times).

## 4 Conclusions

Intracranial grid placement may provide users with unique opportunity to control real and virtual worlds with high speed and accuracy. ECoG-based P300 speller can provide comparable results to the EEG-based one in terms of accuracy. In fact, it can surpass EEG-based P300 speller accuracy. Moreover, this accuracy can be further improved. Indeed, the identification of the “best” locations of ECoG electrodes plays very important role in ensuring high accuracy of P300 speller performance. Our proposed signal processing algorithm based on the duration of the significant difference between TARGET and NON-TARGET responses may be of high value to achieve this goal. It is important note, that this approach is based on information from invasively implanted ECoG electrodes. Because the location of the BCI implantation site may play a critical role in performance, we believe that future approaches should explore non-invasive procedures, such as magnetoencephalography (MEG), to enhance ECoG electrode placement. We suggest that one of the optimal ways to achieve maximal P300 speller performance is: (1) use intracranially/subdurally implanted electrodes; (2) navigate electrode implantation with non-invasive methods; and (3) utilize quantitative approaches/algorithms based on available ECoG information for the choice of the best electrodes.

## References

- Birbaumer, N., Gallegos-Ayala, G., Wildgruber, M., Silvoni, S., & Soekadar, S. R. (2014). Direct brain control and communication in paralysis. *Brain Topogr*, 27, 4-11.
- Farwell, L. A., & Donchin, E. (1988). Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroen Clin Neuro*, 70, 510-523.
- Guger, C., Daban, S., Sellers, E., Holzner, C., Krausz, G., Carabalona, R., . . . Edlinger, G. (2009). How many people are able to control a P300-based brain-computer interface (BCI)? *Neurosci Lett*, 462, 94-98.
- Shih, J. J., & Krusienski, D. J. (2012). Signals from intraventricular depth electrodes can control a brain-computer interface. *J Neurosci Meth*, 203, 311-314.
- Silvoni, S., Cavinato, M., Volpato, C., Ruf, C. A., Birbaumer, N., & Piccione, F. (2013). Amyotrophic lateral sclerosis progression and stability of brain-computer interface communication. *Amyotroph Lateral Scler Frontotemporal Degener*, 14, 390-396.
- Speier, W., Fried, I., & Pouratian, N. (2013). Improved P300 speller performance using electrocorticography, spectral features, and natural language processing. *Clin Neurophysiol*, 124, 1321-1328.

# Improving the Performance of Dry Based Electrodes for P300 brain-computer interfaces

J.M. Clements<sup>1</sup>, L.M. Collins<sup>1</sup>, E.W. Sellers<sup>2</sup>, S. Laughlin<sup>2</sup> and C.S. Throckmorton<sup>1</sup>

<sup>1</sup> Duke University, Durham, NC, U.S.A

jillian.clements@duke.edu, lcollins@ee.duke.edu, cst@ee.duke.edu

<sup>2</sup> East Tennessee State University, Johnson City, TN, U.S.A.

sellers@etsu.edu, bcilab@etsu.edu

## Abstract

We present a performance comparison of gel and dry based electrode caps for use with the P300 speller system, and investigate the potential for a data-driven dynamic data collection algorithm to compensate for the lower signal-to-noise ratio of P300 responses recorded via dry electrode systems. In static data collection, performance with a dry electrode system resulted in a substantial loss in performance. Using dynamic data collection, this loss in performance was reduced; however, dynamic data collection did not fully compensate for the lower SNR. Additional work is likely needed to further improve the performance of dry electrode systems.

## 1 Introduction

Despite encouraging improvements over the last decade, the P300 Speller remains primarily a research device rather than a home-based communication aid. One limitation that inhibits its widespread use outside the research lab is the complexity of setting up the system. Standard P300 systems employ gel based electroencephalographic (EEG) electrode caps to record electrical activity along the scalp. These caps require a conductive gel to be applied to each electrode in order to ensure electrical contact with the users scalp, and the caregiver must check each electrode to ensure that low impedances levels have been achieved. The process can require thirty minutes or more for an experienced lab technician. For people with locked-in syndrome who need to use the system on a daily basis, minimizing this set up time is highly desirable.

In an effort to improve the set-up time of the system, the use of dry electrode caps for P300 Spellers has been proposed. Dry electrode caps are much faster and easier to apply since they do not require any gel to be applied to the users scalp. However, despite sophisticated pre-amplification, recording EEG with dry electrodes tends to introduce a greater amount of noise into the system. The lower signal-to-noise ratio (SNR) can decrease the detectability of the P300 potential which may negatively impact both the spelling speed and accuracy of the P300 Speller. To compensate for the lower SNR, we propose using a data-driven dynamic stopping algorithm that relies on a Bayesian update process to determine the amount of data collection needed based on a probabilistic level of confidence that a character is the target [1].

In this study, we first compare the differences in performance of gel and dry based electrode caps in the standard static data collection environment for online testing. We then present the preliminary data performance results of each cap using a dynamic stopping algorithm.

## 2 Methods

Data collection for this study took part in two separate experiments. EEG responses using the standard static stopping criterion were collected from seventeen healthy participants at East Tennessee State University (ETSU), while the results from the dynamic stopping criterion were collected from ten healthy participants at Duke University. Participants in both experiments completed two P300 Speller sessions; one with a gel based electrode cap and one with a dry based electrode cap.

EEG responses were measured using electrodes positioned on standard 32-channel caps according to the International 10-20 system and connected to a computer via a 16-channel GugerTec g.USBAMP Biosignal Amplifier. The dry electrode cap utilized the GugerTec g.SAHARA active dry electrode system, which is comprised of 8-pin golden alloy electrodes. The gel electrode cap was purchased from Electro-Cap International, Inc. Eight electrodes (Fz, Cz, P3, Pz, P4, P07, P08, and Oz) were used for data collection and classification. These electrodes have been demonstrated to provide adequate information for P300 Speller communication [2]. The EEG responses were sampled at a rate of 256 Hz.

Participants were presented with a 9x8 grid of characters, which was flashed based on the checkerboard paradigm [3]. Each target character was flashed twice in a sequence of 24 flashes. The flash duration was set at 62.5 ms followed by an inter-stimulus interval of 62.5 ms, with an inter-target interval of 3.5 s. The number of sequence responses collected per target character differ between experiments and are detailed in Sections 2.1 and 2.2.

### 2.1 Static Data Collection

For the static data collection experiment, both the gel and dry based electrode cap sessions consisted of three calibration runs and three online test runs. At the start of each calibration session, the participant was asked to copy-spell three six-character words randomly drawn from a subset of words from the English language (18 characters total). For each character presented, EEG responses to five flash sequences, or 120 flashes (5 sequences x 24 flashes/sequence = 120 flashes), were collected. These data were preprocessed and features were extracted according to a method described by Krusienski *et al.* [4]. Stepwise linear discriminate analysis (SWLDA) was used to classify the extracted features. Responses to five flash sequences per target of three six-character words were collected for the three online test runs.

### 2.2 Dynamic Data Collection

The calibration runs for both sessions in the dynamic data collection experiment were gathered in a similar manner to the static data. However, due to the performance results of the static data (Section 3), we chose to collect more data to improve the weights of the classifier. Each participant in this experiment was asked to copy-spell five six-character words randomly drawn from a subset of words from the English language (30 characters total). For each character presented, EEG responses to seven flash sequences, or 168 flashes (7 sequences x 24 flashes/sequence = 168 flashes), were collected. Preprocessing techniques, feature extraction, and classification methods were identical to the static data collection experiment.

Five online test runs were collected in each session using the dynamic stopping criterion presented in Section 1. Instead of having a pre-set number of sequences for data collection, the dynamic stopping algorithm automatically determined the necessary amount of data to collect for each target character. The amount of collected data was controlled by a threshold of 90% on the probabilities that each character in the matrix was the target character. The character

probabilities were updated after each response to a flash was collected and data collection stopped once one of the character probabilities increased above 90%.

### 3 Results

The results from the static data collection experiment illustrate the impact of the lower SNR resulting from the dry electrode caps. In Figure 1, accuracy and bit rate are plotted for both the dry and gel cap sessions of both experiments. Assuming a binomial distribution on the probability of correctly selecting a set of characters, chance level accuracy was 11.1% for the static data collection experiment and 10.0% for the dynamic data collection experiment [5] [6]. Bit rate is a measure of communication systems that incorporates accuracy, speed, and the number of selectable characters presented [7].

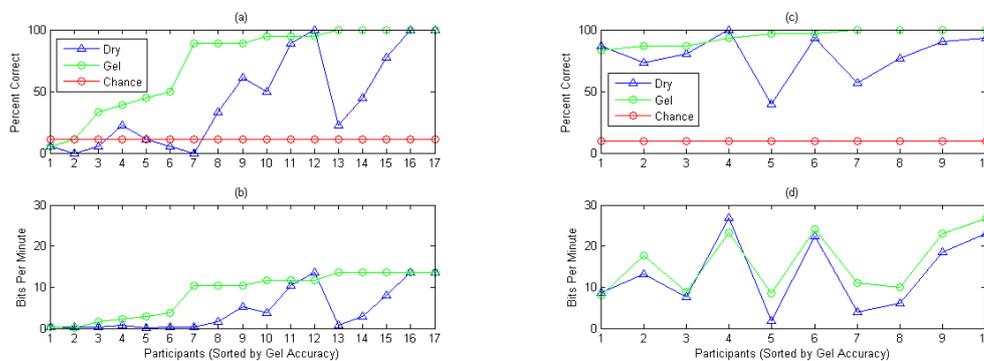


Figure 1: Comparison of dry and gel electrode caps in terms of (a) static data collection accuracy, (b) static data collection bit rate, (c) dynamic data collection accuracy, and (d) dynamic data collection bit rate. Participants are sorted by gel cap accuracy.

Figure 1(a) and Figure 1(b) display the accuracy and bit rate performance for the results of the static data collection experiment. Thirteen out of the seventeen participants performed worse when wearing the dry electrode cap than when wearing the gel electrode cap. An average decrease of 30% in accuracy and 4.1 (bits/min) in bit rate was observed from the dry electrode session of this experiment.

The results of the dynamic data collection experiment are shown in Figure 1(c) and Figure 1(d). These results suggest that the dynamic stopping algorithm may improve the performance of the dry electrode cap by collecting additional data in the online test runs to help compensate for the added noise. Although the performance of the dry electrode cap remains lower when compared to the results of the gel electrode cap, the difference between gel and dry electrode accuracy and bit rate decreased with an average decrease of 15% in accuracy and 3.5 (bits/min) in the bit rate. The increase in data collection for the dry electrodes can be observed in Figure 2. For comparison, the number of flashes that would have been collected for a static data collection of 5 sequences per target character is included. Dynamic data collection increased the amount of data collected across all participants for the dry electrode system, indicating that the data-driven dynamic data collection algorithm can detect and respond to the need for increased data collection in low SNR scenarios. However, the increased data collection was not able to fully compensate for the low SNR responses measured using dry electrodes.

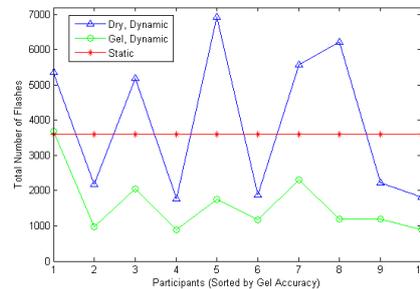


Figure 2: Amount of data collected for each participant in the dynamic data collection experiment. A simulated static data collection was included (5 sequences per target character).

## 4 Conclusion

Using a dry electrode cap with the P300 Speller would greatly reduce the complexity and time it takes to set up the system. However, dry electrodes can reduce the SNR of the recorded responses, reducing their potential utility for a home-based communication aid. Preliminary results using a data-driven dynamic stopping algorithm compensates for the additional noise by collecting more data, but performance is still lost when compared to the gel electrode cap and the time to complete the task is increased. Additional work is needed to further improve the performance of dry electrode caps in P300 Speller systems.

## Acknowledgments

This research was supported by NIH/NIDCD grant number R33 DC010470-03.

## References

- [1] C. Throckmorton, K. Colwell, D. Ryan, E. Sellers, and L. Collins. Bayesian approach to dynamically controlling data collection in p300 spellers. *IEEE Trans Neural Syst Rehabil Eng*, 21(3):508–17, 2013.
- [2] D. J. Krusienski, E. W. Sellers, F. Cabestaing, S. Bayouth, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw. A comparison of classification techniques for the p300 speller. *Journal of Neural Engineering*, 3:229–305, 2006.
- [3] G. Townsend, B. K. LaPallo, C. B. Boulay, D. J. Krusienski, G. E. Frye, C. K. Hauser, and E. W. Sellers. A novel p300-based braincomputer interface stimulus presentation paradigm: moving beyond rows and columns. *Clinical Neurophysiology*, 121(7):1101–1120, 2010.
- [4] D. J. Krusienski, E. W. Sellers, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw. Toward enhanced p300 speller performance. *Journal of neuroscience methods*, 167(1):15–21, 2008.
- [5] G. R. Mller-Putz, R. Scherer, C. Brunner, R. Leeb, and G. Pfurtscheller. Better than random? a closer look on bci results. *International Journal of Bioelectromagnetism*, 10(1):52–55, 2008.
- [6] S. Silvoni, C. Volpato, M. Cavinato, M. Marchetti, K. Priftis, A. Merico, and F. Piccione. P300-based braincomputer interface communication: evaluation and follow-up in amyotrophic lateral sclerosis. *Frontiers in neuroscience*, 3, 2009.
- [7] D. J. McFarland, W. A. Sarnacki, and J. R. Wolpaw. Braincomputer interface (bci) operation: optimizing information transfer rates. *Biological psychology*, 63(3):237–251, 2003.

# Online artifact reduction and sequential evidence accumulation enhances robustness of thought-based interaction

Reinhold Scherer<sup>1</sup>, Johanna Wagner<sup>1</sup>, Martin Billinger<sup>1</sup>  
and Gernot Müller-Putz<sup>1</sup>

Institute for Knowledge Discovery, Graz University of Technology, Graz, Austria  
{reinhold.scherer, johanna.wagner, martin.billinger, gernot.mueller}@tugraz.at

## Abstract

Electroencephalogram-based (EEG) brain-computer interface (BCI) technology should be able to robustly decode the users intent in the presence of artifacts. These requirements are particularly important for functionally disabled individuals with involuntary movements such as individuals with cerebral palsy (CP). In this paper we show that our novel online artifact reduction method in combination with a scanning mechanism that sequentially accumulates evidence for decision making enables able-bodied individuals to successfully operate a steady-state visual evoked potential (SSVEP) BCI on a Tablet computer while walking on a treadmill.

## 1 Introduction

Protocols used to calibrate and operate electroencephalogram-based (EEG) Brain-Computer Interfaces (BCIs) [3] generally demand from users the following, among others: Firstly, to carefully follow instructions and perform defined mental activity in response to a presented cue, and secondly, to avoid movements during periods of mental activity. The former allows labeling of EEG segments and thus the calibration of translation-rules. The latter aims at minimizing muscle artifact contamination. Artifacts are undesirable signals that may change EEG characteristics or even be mistakenly used as the source of control in BCI systems.

We are currently exploring the usefulness of BCI technology for users with cerebral palsy (CP). More precisely for individuals with dyskinetic forms of CP (DCP). Individuals with DCP have hypertonia and hypotonia mixed with involuntary movements and spasms. Hence, EEG will be heavily contaminated with artifacts and individuals, depending on affected muscles, may occasionally not be able to focus on and perceive cue information (e.g. look at screen).

To address the raised issues, we started working on both a novel artifact removal method for online processing of EEG data [1] and a scanning mechanism that sequentially accumulates evidence for decision making. In this paper, we introduce a steady-state visual evoked potential (SSVEP) BCI that incorporates both methods and report on first online control results.

## 2 Methods

### 2.1 Adapted online SSVEP-based BCI speller system

**Conceptual design.** The BCI is based upon a user-facing tablet with modules for both SSVEP and oscillatory activity (imagery) based communication. In the current manuscript only the SSVEP module is described. The user front-end of the system runs on an Android tablet, while

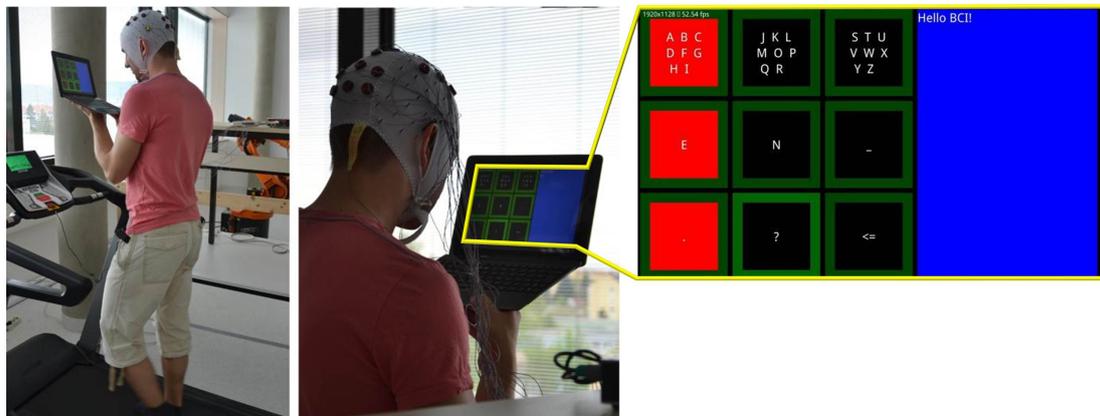


Figure 1: Participant walking on a treadmill while spelling with the SSVEP BCI system. The User Interface presented on an Android tablet (ASUS Transformer Pad TF700) is shown on the right side.

data acquisition and processing is performed on a Windows operating system computer. An operator computer can be used for monitoring and controlling experimental procedures.

Figure 1 shows the user front-end running on an Android tablet. The screen is split in two parts. The left side contains a 3x3 grid of stimulation targets, and to the right there is a speller where the participants see the letters they selected. One row or column flickers at a time in color red with a frequency of 7.5 Hz (SSVEP stimulation frequency). This frequency was selected after several tests with healthy subjects as working most robustly. The stimulation period, by default, is set to 4 seconds. The BCI evaluates data from seconds 2-4 to allow the SSVEPs to settle at a steady state. To provide feedback about target selection, target rows and columns which the user is concentrating are highlighted with given intensity (green surrounding boxes are enhanced) while flickering. The SSVEP scanning paradigm allows the user to select any number of targets with only one single stimulation frequency. One row or column flickers at a time, switching to the next row or column after 4 seconds. This takes at least six steps to uniquely select one out of nine targets. The order of stimulation was random in this study. The current design of the system allows in a first step to select one out of three different menus representing certain groups of letters. By selecting one of these menus the user is then able to select the target letter in a second step.

**Online Signal Processing and Analysis.** Processing of the signals is performed with Matlab/Simulink (MathWorks, Natick, MA, USA). Many BCI users with CP exhibit high levels of spontaneous movement, therefore a fully automated method for the removal of artifacts such as electromyogram (EMG) was developed [1] and integrated in the online BCI system. EEG is first decomposed via Wavelet decomposition methods. Independent components are then derived from decompositions of the signal and removed from the signal in instances when they exceed pre-defined thresholds. In the SSVEP based BCI, classification is performed via the canonical correlation analysis (CCA) method described in [4] and applied in [2]. Correlations are found between EEG recorded over occipital cortex and the SSVEP stimulation frequency. The largest correlation coefficient was used to identify the stimuli the user is attending to. During the evaluation period (2-4 s) of a flickering target, the mean of the correlation coefficient for that target is adaptively updated. The mean for unfocused targets reaches a lower value

than the mean of a focused target. A selection is made when the mean of a target reaches a threshold of 0.35 (determined empirically). Feedback about the progress towards selection is presented to the user in form of green colored frames around the target (Figure 1). Color black corresponds to a mean of 0, and color light green corresponds to a mean close to the threshold. The color is updated whenever the mean for that target changes by a certain step size.

The artifact correction method worked on blocks of one second of EEG data. The BCI system requires operation on smaller time steps. To ensure smooth operation and allow frequent updates of the feedback, artifact correction and BCI run as separate processes. Communication and data exchange was implemented through shared memory. Due to the block-wise cleaning of EEG, however, the BCI feedback presentation was delayed by 1 second.

## 2.2 Evaluation of the System

To evaluate the online system and test how the online artifact removal method affects the performance of the BCI we performed a study with four able-bodied male participants ( $27 \pm 3$  yr). All participants had previous experience with SSVEP BCIs. In the experiment participants had to spell with the SSVEP BCI system during different conditions which were more or less prone to produce artifacts in the EEG. Participants were: (i) sitting in a shielded, shaded and sound-proof measurement cabin with the tablet placed about 70 cm in front of them on a table (condition CAB-SIT), (ii) standing in the measurement cabin holding the tablet in their hands (condition CAB-STND), (iii) standing outside the measurement cabin on a treadmill in a room with daylight and holding the tablet in their hands (condition TRD-STND), and (iv) walking on a treadmill with slow speed (1.4-1.6 km/h) holding the tablet in their hands (see Figure 1). Condition (iv) was performed twice, once with the online artifact removal method (condition TRD-WLK, see Figure 1) and once without artifact correction (condition TRD-WLK-OFF). All other conditions were performed with artifact correction enabled. Before performing conditions (iv) participants trained spelling during walking on the treadmill for ten minutes. Conditions were randomized. Participants were asked to write one word of four letters (German words: Mund, Zaun, Auto, Haus, Werk, Baum) during each run. Participants had ten minutes respectively to complete the task.

EEG was recorded from 16 electrodes placed over cortical areas according to the international 10-20 system. Electrode positions included AFz, FC3, FCz, FC4, C3, Cz, C4, CP3, CPz, CP4, PO3, POz, PO4, O1, Oz, O2. Reference and ground were placed at position Pz and P5, respectively. The g.GAMMAsys system with g.LADYbird active electrodes (Guger Technologies, Graz, Austria) and one g.USBamp biosignal amplifier was used to record EEG at a rate of 512 Hz (Notch 50 Hz, 0.5-100 Hz band pass).

## 3 Results

Results are summarized in Table 1. Arithmetic mean values, averaged over all participants for each condition individually, show that the the number of correct and erroneous selections are in a comparable range for conditions CAB-SIT, CAB-STND and TRD-STND. Also the selections per minute are in a comparable range over these conditions. When walking on the treadmill (TRD-WLK) there is a slight decrease in correct selections and in total 0.9 selections per minute. Disabling the artifact rejection method during treadmill walking (TRD-WLK-OFF) degrades the performance to 4.7 correct selections per run and 0.67 selections per minute. The average number of scan steps needed to correctly select an item was in the range 3.8-4.8 selections.

Table 1: Number of correct/erroneous selections and the number of selections per minute for each participant (P1-P4) and condition. Additionally the average number of scan steps needed to make a correct selection (Steps) for each condition is shown.

	CAB-SIT	CAB-STND	TRD-STND	TRD-WLK	TRD-WLK-OFF
P1	07/4, 1.0	09/1, 1.6	09/2, 1.5	10/5, 1.6	12/3, 1.6
P2	11/2, 1.6	10/1, 1.4	10/2, 1.2	09/1, 0.9	5/1, 0.7
P3	07/1, 0.7	08/2, 0.9	07/0, 1.3	07/3, 0.9	2/1, 0.4
P4	08/0, 1.9	08/0, 1.5	06/1, 0.6	01/0, 0.3	0/0, --
Mean	8.2/1.7, 1.30	8.7/1.0, 1.35	8.0/1.2, 1.15	6.7/2.2, 0.90	4.7/1.2, 0.67
Steps	4.5	3.8	4.6	4.8	4.4

## 4 Discussion

The preliminary results of our study show that when EEG data is contaminated by artifacts the performance of the BCI system improves when using the developed online artifact correction method. Participant P4 had problems to spell outside the box as the light in the room was very bright and dazzled him. In total for three of the participants (P2-P4) the system performed better with artifact correction method, and for one participant (P1) the system performed equally well with and without artifact correction.

The implemented scanning mechanism sequentially accumulates evidence until a decision can be made. Hence, providing user on-demand access is an intrinsic feature of this strategy. Robustness of detection can be enhanced by optimizing scanning time and selection threshold for each individual. Selection time can be enhanced by optimizing the scanning protocol.

One big issue when designing the experimental paradigm was to find a strategy to naturally elicit artifacts. As a first step, we asked participants to walk on a treadmill and hold the tablet in their hand. Since the results of the study suggest that the developed methods work at a fundamental level, in the next step we will explore their performance with CP users and adapt methods and scanning protocols following user-center design principles.

## Acknowledgments

This work was supported by the FP7 Framework EU Research Project ABC (No. 287774). This paper only reflects the authors views and funding agencies are not liable for any use that may be made of the information contained herein.

## References

- [1] I. Daly, R. Scherer, M. Billinger, and G. Müller-Putz. FORCE: Fully Online and automated artifact Removal for brain-Computer interfacing. *IEEE Trans Neural Systems Rehab Eng*, accepted.
- [2] P. Horki, C. Neuper, G. Pfurtscheller, and G. Müller-Putz. Asynchronous steady-state visual evoked potential based bci control of a 2-dof artificial upper limb. *Biomed Tech*, 55(6):367–374, 2010.
- [3] R. Scherer, G. R. Müller-Putz, and G. Pfurtscheller. Flexibility and practicality: Graz brain-computer interface approach. *International Review of Neurobiology*, 86:119–131, 2009.
- [4] G.A.F. Seber. *Multivariate Observations*. Wiley, 1984.

# P300-based BCI to drive an assistive device: Usability Evaluation with Stakeholders

Francesca Schettini<sup>1,2</sup>, Angela Riccio<sup>1,3</sup>, Mario Caruso<sup>2</sup>;  
Massimo Mecella<sup>2</sup>, Donatella Mattia<sup>1</sup>, Febo Cincotti<sup>1,2</sup>.

<sup>1</sup> Neuroelectrical Imaging and BCI Lab. IRCCS Fondazione Santa Lucia, Rome, Italy;

<sup>2</sup> Department of Computer, Control, and Management Engineering “Antonio Ruberti”, “Sapienza”  
University of Rome; <sup>3</sup> Department of Psychology, “Sapienza”, University of Rome  
f.schettini@hsantalucia.it

## Abstract

This work aims at evaluating the usability of an assistive device for people with amyotrophic lateral sclerosis, designed to permit a wide range of input modalities to support communication and interaction with the environment. Nine stakeholders (assistive technology experts, medical doctors and BCI researchers) were involved in the study, which included four experimental conditions, each considering a different input modality: touchscreen, buttons (scan mode), headtracker, and a P300-based brain computer interface. The latter exhibited a lower effectiveness and efficiency with respect to the other input devices. However no differences were found among the four conditions in terms of ease of access, ease of use, ease of understanding, usefulness and satisfaction.

## 1 Introduction

The “user-centered design” (UCD; ISO 9241-210) implies that end-users play an active role in the device design and development following an iterative process. The UCD approach has been recently introduced in the Brain Computer Interface (BCI) field of research (Zickler et al. 2011; Kubler et al. 2013). The Brindisys project ([www.brindisys.it](http://www.brindisys.it)) recently deployed a prototype assistive device (AD) providing functionalities which are seamlessly accessible through several conventional/alternative input channels, among which a P300-based BCI, to enhance/allow basic communication and environmental interaction of people with Amyotrophic Lateral Sclerosis (ALS). The multimodal access to this AD prototype provides end-users with an adaptable system coping with the different stages of the disease. This study aims at evaluating the usability of the proposed prototype AD by involving three different categories of stakeholders, namely medical doctors working with patients with ALS, psychologists experienced in the assistive technology (AT) field and BCI researchers with previous experience with ALS. In fact, stakeholders have a broad knowledge about the different stages of the disease and about the conventional ATs currently adopted by end-users, making their opinion extremely valuable for the evaluation of a new AT.

## 2 Methods

### 2.1 Description of the Assistive Device Prototype

The AD prototype can be operated through several input modalities: touchscreen, hardware buttons, headtracker and a P300-based BCI. To ensure portability and affordability, the prototype was developed on a 10" tablet and the software written in Java and C++ running on the Windows operating system (Caruso et al. 2013). The selection of available functionalities has been performed in agreement with the results of a preliminary user survey (Schettini et al. 2014). In the domain of interpersonal communication, a Graphical User Interface (GUI) running on the tablet, provided three main functions: (i) an alarm sound to draw the attention of a caregiver; (ii) a simple text editor, for both face-to-face and remote (SMS) communication; and (iii) an interface to select predefined sentences for quick communication. In the domain of environmental control, functionalities available on the GUI included on/off switching of lights and appliances, TV and music players remote control. These functionalities required the deployment of a "domotic kit", i.e. dedicated hardware packed in briefcase, which includes: a WiFi router (communication with the tablet), three controllable mains sockets (appliances), an infrared controller (TV remote), and an UMTS router (internet monitoring).

As for the BCI input, stimulation timing and data acquisition were managed by the BCI2000 framework. A custom software program managed the communication between BCI2000 and the GUI, and generated the visual stimuli (green grids), necessary to generate evoked potentials, on top of the aforementioned input-independent GUI.

### 2.2 Participants and Experimental Protocol

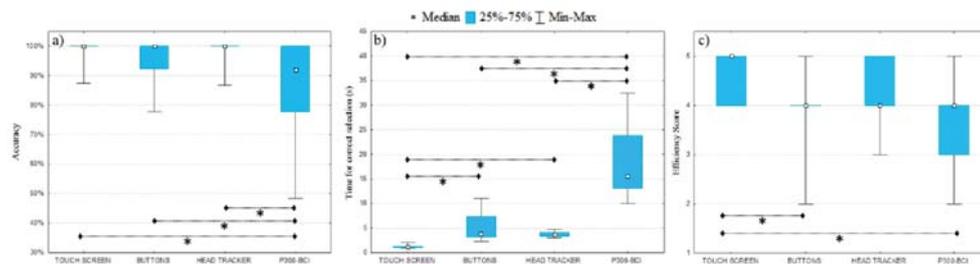
Nine stakeholders (mean age = 37.8±5.6) were involved in the study. Three of them were medical doctors working with patients with ALS, three were psychologists with experience in AT for people with ALS and three were BCI researchers (an engineer, a neuropsychologist and a neurophysiology technician) with experience in experimentation with persons with ALS.

The experimental protocol consisted of a single session in which the stakeholders evaluated the AD prototype in four conditions: (i) touch screen, (ii) two buttons (one used to scan the icons, the other one to select the icon of interest), (iii) a head tracker with "dwell" selection and (iv) the P300-based BCI. At the beginning of the session, each participant watched a video-tutorial describing the AD prototype functionalities and how to use the different input devices. Participants were then given 10 minutes to familiarize with the prototype at their will, using a touchpad. The four conditions were presented in a randomized order. Participants were required to complete a communication task and an environmental control task twice (2 runs for task, 4 runs for condition). A minimum of 12 and 10 selections were required to complete the first and the second task, respectively. The experimenter did not provide participants with indications about the sequence of individual selections (menu navigation) to complete the tasks. At the end of each condition, stakeholders were requested to fill a questionnaire about ease of access, ease of use, ease of understanding, usefulness, satisfaction and perceived efficiency. Each variable was scored from 1 to 5, by means of a likert scale. Participants were requested to evaluate the system taking into account the broadest needs of potential users with ALS. Before performing the BCI condition, each participant carried out 6 calibration runs (no feedback was provided), selecting four items from grids of three different sizes (2 by 2, 4 by 4, and 6 by 6). Parameters of the linear classifier were extracted applying a stepwise linear discriminant analysis on the ensemble of the calibration runs. The number of stimuli repetitions to use during the online tasks was defined by means of a 6-fold cross-validation on the calibration runs, and set as the minimum number of repetitions needed to achieve the highest accuracy.

## 2.3 Usability assessment

According to the UCD methodology we assessed the overall usability of the AD prototype for each condition within three domains: (i) effectiveness, assessed as the number of correct selections performed (including selections needed to correct errors) divided by the total number of selections performed to complete the task; (ii) efficiency assessed as the average time for correct selection, i.e. the total time (in seconds) to complete the task divided by the number of correct selections; (iii) and satisfaction, assessed by means of a questionnaire. For effectiveness and efficiency, a comparison between the first and the second run has been performed for each task and subject, in order to evaluate the learnability of the AD prototype as well as the overall simplicity in getting acquainted with it. A non-parametric one way Friedman ANOVA for repeated measures was performed for each variable. A Wilcoxon test was performed as post-hoc analysis for the variables with a significant F test.

## 3 Results



**Figure.1 a) Accuracy achieved during the tasks in all conditions; b) Time for correct selection in all conditions; c) Efficiency perceived by stakeholders**

With regard to effectiveness, (Figure 1.a.), the ANOVA showed significant differences among the four conditions ( $\chi^2=25.59$ ;  $p<.01$ ). The P300-BCI exhibited significant lower accuracy ( $p<.05$ ) with respect to other conditions. Figure 1.b shows the values of time for correct selection. The ANOVA pointed out significant differences among the four conditions ( $\chi^2=49.02$ ;  $p<.01$ ). P300-based BCI exhibited higher ( $p<.05$ ) time for correct selection with respect to the other conditions, the touch screen being the fastest condition ( $p<.05$ ), as assessed by the post-hoc test.

These differences influenced the efficiency perceived by participants and assessed by means of the questionnaire. Indeed, Friedman ANOVA showed a significant difference among the four conditions in the efficiency variable ( $\chi^2=7.9$ ;  $p<.05$ ). As a result of the post-hoc test, efficiency of the touch screen condition resulted higher than the two buttons condition ( $p<.05$ ) and the BCI condition ( $p<.05$ ). No significant differences were found by analyzing the other 5 variables (ease of access, ease of use, ease of understanding, usefulness and satisfaction).

With regard to the comparison between the first and the second run, the ANOVAs point out significant differences neither in terms of accuracy increment ( $\chi^2=1.74$ ;  $p=.62$ ) nor in terms of time for correct selection ( $\chi^2=2.20$ ;  $p=.53$ ) among the four conditions (Table 1).

The analysis of feedbacks about the overall usability of the system highlighted some usability issues regarding both the structure of the GUI (e.g. the pause function resulted confusing), and some weak points with specific input devices (e.g. the alarm call menu was not well designed for the buttons input). Problems identified in usability were not directly related to the inclusion of the BCI as control input.

**Table 1. Comparison between the first and the second run. White and gray cells denote an increment and a decrement of the value from the first to the second run respectively.**

	Accuracy				Time for Correct selection (s)			
	Touch Screen	Buttons	Head tracker	P300-BCI	Touch Screen	Buttons	Head tracker	P300-BCI
Mean	0,05%	2,51%	0,37%	3,19%	0,29	1,05	0,18	1,55
Median	0,00%	0,00%	0,00%	2,05%	0,26	1,14	0,11	0,22
I Quartile	0,00%	9,38%	0,00%	3,33%	0,18	0,05	0,01	2,41
III Quartile	0,00%	0,00%	0,00%	5,53%	0,33	2,43	0,39	1,99
Min Value	0,00%	12,06%	6,67%	28,48%	0,12	1,98	0,23	10,69
Max Value	0,42%	10,00%	3,33%	12,50%	0,73	3,40	0,64	4,19

## 4 Discussion

Despite not conclusive, these results confirm the feasibility of a single AD accessible from a broad range of input modalities, including a BCI. In fact, no differences in terms of simplicity in getting acquainted with the prototype were pointed out by stakeholders. The P300-BCI exhibited lower efficiency and effectiveness with respect to the other access conditions, but this did not affect the ease of access, ease of use, ease of understanding, usefulness and satisfaction perceived by participants with BCI. Indeed no differences were found between the four input conditions. This confirms that the overall system has been perceived by stakeholders as a single multimodal AD system.

## Acknowledgments

The work is partly supported by the Italian Agency for Research on ALS – ARiSLA, project “Brindisys”.

## References

- Caruso, M. et al., 2013. My-World-in-My-Tablet: An Architecture for People with Physical Impairment. In M. Kurosu, a c. di *Human-Computer Interaction. Interaction Modalities and Techniques*. Berlin, Heidelberg: Springer Berlin Heidelberg, pagg. 637–647.
- Kübler, A., et al., 2013. Applying the user-centred design to evaluation of Brain-Computer Interface controlled applications. *Biomed Tech (Berl)*. doi:10.1515/bmt-2013-4438
- Schettini, F. et al.; 2014. An assistive device with conventional, alternative and Brain-computer interface inputs to enhance interaction with the environment for people with Amyotrophic Lateral Sclerosis: a feasibility and usability study. *Archives of Physical Medicine and Rehabilitation* in press.
- Zickler, C. et al., 2011. A brain-computer interface as input channel for a standard assistive technology software. *Clinical EEG and neuroscience*, 42(4), pag.236–244.

# Frontal Alpha Asymmetry Neurofeedback for Brain-Computer Interfaces

Marc Cavazza<sup>1</sup>, Gabor Aranyi<sup>1</sup>, Fred Charles<sup>1</sup>, Julie Porteous<sup>1</sup>,  
Stephen Gilroy<sup>1</sup>, Gilan Jackont<sup>2</sup>, Ilana Klovatch<sup>2</sup>, Gal Raz<sup>2</sup>,  
Nimrod Jakob Keynan<sup>2</sup>, Avihay Cohen<sup>2</sup> and Talma Hendler<sup>2</sup>

<sup>1</sup> Teesside University, School of Computing, Middlesbrough, United Kingdom

<sup>2</sup> Functional Brain Center, Tel Aviv Sourasky Medical Center, Tel Aviv, Israel  
m.o.cavazza@tees.ac.uk, hendlert@gmail.com

## Abstract

We report the development of an affective BCI based on frontal alpha asymmetry neurofeedback, which has previously been used in clinical experiments. Our results evidence a pattern of high-performance for some subjects, combined with high illiteracy, with 52.3% of subjects succeeding in the neurofeedback task. We suggest that individual asymmetry baseline values may be one of the factors explaining BCI illiteracy in this context.

## 1 Introduction and Rationale

Frontal alpha EEG asymmetry (henceforth FAA) was identified by Davidson [2] as a correlate of approach/avoidance, a high-level emotional dimension that plays an important role in many aspects of human psychology. FAA was later demonstrated to be a marker of clinical depression, which led to the proposed use of neurofeedback (NF) techniques for depression therapy [6], demonstrating this marker to be amenable to NF.

We have explored the use of FAA to implement an affective BCI for interactive media based on NF [5]. In this system, a virtual narrative is generated in the form of a real-time 3D animation, in which the user can intervene to support the central character by expressing positive thoughts, which in turn will drive the evolution of the narrative towards a happier ending for the character. A previous feasibility pilot has demonstrated a success rate of 50% for an all-male sample of 12 subjects. Furthermore, these experiments were carried out under an EEG/fMRI installation to provide a first level of anatomical validation, which confirmed our original hypothesis [4]. In this paper we present the results of follow-up work in which we explore the determinants of successful NF, using a larger number of subjects and analysing training data.

Following previous work we decided to use the  $A_2$  score  $(F_4 - F_3)/(F_4 + F_3)$  (with  $F_4(R)$  and  $F_3(L)$  electrodes) as a measure of FAA for NF input. However, because of the intrinsic fluctuations in  $A_2$  and the need to provide a stable signal to support visual feedback (here, the central character's skin colour saturation) that reduces jitter, we used a 4-point moving average as a simple low-pass filter (henceforth  $MA_2$ ). As a baseline, we used the  $A_2$  average value obtained during a 2min calibration period (which was discussed by Allen et al. [1] as the minimum reliable calibration time).

Since our NF signal is visual and continuous, we defined a mapping function between each subject's individual baseline and an empirically determined "maximum" value for  $MA_2$  onto the 0-100% saturation range. The rationale for determining such a maximum was that the activity in the right hemisphere cannot be reduced infinitely hence the  $A_2$  score should be seen as asymptotic. We conducted a calibration experiment with 16 subjects to empirically determine

the mapping of  $MA_2$  values to saturation, where we used the variation in the subject's  $MA_2$  values to determine maximum saturation. In order to avoid a few high scores raising the value where maximum saturation occurs (thereby decreasing the amount of feedback received for scores that are likely to occur more frequently) we decided to define the maximum feedback below the maximum of  $MA_2$  values ( $\sim 0.7$ , across all subjects). When applying a  $mean + 1.64SD$  filter<sup>1</sup>, the average difference between subjects' maximum  $MA_2$  value and their  $MA_2$  mean (i.e. baseline) was 0.23 ( $SD = .06$ ). It should be noted that this finding is consistent with Zotev et al. [8] setting variations to  $A_2$  of 0.2 as a success criterion in their own NF experiments. This led us to determine the  $MA_2$  value resulting in maximum saturation for each subject as  $\min(0.7, baseline_{subject} + 0.2)$ , making saturation scores subject specific.

The next step was to define a success score for a NF trial. This consists in the ability to increase  $MA_2$  and to sustain this increase over a non-trivial amount of time (so as to distinguish it from spontaneous fluctuations). In order for the success score to be normalised across subjects, we used the above saturation scores instead of raw  $MA_2$  since they already incorporate the individual's  $A_2$  baseline. We assign a maximum score of 100 to an ideal result of 100% saturation over the full 30s NF window. This score is meant to account for any combination of above-threshold amplitude and time, essentially integrating above-threshold values over time. For instance, a score of 10 would correspond to either 100% saturation over 3s, 50% saturation over 6s, 30% saturation over 10s ... this is meant to cover the very different NF patterns observed across subjects, while appearing a stricter criterion than that of [3] who imposed above threshold values to be sustained for 500ms as well as Rosenfeld [6] and its "number of hits above threshold" metric.

Finally, a trial can only be considered successful if  $MA_2$  actually increases during NF over its previous resting value. Following several authors (including Zotev et al. [8] who have used as a reference the average asymmetry scores during the preceding 40s), we have defined as a criterion for success for a NF trial that it reaches a score of 10, such score being also higher than that of the preceding resting period of 15s in-between trials (scores being normalized for respective durations of resting and NF periods during trials), as shown in Figure 1.

## 2 Methods

We recruited 36 subjects (17 male, 19 female), whose average age was 30.4 years ( $SD = 9.25$ ,  $range : [20, 52]$ ). Experiments were approved by our local ethics committee, and subjects were issued detailed consent forms; all data were anonymised. Subjects were sat in a comfortable armchair in a quiet room with dimmed lighting, and they were given instructions on how to relax to minimise muscular artefacts as well as avoiding blinking as much as possible. They were introduced to the concept of a NF loop in simple terms, with the character's skin colour saturation introduced as a visual indicator of the magnitude of mental support (as shown in Figure 2). Instructions were deliberately generic ("*express positive thoughts towards the character.*"), in order to avoid influencing users cognitive strategies towards any implicit or explicit one.

Each training block consisted of a 30s NF session preceded by a 15s resting period during which subjects were instructed to relax and remain staring at a black screen (Figure 2-a). Each subject performed 12 successive training trials for a total duration of 9min. All subjects completed the training session. We defined a subject as successful in training if s/he succeeded in 6 out of 12 trials. After the training session, each subject participated in one session of the

<sup>1</sup>Filtering out the top 5% values assuming near normality after [6].

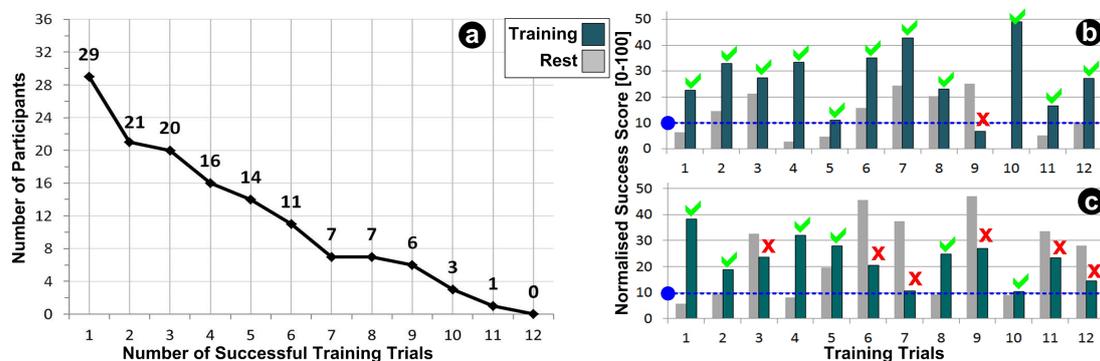


Figure 1: Training results. (a) Number of successful training trials per participant, (b) comparing scores to previous resting periods for each trial for one subject, and (c) for another subject, rest values are greater than NF for trials 3, 6, 7, 9, 11, 12, which are all rated “failed”. The blue line represents success threshold as a score of 10 (see text).

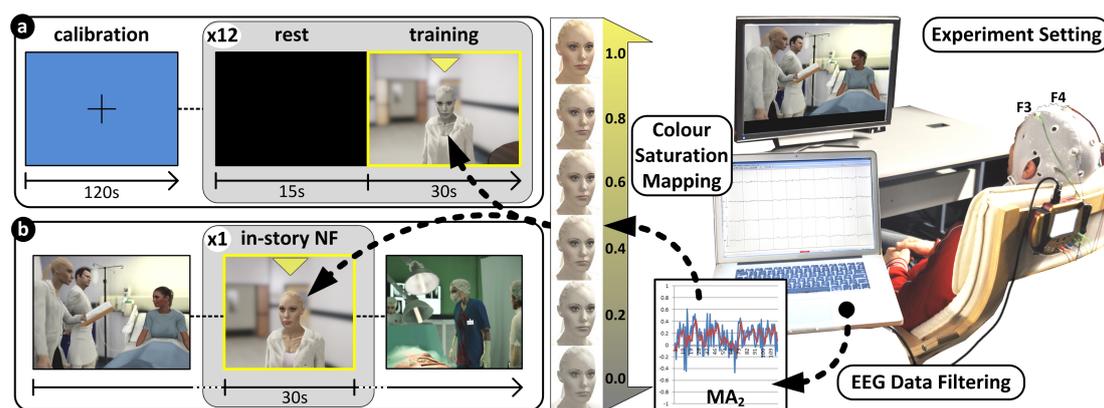


Figure 2: Experiment setup. (a) Initial calibration, followed by 12 training trials pairs (Rest/NF). (b) BCI-based narrative. During NF,  $MA_2$  is mapped to the colour saturation of the character in need of support by the user (see text).

BCI-enabled Interactive Narrative [4] (Figure 2–b).

EEG data was acquired using an 8-channel Brain Products V-Amp system. Data was recorded at a sampling rate of 250Hz and collected on a PC running BrainVision RecView<sup>2</sup>. Alpha band (8–12 Hz) power was extracted online from electrodes  $F_3$  and  $F_4$ , sampled at 1Hz with a reference electrode at  $FCz$ . The pre-processing algorithm was compiled from Matlab R2013b to Microsoft .NET, so that it could be executed within the BrainVision RecView EEG Recorder system. The Matlab.NET compiled DLL calculated the  $A_2$  instantaneous value once filtered through the calculation of  $MA_2$  and passed this value back to the NF system, which in turn was used to drive the feedback visuals.

<sup>2</sup><http://www.brainproducts.com/>

### 3 Results and Discussion

Figure 1-a plots the number of subjects that performed successfully on a given number of trials during NF training. The most striking finding is to observe FAA NF as a high-performance, high-illiteracy BCI: the average in-story score for successful subjects is actually 20, twice the success threshold we have defined, together with a global success of 52.3%. However, the proportion of subjects failing the task (although not all failure can be assimilated to BCI illiteracy) is much higher than previous reports of BCI illiteracy, which were in the region of 20-30% [7]. Since BCI illiteracy is considered to be specific to the BCI technique considered, we investigated the role of the  $A_2$  baseline as a possible determinant of illiteracy by calculating the biserial correlation between the individual threshold value and training success,  $r_b = -.69, p = .002$ , which suggested a strong and statistically significant negative relationship between baseline and training success. On the other hand, the contribution of training, or lack thereof, to the observed illiteracy failure rate figures, is more difficult to assess without extending our experiments to the same level of training as the one described for clinical applications (previous authors having reported hours of training over multiple sessions). However, our ethical approval was limited to one short session, as these are unlikely to alter subjects trait variables over a prolonged period. Another possible direction to explore to improve performance consists in the baseline measure itself. Following previous literature, we have alternated epochs of eyes open and closed during baseline measures, but we have observed by retrospective analysis<sup>3</sup> that only using eyes open signals (as Davidson reported for experiments with films [2]) would have resulted in a lower baseline, hence potentially higher success rates.

### References

- [1] J. J. B. Allen, J. A. Coan, and M. Nazarian. Issues and assumptions on the road from raw signals to metrics of frontal EEG asymmetry in emotion. *Biological Psychology*, 67(1):183 – 218, 2004.
- [2] R. J. Davidson, P. Ekman, C. D. Saron, J. A. Senulis, and W. V. Friesen. Approach-withdrawal and cerebral asymmetry: Emotional expression and brain physiology: I. *Journal of personality and social psychology*, 58(2):330, 1990.
- [3] Á. M. Dias and A. van Deusen. A new neurofeedback protocol for depression. *The Spanish Journal of Psychology*, 14(01):374–384, 2011.
- [4] S. W. Gilroy, J. Porteous, F. Charles, M. Cavazza, E. Soreq, G. Raz, L. Ikar, A. Or-Borichov, U. Ben-Arie, I. Klovatch, and T. Hendler. A brain-computer interface to a plan-based narrative. In *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence*, pages 1997–2005. AAAI Press, 2013.
- [5] C. Mühl, A.-M. Brouwer, N. van Wouwe, E. L. van den Broek, F. Nijboer, and D. K. J. Heylen. Modality-specific affective responses and their implications for affective BCI. In *Proceedings of the Fifth International Brain-Computer Interface Conference*, pages 120–123. Verlag der Technischen Universität, 2011.
- [6] J. P. Rosenfeld, G. Cha, T. Blair, and I. H. Gotlib. Operant (biofeedback) control of left-right frontal alpha power differences: Potential neurotherapy for affective disorders. *Biofeedback and Self-Regulation*, 20(3):241–258, 1995.
- [7] C. Vidaurre and B. Blankertz. Towards a cure for BCI illiteracy. *Brain Topography*, 23(2):194–198, 2010.
- [8] V. Zotev, R. Phillips, H. Yuan, M. Misaki, and J. Bodurka. Self-regulation of human brain activity using simultaneous real-time fMRI and EEG neurofeedback. *NeuroImage*, 85:985–995, 2014.

<sup>3</sup>Repeated measures t-test confirmed that eyes-open baseline ( $M = .25, SD = .13$ ) was significantly lower than eyes-closed baseline ( $M = .35, SD = .14$ ),  $t(35) = 6.61, p < .001, r = .75$ (large).

# Towards Aphasia Rehabilitation with BCI

Michael Tangermann<sup>1</sup>, Norah Schnorr<sup>2</sup> and Mariacristina Musso<sup>2</sup>

<sup>1</sup> University of Freiburg, Dept. Computer Science,  
Excellence Cluster BrainLinks-BrainTools, Freiburg, Germany  
[michael.tangermann@blbt.uni-freiburg.de](mailto:michael.tangermann@blbt.uni-freiburg.de)

<sup>2</sup> University of Freiburg, Dept. Computer Science,  
Excellence Cluster BrainLinks-BrainTools, Freiburg, Germany

<sup>3</sup> University Medical Center Freiburg, Dept. Neurology,  
Excellence Cluster BrainLinks-BrainTools, Freiburg, Germany

## Abstract

Cognitive states can be monitored and exploited in real-time by brain-computer interface (BCI) systems. Comparable to motor rehabilitation after stroke, we propose a simple BCI-supported paradigm for the cognitive rehabilitation of speech production deficits in aphasia patients. The paradigm thrives to close the loop between top-down execution and bottom-up perception, while constraining the execution of a rehabilitation trial to BCI-detectable attention patterns. In an offline study, the basic characteristics of the novel paradigm were explored with  $n=9$  normal hearing healthy students in a free-field spatial auditory setup. Event-related potentials (ERPs) of the electroencephalogram were analyzed and the possibility for an online classification of target vs. not-target words was explored. Despite differing in delay from ERP components evoked by artificial tones, the ERPs evoked by words can be classified on a comparable level in single trial, encouraging future tests of the attention-constraint paradigm in closed-loop.

## 1 Introduction

The inability to speak certain words (e.g. after stroke) can have different causes, ranging from a lesion within the dual loop model of language [4, 6] to disturbed attention mechanisms [1, 3]. Several cognitive models propose that word production may derive from high and rapid online interaction between top-down and bottom-up processing at different levels (i.e. formulation of conceptual representation into a linguistic form and an articulation level).

It is intriguing to engage both interaction directions in a rehabilitation approach. Practically this requires to close the loop from (1) the willful attempt of a word production or at least the concentration onto a word to (2) producing the sound of the intended word and thereby providing sensory feedback of the (artificially) spoken word. Transferring methods from spatial auditory BCI paradigms, we make a first attempt towards such an attention-constrained rehabilitation paradigm: a BCI system, detects attention onto a difficult-to-produce target word among a sequence of several words. Upon detection, a trial ends by playing the target word, which closes the loop via auditory feedback.

## 2 Methods

### 2.1 Participants

Two pilot sessions were performed. Data thereof was used to fine-tune experimental parameters, but excluded from further analysis. After providing written informed consent,  $N=9$  healthy students (5 male and 4 female, age 22–26 yrs, no history of hearing defects, native German speakers) participated in the study.

## 2.2 Experimental Setup

Subjects were seated in a ring of six speakers (AMUSE paradigm, [5]). A screen provided a fixation cross, indicated periods for relaxation and active trials. At trial start, subjects received static visual and auditory information about the current target direction.

**Stimuli and Trial Structure.** The following one-to-one relation between sentences and end-of-sentence words had been learned by subjects during a familiarization phase:

Wir müssen putzen, überall ist ... Dreck.  
 Komm, wir gehen raus an die frische ... Luft.  
 Zum Frühstück mache ich mir ein belegtes ... Brot.  
 Wir machen Picknick auf der Wiese im ... Gras.  
 Alle Kinder trinken gerne ... Saft.  
 Er hat schon wieder gewonnen - er hat so viel ... Glück.

Prior to presenting a rapid stimulus sequence consisting of the six end words, one of the sentences was cued from a loudspeaker. It indicated the target direction of the current trial. (Target directions were pseudo-randomized between trials.) The last word of the sentence was missing, and subjects were instructed to focus their attention to complement the missing word. The subsequent rapid sequence of 90 word stimuli per trial was presented with a stimulus onset asynchrony (SOA) of 175 ms. Comprising 245 ms each, word stimuli slightly overlapped in time. During a run, the six words were tied one-to-one to the loudspeakers. This mapping was changed between runs and balanced over runs.

**Structure.** Subjects were familiarized with a six-class AMUSE spatial auditory setup by first slowly. During the familiarization phase only, subjects had to count and indicate by finger movements when they had perceived target stimuli. Subsequently, three main blocks were performed with an SOA of 175 ms and one last block with 300 ms. Only data recorded with 175 ms SOA entered the analysis. Each block consisted of six runs, every run of six trials. A trial contained 90 stimuli (15 iterations of the six stimuli). As each iteration contained only one target stimulus and five non-target stimuli, this procedure totaled to an amount of  $3 * 6 * 6 * 15 * 1 = 1620$  target epochs and  $3 * 6 * 6 * 15 * 5 = 8100$  non-target epochs.

**Electroencephalogram Recordings.** During a single session, electroencephalogram (EEG) was recorded from 63 passive Ag/AgCl electrodes (EasyCap / BrainAmpDC amplifiers). The electrodes were placed according to the 10-20-system and referenced against the nose. In addition, one channel was placed below the right eye. Signals were sampled at 1 kHz, band-filtered between 0.01 Hz and 250 Hz and stored for offline processing.

## 2.3 Offline Data Analysis

Data of one subject had to be removed due to excessive eye blinks. Data of the remaining eight subjects was bandpass-filtered between 0.5 Hz and 20 Hz. For each stimulus, an epoch was windowed at [-175 ms , 1200 ms] relative to stimulus onset. Epoch offsets were corrected based on the interval [-175 ms , 0 ms]. Epochs were removed, if the variance of their amplitude exceeded three standard deviations or if max-min amplitudes exceeded  $80 \mu V$ . On average, 37 target epochs and 182 non-target epochs were dropped per subject. Remaining epochs were averaged over subjects, but separately for target- and non-target stimuli to reveal the grand-average ERP structures (time courses and scalp maps) of the word paradigm.

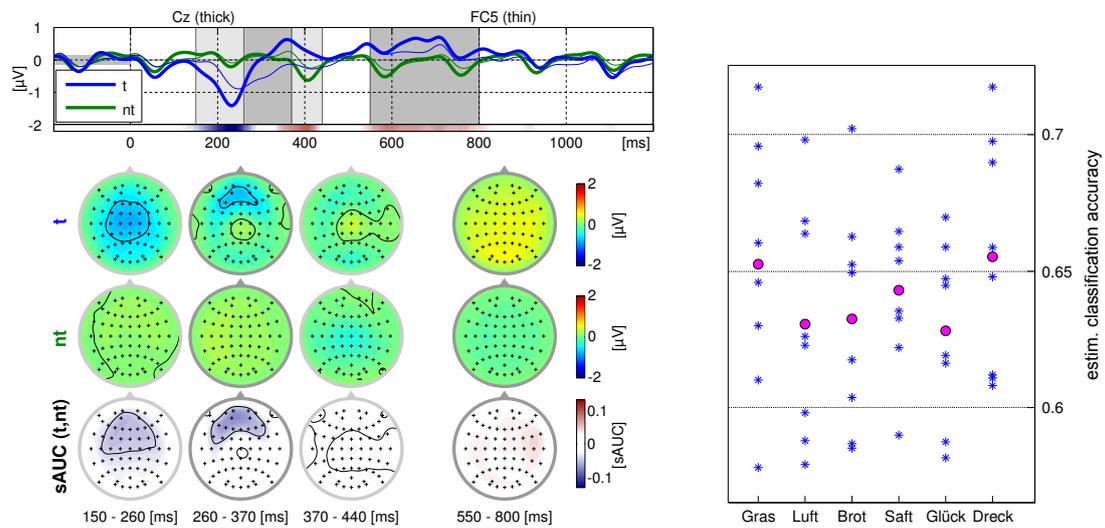


Figure 1: **Left:** evoked ERP components (grand average,  $n=8$ ). TOP ROW: average time series of EEG channels Cz and FC5 for target (blue) and non-target (green) epochs. A color bar below the time series indicates the class-discriminative information (signed AUC) values contained over time in the Cz channel. MIDDLE ROWS: average potential maps evoked by target stimuli (t) and non-target stimuli (nt) of three earlier intervals and a later one (see corresponding light and dark gray shadings of top row). BOTTOM ROW: maps of channel-wise class-discriminative information content as indicated by signed AUC values of the same intervals. **Right:** Classification accuracies estimated for six classifiers trained separately for each word. Asterisks indicate the eight subject-specific accuracies, grand average results are indicated by colored circles.

Prior to classification with a shrinkage-regularized linear discriminant analysis (LDA), 15 average potentials (six intervals of 20 ms duration between 170 ms and 290 ms and nine intervals of 60 ms length between 290 ms and 830 ms) were extracted from each EEG channel, resulting in  $15 \times 63 = 945$  dimensions per epoch. Classification values are given relative to a chance level of 0.5 and were estimated by five-fold chronological cross-validation.

To get an overview over the novel paradigm and to estimate its single-trial classification level, a binary target vs. non-target classifier was trained for each subject. The average performance over all subjects is reported.

In addition, six separate binary classifiers were trained per user. For each classifier, only epochs from one of the six words were included. Though this strategy reduced the available amount of data by a factor of 1/6 and thereby lead to classifiers of decreased performance, it allowed us to investigate systematic differences between words.

### 3 Results

The evoked ERP components for target- and non-target words varied strongly between users. For this reason, the left plot of Figure 1 must be taken with caution. It shows the grand average ERP response for two channels and four selected time intervals. Over subjects, the word stimuli elicited an attention-modulated, class-discriminative fronto-central to fronto-bilateral negativity

around 230 ms post stimulus onset, and a subsequent class-discriminant central positivity ending at approx. 800 ms with peak amplitudes at more (right-) lateralized electrodes. Compared to the original AMUSE setup, which also shows these two class-discriminative features, the latencies on average are longer for the word stimuli. The excessive length of the class-informative late positivity in fact is caused by averaging over shorter corresponding intervals of single subjects, which strongly differ in latency and duration.

Classification analysis reveals, that the word stimuli can be classified with accuracies between 65 % and 77 %. The grand average performance over subjects is 71.2 %, which slightly improves over the brisk AMUSE tone stimuli [5] (68.5 %).

Direction-related performance differences with word stimuli by and large replicate findings of the AMUSE reports (not shown). Thus it is interesting to see, if the six words, which have been randomized over the loudspeaker directions, result in similar classification performances. The right plot of Figure 1 shows the results of six classifiers trained specifically on epochs of one word each. Fortunately the spread of binary classification accuracies between words is small (between 0.625 and 0.655).

## 4 Discussion and Future Work

The data presented is an indicator for the effectiveness of the proposed attention-constrained word paradigm to elicit target and non-target responses, such that profound problems are not expected when transferring the paradigm to an online setup. As next steps, the trade-off between delayed high-quality feedback (cp. to current auditory BCI routines) and immediate, but more unreliable feedback must be addressed. Furthermore, the feasibility of the paradigm with individually chosen word sets needs investigation, as well as the use of longer words. The far goal is to test the potential effectiveness of the attention-constrained aphasia rehabilitation training procedures, which of course can only be verified in randomized, blinded online studies with the target user group of chronic stroke patients. Tight time budgets in clinics or for home-use should be accommodated for by reducing the calibration time to a minimum. A combination of the attention-constrained paradigm with recently proposed transfer learning approaches in combination with unsupervised adaptive classification [2] may provide a solution.

**Acknowledgments** The authors appreciate their funding by the German Research Foundation within the Excellence Cluster BrainLinks-BrainTools (grant *EXC 1086*).

## References

- [1] F. Burgio and A. Basso. Memory and aphasia. *Neuropsychologia*, 35(6):759–766, 1997.
- [2] P.-J. Kindermans, M. Tangermann, K.-R. Müller, and B. Schrauwen. Integrating dynamic stopping, transfer learning and language models in an adaptive zero-training erp speller. *J Neural Eng*, 11(3):035005, 2014.
- [3] L. L. Murray. Review attention and aphasia: Theory, research and clinical implications. *Aphasiology*, 13(2):91–111, 1999.
- [4] J. P. Rauschecker and S. K. Scott. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nature neuroscience*, 12(6):718–724, 2009.
- [5] M. Schreuder, B. Blankertz, and M. Tangermann. A new auditory multi-class brain-computer interface paradigm: Spatial hearing as an informative cue. *PLoS ONE*, 5(4):e9813, 2010.
- [6] C. Weiller, T. Bormann, D. Saur, M. Musso, and M. Rijntjes. How the ventral pathway got lost—and what its recovery might mean. *Brain and language*, 118(1):29–39, 2011.

# A new descriptor of neuroelectrical activity during BCI–assisted Motor Imagery training in stroke patients

Manuela Petti<sup>1,2</sup>, Donatella Mattia<sup>1</sup>, Floriana Pichiorri<sup>1</sup>, Laura Astolfi<sup>1,2</sup>, Febo Cincotti<sup>1,2</sup>

<sup>1</sup>IRCCS Fondazione Santa Lucia, Rome, Italy. <sup>2</sup>Dept. of Computer, Control, and Management Engineering, Univ. of Rome “Sapienza”, Italy.  
manuela.petti@uniroma1.it

## Abstract

Recent BCI applications stroke motor rehabilitation have raised important concerns regarding the type of brain activity which one would train in agreement with an evidence-based approach in rehabilitation. In this pilot study we proposed an offline analysis on EEG data acquired during a BCI-assisted motor imagery training performed by a stroke patient, with the aim of defining an index for the evaluation of the training achievements across session. The proposed  $h$  parameter would be independent from the selected BCI training setting and would better describe the physiological properties of the patterns generated during training, allowing a more appropriate evaluation of the training achievements than the behavioral performance (i.e. percentage of hit target).

## 1 Introduction

Nowadays, Brain Computer Interface (BCI) represents a promising technology to support motor and cognitive rehabilitation after stroke. In such rehabilitative context, BCI application aims at increasing the neuroelectric or metabolic brain responsiveness, which in turn would lead to a better recovery of function.

The Electroencephalographic (EEG) -based BCI operated by motor imagery (MI) can provide a valuable approach to support mental motor practice to enhance arm motor recovery after stroke [Mattia et al., 2012]. However, as stroke cortical lesions may result in a functional reduction/derangement of neuroelectrical activity generated over the ipsilesional hemisphere there is a need for further implementation of the procedures for recognition of those EEG patterns which are reinforced during the BCI-supported training of MI. Furthermore the online classification of trials as successful and failed also relies on a arbitrary choice of parameters and gains that do not strictly reflect the intrinsic properties (i.e. the level of desynchronization of SMR) of the patterns of activity recorded during the training.

The aim of this study was to define an index which would be independent from the settings adopted for the online control and thus, would describe the properties of neuroelectrical activations across BCI training sessions more appropriately than the hit rate (behavioral performance). To this purpose, we performed an offline analysis of EEG data sets acquired from stroke patients who underwent a MI-assisted BCI training aiming at promoting functional motor recovery of the paralyzed upper limb [Pichiorri et al., 2011]. The estimated index was monitored across training sessions and used to sort trials according to their intrinsic properties.

## 2 Methods

### 2.1 Experimental design

EEG data were collected from 14 stroke patients (age:  $64 \pm 8$  years; first ever, unilateral stroke causing hemiparesis/plegia) who underwent a BCI-assisted MI training. The training was preceded by an EEG screening session. EEG signals were recorded from 61 scalp positions (sampling rate 200 Hz) and such data were used to extract the features for the online control. Such control features were spatially selected only over the damaged (stroke) hemisphere (two channels) within an EEG frequency range relevant for sensorimotor function (10-15 Hz). The training protocol included 4 weeks of MI-based BCI training (3 sessions per week), during which the patient was asked to control the movements of a virtual representation of his own stroke-affected hand throughout the imagination of simple hand movements (visual neurofeedback). Each training session included 4 up to 8 runs (20 trials per run). Trials consisted of a baseline period (4 sec) followed by MI (max 10 sec). EEG signals (sampling rate at 200 Hz) during training were collected from 31 electrode positions (fronto-central, central, centro-parietal and parietal lines).

### 2.2 Offline analysis

After frequency (1-60 Hz band-pass and 50 Hz notch filters) and spatial (Common Average Reference; CAR) signal filtering, the power spectral densities (PSD) over all channels were computed by means of the Welch method [Welch, 1967]. We defined the new parameter  $h$  describing the activity associated to the selected features, elicited during each trial. The parameter is defined as follows:

$$h = \alpha * t(ch_1, bin) + \beta * t(ch_2, bin) \quad (1)$$

where channels  $ch_1$  and  $ch_2$  and  $bin$  (2Hz-frequency range) are the features selected during the initial screening for each subject,  $\alpha$  and  $\beta$  are multiplicative constants ( $\alpha + \beta = 1$ ), while  $t$  is the result of the Student's  $t$  test performed, for each trial, between the values of the PSD associated to the samples of task phase and those of baseline phase. The constants  $\alpha$  and  $\beta$  were set at the same value ( $\alpha = \beta = 0.5$ ).

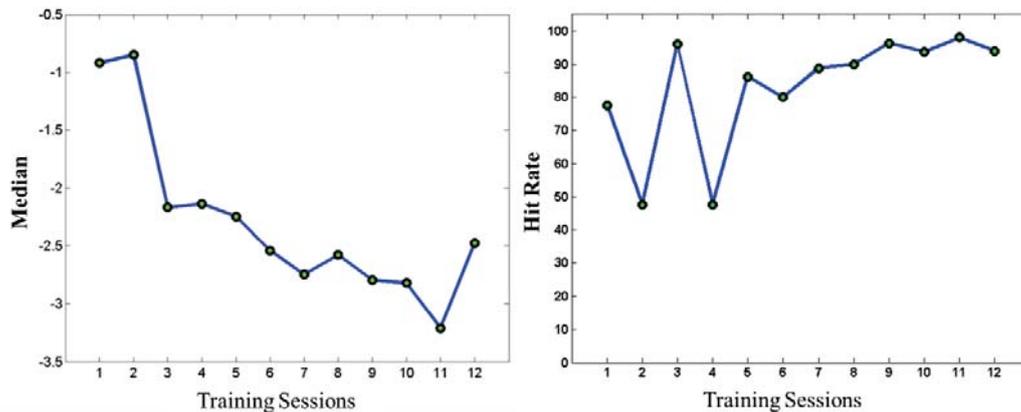
The parameter  $h$  was computed for each trial over the total of 12 training sessions and the distribution of  $h$  was built for each session (see Figure 2). The median of such distributions was considered as a synthetic descriptor of the features and was used to *monitor* the EEG pattern evolution session by session. We also evaluated 3 percentiles for each single distribution: 25% (first quartile), 50% (median) and 75% (last quartile) to investigate how trials would be distributed according to the  $h$  parameter. For each of the considered percentile, the overall trials were then separated into two groups: one group included all trials associated with  $h$  values below the  $h$  threshold relative to the percentile under investigation and the second group included all remaining trials.

To identify the spatial distribution of the spectral activity related to the selected trials, we computed the statistical scalp maps by contrasting the MI and baseline PSD values (student's  $t$ -test; significance level 5%) relative to each trial for the trained 2Hz-frequency range. False Discovery Rate correction for multiple comparisons was applied to avoid the occurrence of type I errors [Benjamini et al., 2001].

## 3 Results

The results are showed for one exemplary training data set acquired from a stroke patient with left affected hemisphere whose selected features were C3 and Cp3 electrode position at frequency bin of 10-12 Hz.

The graphs in Figure 1 show the trend of the  $h$  median values relative to each trial (left panel) and of the correspondent hit rate percentage (behavioral performance) (right panel) as a function of training sessions. In particular, Figure 1 shows that the median follows the positive outcome of the training with a decreasing trend along the investigated sessions. The hit rate, instead, after large oscillations in the first half of training sessions, saturated around 90% for the second half, not showing a clear trend along the intervention.



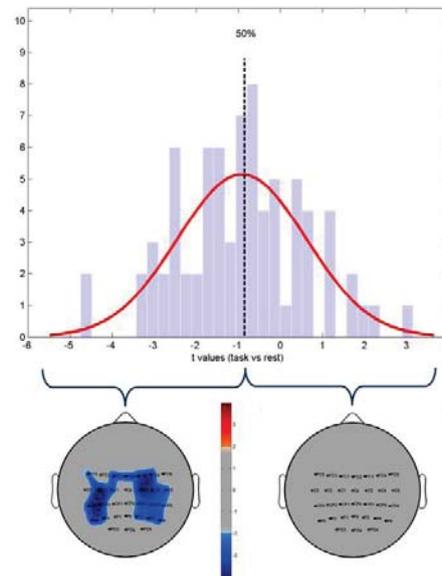
**Figure 1:** Median value of  $h$  parameter (left panel) and hit rate percentage (right panel) across training sessions.

The Figure 2 illustrates the distribution of the  $h$  parameter (top panel) and the relative statistical scalp maps (bottom panel) obtained from the second training session data set, considering the two groups of trials divided by the median. Note that only the statistical scalp map relative to the trials distributed to the left of the median (i.e. stronger desynchronization) showed a significant pattern (blue color in the left side map) as compared to that relative to the trials distributed to the right of median.

## 4 Discussion and Conclusion

In this study we defined a new parameter ( $h$ ) to measure the MI-induced desynchronization during BCI training for motor rehabilitation after stroke. The  $h$  distribution associated to each session allowed to evaluate how the trials are distributed with respect to the parameter. The median of the distributions proved to be more stable than the hit rate and suggested an increase of spectral desynchronization associated to MI across training. The statistical scalp maps obtained from trials to the left of the median provided a topological description of the activation underlying the execution of “good” trials (desynchronization). It was possible to achieve such result from the very early sessions thanks to the use of the  $h$  index.

Such preliminary results suggest that the proposed approach could be useful in optimizing a BCI-based intervention for neurorehabilitation purposes. In fact, future developments of the new described methodology will be oriented to the definition of an automatic procedure to detect the correct and specific threshold between successful and failed trials for the online analysis during MI-based BCI training.



**Figure 2:** Distribution of the  $h$  parameter (top panel) and statistical scalp maps (bottom panel) obtained from the second training session data set. Top panel: dashed line indicates the median (percentile 50%). Bottom panel: the color bar codes for t-value associated with each pixel: gray color = not significant differences between baseline and task trial PSD values; hot (yellow-red) = significant synchronization; cold (blue) = significant desynchronization.

**Acknowledgements.** Research partially supported by the Italian Ministry of Health (grant:RF-2010-2319611).

## References

- D. Mattia, F. Pichiorri, M. Molinari, R. Rupp, "Brain computer interface for hand motor function restoration and rehabilitation", Towards Practical Brain-Computer Interfaces for Hand Motor Function Restoration and Rehabilitation, 131-153, 2012, Springer.
- F. Pichiorri; F. Cincotti; F. de Vico Fallani; I. Pisotta; G. Morone; M. Molinari; D. Mattia; "Towards a brain computer Interface-based rehabilitation: from bench to bedside", Proceedings of the 5th international brain-computer interface conference, Verlag der Technischen, 268-271, 2011.
- Welch, (1967) "The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Averaging Over Short, Modified Periodograms", IEEE Transactions on Audio Electroacoustics, AU-15, 70-73.
- Y. Benjamini and D. Yekutieli, "The Control of the False Discovery Rate in Multiple Testing under Dependency," The Annals of Statistics, vol. 29, no. 4, pp. 1165-1188, Aug. 2001.

# Correlation of EEG Band Power and Hand Motion Trajectory

Attila Korik<sup>1</sup>, Nazmul Siddique<sup>1</sup>, Ronen Sosnik<sup>2</sup> and Damien Coyle<sup>1</sup>

<sup>1</sup>Intelligent Systems Research Centre, University of Ulster, Derry, U.K.

<sup>2</sup>Hybrid BCI Lab, Holon Institute of Technology, Holon, Israel

korik-a@email.ulster.ac.uk, nh.siddique@ulster.ac.uk,  
ronens@hit.ac.il, dh.coyle@ulster.ac.uk

## Abstract

A preliminary analysis of the correlation between high resolution EEG and three dimensional (3D) hand motion trajectories is presented. The study involved assessing which EEG frequency components and cortical areas show the most significant correlation with hand motion trajectory towards five targets positioned in different locations in space. The time shift between time-power pattern of the selected EEG frequency and related motion trajectory is also analyzed. The results show strong correlation between EEG and kinetic data in low frequency (0.5-4Hz) range as well as significant correlations in the 28-36Hz range. The results indicate that these EEG frequency bands may best be used to develop an EEG-based brain-computer interface (BCI) for the decoding of 3D hand motion trajectories.

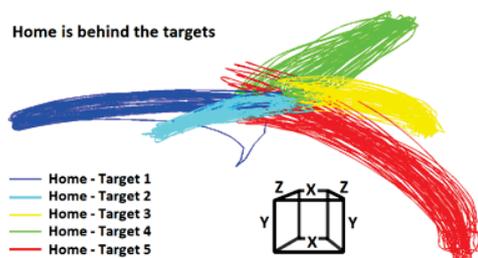
## 1 Introduction

Motion trajectory prediction based BCIs aim to exploit the relationship or correlates between EEG signals and limb motion to decode an imaginary limb movement in 3D space. To date only a limited number of studies have investigated EEG and kinetic data associated with 3D limb movements [1], [2]. Studies have shown that band-pass filtered low frequency EEG components around 2Hz convey a lot of information to the decoding task. Although Antelis et al [3] have drawn attention to possible misinterpretation of the results when using low frequency based motion prediction models, Paek et al. [4] have recently demonstrated the feasibility of decoding finger kinematics from low frequency scalp EEG signals. To date a comprehensive assessment of the spatial and spectral EEG correlates of real and imagined 3D hand motion trajectory has not been conducted. This paper aims to address this with an analysis of subjects undergoing a high resolution EEG recording whilst performing real and imagined 3D hand movements towards five targets positioned in different location in space. Sub-bands in the 0.5-40Hz EEG spectrum which show the highest level of correlation with movement trajectory are identified. The correlation strength of different cortical areas and different phase shifts between EEG and kinetic records are also assessed to learn more about the relationship between EEG signals and kinetic data.

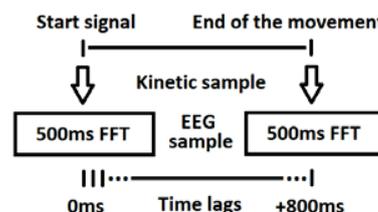
## 2 Experimental Task and Data Collection

The experimental task involved moving the right dominant hand between a home position ("H") and one of five target positions and return to home position. Target 1, 2 and 3 lie in the shoulder plane forming 45°, 67.5° and 90°, respectively, between the torso and the shoulder. Target 4 and 5 lie 45°

below and above the shoulder plane, forming 90° between the torso and the shoulder. Real movement blocks to a particular target were followed by imagined movement blocks to the same target. Participants could choose a home position that varies between subjects and blocks. Figure 1 illustrates motion trajectories of Subject 1. The task cue was synchronized with an auditory signal. Movements were followed by a rest phase. Both fast movements and slow movement blocks were interleaved where the length of motion and rest phases was 800ms and 500ms for slow and fast movement, e.g., Block 1 – movement H → 1 → 1 → H →H (0.8s each). The duration of the blocks was 48s, with an approximate inter-block interval (IBI) of 30s. The number of registered blocks was 20 from which only ten were considered in this study (slow and fast real movements between home and target positions). The analysis of the imagined movement blocks will be the subject of future study. Datasets containing parallel registered Electroencephalogram (EEG), Electromyogram (EMG) and kinetics data were acquired from six healthy right handed male human subjects (age range 25-42 years). EEG signals were registered in 62 channels + 1 electro-oculogram (EOG) at 1200 Hz. EMG was recorded from the Biceps with a sample rate of 2000Hz. Kinetic data were recorded from the right dominant hand, elbow and shoulder at 30 frame per seconds (FPS) using a 3D Microsoft Kinect camera system. All datasets were acquired at the Hybrid BCI lab at Holon Institute of Technology (HIT), Israel.



**Figure 1:** Hand motion trajectories of Subject 1. This figure is prepared by smooth filtered, valid kinetic data.



**Figure 2:** Illustration of synchronization between FFT windows and kinetic data in case of slow tasks.

### 3 Preprocessing

EEG and kinetic data were stored in a pointer based structure that enabled variable length of epochs without re-slicing and reduced memory requirement. Baseline shift was removed and the EEG was filtered by 0.5-40Hz, eight-order, band-pass Butterworth filter. The Fast Fourier Transformation (FFT) was applied for calculating power value of EEG frequency components in 2Hz wide bands (non-overlapped) between 0 and 40 Hz. We used 500ms width FFT windows and 33.3ms time lags between two windows. This lag has been chosen for ease of synchronization with kinetic data sampled at 30FPS. Figure 2 illustrates the setup of FFT window for slow tasks i.e., 800ms. Smoothing filter has been applied on kinetics data for noise reduction as the band-pass filtering causes a distortion in data at the beginning of motion. The filter calculated a mean value of five adjacent samples.

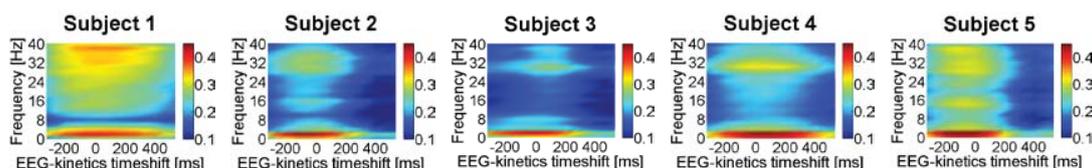
Time synchronization is crucial in this comparative analysis. We used the sample rate converted EEG triggers for synchronizing the kinetic data with the EEG data where the first trigger was recorded simultaneously for both data types. Task compliance validation was performed manually. Tasks were considered valid only when converted EEG triggers matched the beginning of the motion. Subject 6 was discarded due to inadequate kinetic records.

### 4 Correlation Analysis

The correlation coefficient between a power pattern (belonging to one of the computed FFT frequencies) and related kinetic data was calculated for each valid trial.

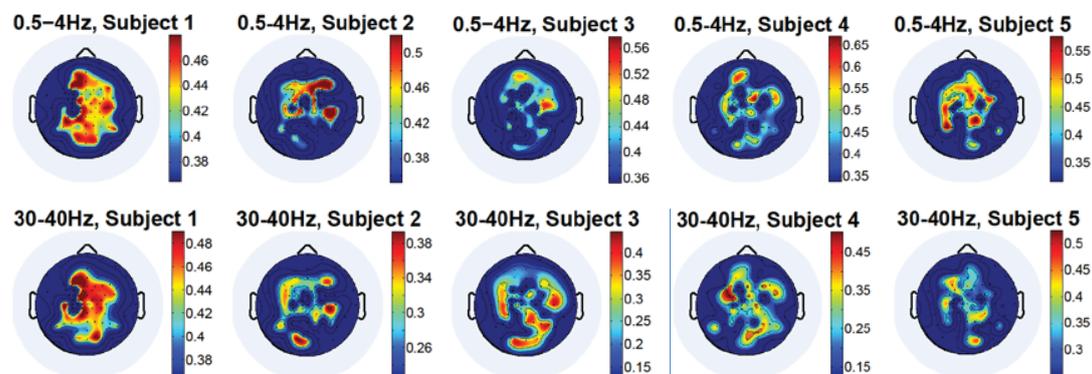
The kinetic data consist of an x, y or z Cartesian vector component of hand coordinates in the 3D space or joint angle at the elbow, hand and shoulder reference points. We analyzed the correlation between the two descriptors during the movement hence the size of the correlation window matched the time interval of the related task (slow tasks 800ms, fast tasks 500ms). Pearson's linear correlation was chosen as we were interested in the correlation strength between two row vectors. The first row vector contained EEG band power values which were gained from FFT at analyzed time lags, the second one contained the kinetic data (see Figure 2). The *corrcoef()* Matlab function was used for computation of correlation coefficients ( $R$ ) and  $p$  values from Student's t-test. The most significant  $R$  values were used for further analysis ( $p < 0.01$ ). As we were only interested in the correlation strength between the two descriptors, the correlation sign was overlooked.

We also aimed at detecting the time shift between EEG and kinetic record wherein correlation is maximal. Therefore, we repeated the above calculations for different time shifts between EEG and kinetic patterns. Figure 3 summarizes the distribution of correlation level between the EEG signal and related Hand(z) trajectory, where the vertical axis represents the different EEG frequencies and the horizontal axis represents the time shifts between EEG signal and kinetic record. (e.g., the -200ms column contains correlation intensity for an EEG compared with kinetic data registered 200ms earlier which is an already executed motion thus the +200ms column compares EEG and planned motion). Colour code indicates strength of the correlation where red colour denotes higher correlation levels.

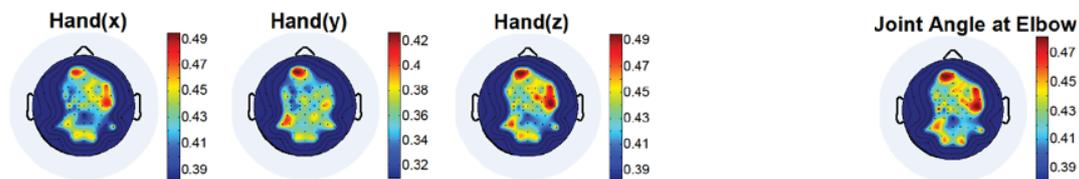


**Figure 3:** Correlation level distribution between EEG patterns and Hand(z) trajectories for different EEG frequencies and time shifts between EEG signal and kinetic data. Colour code is an indicator of correlation level.

We have found the highest level of correlation around 2Hz in the low EEG frequency range. Additionally, the 28-36Hz band also shows high correlation level although it is not as pronounced as that in the 2Hz range. The distribution of correlation levels was similar for different kinetic components (hand x, y, z and joint angle at the elbow). Analysis of different velocities (slow and fast) and movement directions shows some diversity in the correlation level although the variance was not significant. The correlation was maximal when EEG signals were compared with parallel registered kinetic data whose results are illustrated at 0ms column in Figure 3.



**Figure 4A:** Correlation between power pattern of EEG bands and Hand(z) trajectory of subject 1-5 (slow tasks)



**Figure 4B:** Comparison of the correlation levels of EEG patterns and kinetic components (0.5-4Hz, slow tasks)

Figure 4A and Figure 4B illustrates distribution of correlation levels in case of parallel recorded EEG signals and kinetic data (time shift = 0) for different cortical areas. Dark blue (background) colour indicates areas where we were unable to collect good quality synchronized EEG and kinetic data in our pilot study for valid statistics. Figure 4A compares results from different subjects and Figure 4B shows the difference when x, y or z vector component of hand trajectory or joint angle at the elbow has been used as kinetic data in correlation. The similarity between Hand (z) and joint angle at the elbow is the highest, compared to x or y. The result indicate the EEG associated with the z direction vector component of the hand movement is more correlated with the joint angle than x or y. The increased frontal and occipital correlation in case of 30-40Hz gamma band [5] highlight importance of these cortical areas in motion trajectory prediction. Our results could not unequivocally support this theory because there are insufficient data gained from some important cortical areas.

## 5 Conclusion

We have used a correlation analysis to examine the importance of different EEG frequencies, cortical areas and time shift between EEG signals and kinetic data in motion trajectory prediction. We identified strong correlation between EEG and kinetic data in the delta (0.5-4Hz) EEG range which is consistent with repetitive finger movement study from Paek et al. [4]. Furthermore, the correlation in 28-36Hz band was also more pronounced at pre-movement and at movement onset than in the mu, alpha and beta bands. Gamma band activity or 40 Hz EEG is strongly related to planning of a specific movement [6]. We found that the correlation between EEG and kinetic data is maximal when both are registered in parallel i.e., no lag between EEG and Kinetic data.

## References

- [1] T. J. Bradberry, R. J. Gentili, and J. L. Contreras-Vidal, "Reconstructing three-dimensional hand movements from noninvasive electroencephalographic signals," *J. Neurosci.*, vol. 30, no. 9, pp. 3432–7, Mar. 2010.
- [2] J.-H. Kim, R. Chavarriaga, J. D. R. Millan, and S.-W. Lee, "Three-dimensional upper limb movement decoding from EEG signals," *2013 Int. Winter Work. Brain-Computer Interface*, pp. 109–111, Feb. 2013.
- [3] J. M. Antelis, L. Montesano, A. Ramos-Murguialday, N. Birbaumer, and J. Mínguez, "On the usage of linear regression models to reconstruct limb kinematics from low frequency EEG signals," *PLoS One*, vol. 8, no. 4, p. e61976, Jan. 2013.
- [4] A. Y. Paek, H. Agashe, and J. L. Contreras-vidal, "Decoding repetitive finger movements with brain activity acquired via non-invasive electroencephalography," *Front. Neuroeng.*, vol. 7, no. 3, p. in press, 2014.
- [5] M. Grosse-Wentrup, B. Schölkopf, and J. Hill, "Causal influence of gamma oscillations on the sensorimotor rhythm.," *Neuroimage*, vol. 56, no. 2, pp. 837–42, May 2011.
- [6] G. Pfurtscheller, D. Flotzinger, and C. Neuper, "Differentiation between finger, toe and tongue movement in man based on 40 Hz EEG," *Electroencephalogr. Clin. Neurophysiol.*, vol. 90, no. 6, pp. 456–460, Jun. 1994.

# Applying Brain-Computer Interfaces outside the lab – Piloting a plane with active BCI

Thorsten O. Zander<sup>1,3</sup>, Tim Fricke<sup>2</sup>, Florian Holzapfel<sup>2</sup> and Klaus Gramann<sup>3</sup>

<sup>1</sup> Team PhyPA, Technische Universitaet Berlin, Berlin, Germany

[tzander@gmail.com](mailto:tzander@gmail.com)

<sup>2</sup> Lehrstuhl fuer Flugsystemdynamik, Technische Universitaet Muenchen, Garching bei Muenchen, Germany

[tim.fricke@tum.de](mailto:tim.fricke@tum.de), [florian.holzapfel@tum.de](mailto:florian.holzapfel@tum.de)

<sup>3</sup> Biological Psychology and Neuroergonomics, Technische Universitaet Berlin, Berlin, Germany

[klaus.gramann@tu-berlin.de](mailto:klaus.gramann@tu-berlin.de)

## Abstract

The precursor study presented here describes one further step towards investigating active BCIs in realistic scenarios. We invited six trained pilots to control horizontal flight in a flight simulator via an active BCI. Performance was tracked via standard BCI measures and performance in operational flight tasks. Results indicate that standard BCI setups can indeed be used in realistic scenarios outside of the laboratory.

## 1 Introduction

Previous work has investigated active BCIs [6] based on motor imagery in lab studies (see [1] for example). As BCIs are intended to be applied in real world scenarios the question arises whether the methodology developed in recent years is also reliably working in more complex and less controlled environments. From a human factors perspective it is of interest whether users can use a BCI while interacting with a technical system in a realistic setting [5]. In the precursor experiment we are presenting here we addressed these questions by investigating motor imagery-based BCI control in a flight simulator. Experienced pilots controlled a plane in the horizontal axis while performing operational flight tasks.

## 2 Setup

Six pilots with different levels of experience were invited to participate in our experiment. In each run the participants occupied the left seat in a flight simulator. The simulated outside world view was projected on a cylindrical 180° screen round the fixed-base cockpit. The instruments provided to the subjects comprised classical (backup) instruments (airspeed indicator, attitude indicator, altimeter and magnetic compass) as well as a research display, which showed attitude, airspeed, altitude, vertical speed, flight path angle, heading, turn rate, a brain signal feedback with a delay of 1 second and task specific elements. Two types of flight control were investigated. Control type A enabled the pilot to change the rate of turn directly (direct mapping). Control type B provided a turn rate only if the BCI input exceeded a threshold, and automatically returned to straight flight otherwise.

## 3 Experiment

### 3.1 Calibration

Before online flight control each participant generated calibration data. Based on data from these trials a BCI classifier was trained for each participant individually. Each trial had a sequence of fixation, command and relax states. This sequence was communicated by objects displayed in the center of the screen. Fixation was indicated by a cross displayed for 1 second. It was then replaced by a letter randomly drawn from the set  $\{L,R,F\}$ , which was displayed for 4 seconds. Participants were advised to consistently imagine a specific movement with their left hand (L), their right hand (R) or their foot (F), respectively. Each calibration session had 120 trials (40 trials per class) in total, subdivided by breaks of 60 seconds after 40 trials.

### 3.2 Application

The resulting BCI was then applied to lateral control in simulator scenarios. Altitude and throttle were controlled automatically by the simulator. As all pilots were familiar with the controls and displays of the simulator, tasks were communicated through the displays. Either a marker on the Horizontal Situation Indicator (HSI), indicating a specific angle to take (heading bug), or the alignment of the main axis of the plane (lubber line) with a course select pointer on the HSI (localizer) was used for this purpose. Participants were advised about the tasks by the experimenter before the application experiment began. The application had three stages. Participants flew all stages with both controls (as described in section 2). The order of control types was permuted over participants.

**Turns.** The task was to follow steps of the heading bug on the HSI, i.e. to acquire and hold a given heading. The pilot was advised to always choose the shortest turn to acquire the next heading and, if possible, to turn with standard rate of turn ( $3^\circ/s$ ). The sequence of steps has been selected to be random appearing with an equal number of left and right turns.

**Tracking.** In the tracking task, pilots were instructed to follow the heading bug on the HSI. This time, however, the heading bug oscillates about the initial heading. The forcing function of this tracking task was composed of 10 different sine waves, so that it was randomly appearing.

**Approach with offset localizer and visual landing.** The task was to first intercept the offset localizer, and then track it [3]. As there were no outside visual references in this stage, it was not apparent to the pilot that the localizer was offset. When the aircraft descended below 500ft above ground level, the runway became visible. The pilot now ignored his navigation instruments and continued the approach only by outside visual references. Since the first part of the approach was offset, the pilot was forced to conduct a sidestep maneuver. After that, he tracked the runway centerline. The simulation ended just before touchdown.

## 4 Analyses

BCILAB [2] was used to calibrate classifiers for all pairwise comparisons of classes. Features were extracted with standard parameters for Common Spatial Patterns [4] on time windows between 1 and 3 seconds after task onset. The classifier with best cross-validation estimates was then used in online application. Flight performance was tracked by diagrams indicating the deviation of the optimal target heading.

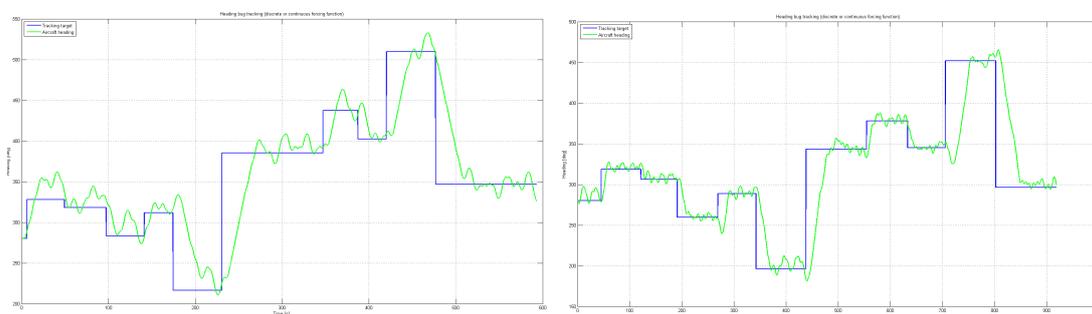


Figure 1: Exemplary flight diagram for control type A (left) and control type B (right) during turns. The blue line indicates the optimal target heading, the green line the actual flow target heading. X-Axis is in seconds, Y axis is heading in degree.

Participant No. 7

CSP, Standard Parameters, 1-3 sec.

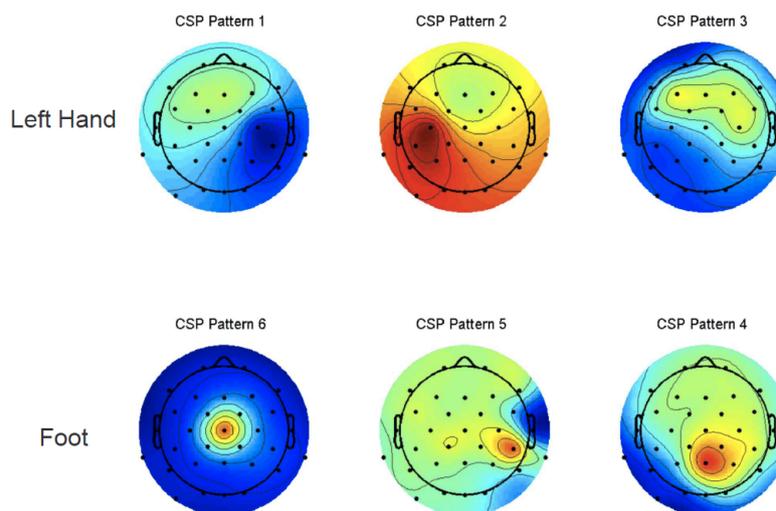


Figure 2: Common Spatial Patterns selected for Participant 6. Discriminating between classes 'left hand' (upper row) and 'foot' (lower row). Pattern 1 focuses on the right motor cortex, while pattern 2 aims at the left motor cortex inversely weighted and pattern clearly weights the central motor cortex.

## 5 Results

Table 1 shows cross-validation estimates of the chosen classifiers (left column), while figure 2 shows the chosen CSPs for the best performing participant (no. 6). Online flight performance is shown exemplarily for participant no. 6 in the 'turns' condition for control type A and control type B (figure 1). Participants were asked whether they felt having control. Participants with high classification accuracies stated to have control while participants with low accuracies did not (see table 1, right column).

	CV acc.	Participant reported having control?
Participant No. 1	98	yes
Participant No. 2	58	no
Participant No. 3	51	no
Participant No. 4	64	no
Participant No. 5	89	yes
Participant No. 6	95	yes

Table 1: Best classification accuracy from pairwise comparisons of the classes via cross validation.

## 6 Discussion and Outlook

The results of this precursor study indicate that a one-dimensional control via active BCIs is applicable in real world applications. Interestingly the group of participants was divided in two subgroups. One group had almost perfect control while the other had no control at all. In future studies we will investigate whether this is a stable effect and whether the participants show a similar result in standardized lab-studies. As the aircraft's flight dynamics have an impact on control as well, we will also investigate further improvements of the flight controller.

## References

- [1] Benjamin Blankertz, Ryota Tomioka, Steven Lemm, Motoaki Kawanabe, and K-R Muller. Optimizing spatial filters for robust eeg single-trial analysis. *Signal Processing Magazine, IEEE*, 25(1):41–56, 2008.
- [2] Christian A. Kothe and Scott Makeig. Beilab: a platform for brain–computer interface development. *Journal of neural engineering*, 10(5):056014, 2013.
- [3] Herman Albert Mooij. Criteria for low-speed longitudinal handling qualities of transport aircraft with closed-loop flight control systems. 1984.
- [4] Herbert Ramoser, Johannes Muller-Gerking, and Gert Pfurtscheller. Optimal spatial filtering of single trial eeg during imagined hand movement. *Rehabilitation Engineering, IEEE Transactions on*, 8(4):441–446, 2000.
- [5] Thorsten O. Zander. *Utilizing Brain-Computer Interfaces for Human-Machine Systems*. PhD thesis, Universitätsbibliothek der Technischen Universität Berlin, 2011.
- [6] Thorsten O Zander and Christian Kothe. Towards passive brain–computer interfaces: applying brain–computer interface technology to human–machine systems in general. *Journal of Neural Engineering*, 8(2):025005, 2011.

# Towards a SSVEP-BCI Based on Depth of Field

Anibal Cotrina<sup>1</sup>, Teodiano Bastos<sup>1</sup>, Andre Ferreira<sup>1</sup>, Alessandro Benevides<sup>1</sup>,  
Javier Castillo-Garcia<sup>1,2</sup>, David Rojas<sup>3</sup> and Alessander Benevides<sup>4</sup>

<sup>1</sup> Post-Graduate Program of Electrical Engineering, Federal University of Espirito Santo, Brazil  
anibal.atencio@ufes.br

<sup>2</sup> Post-Graduate Program of Electronics Engineering, Del Valle University, Cali, Colombia

<sup>3</sup> Computer Vision Center, Autonomous University of Barcelona, Barcelona, Spain

<sup>4</sup> ICT Doctoral School, University of Trento, Trento, Italy

## Abstract

It has been shown that the visual evoked potential amplitude reduces as the stimulus becomes increasingly defocused. Based on depth-of-field theory, which states that subject distance is the range of distance in which an object appears sharp at the retinal image, the present work attempts to verify an alternative SSVEP BCI setup, where the user gaze simultaneously two SSVEP stimuli flickering with different frequencies and located at different distances. This setup relies on the assumption that the focused stimulus is able to elicit distinguishable SSVEP pattern regardless the non-focused stimulus that is also present. Three subjects and two stimuli were considered. Clear SSVEP pattern was elicited when they were asked to focus either the nearest or the farthest stimulus.

## 1 Introduction

Visual evoked response (VEP) is an event related potential that occurs involuntarily in response to a visual stimulus. It has been shown that the defocusing of the retinal image has a greater effect on the latency of this potential [8]. An object is defocused when it is located out of the eye's Depth of Field (DOF), that is defined as the range of distance in which an object appears sharp at the retinal image [6]. The focusing of a target is performed by an accommodation mechanism that is achieved when a neural signal is sent to the ciliary muscle changing the shape of the crystalline lens. It modifies the angle of refraction minimizing automatically the blurriness in the retinal image [3]. Then, the performance of brain-computer interfaces (BCI) based on steady-state visual evoked potentials (SSVEP) could be affected by this optical phenomenon, specifically when the flickering stimuli go out of focus of the subject's eye. As the amplitude of the VEP pattern can be reduced as the retinal image is increasingly defocused, an arising hypothesis is: if a BCI user is gazing simultaneously two stimuli flickering with different frequencies and located at different distances (enough to get only one into the subject's DOF), the focused stimulus will be able to elicit distinguishable SSVEP pattern regardless the non-focused stimulus is also present. Due to the user choose the target stimulus by shifting the focus instead of gaze movements or attention, this work becomes an alternative method for presenting SSVEP stimuli either to traditional SSVEP paradigm or covert attention based SSVEP paradigms. Results showed that clear SSVEP pattern can be elicited when the subjects are asked to focus either the nearest or the farthest stimulus.

## 2 Methods

Three healthy subject without any experience with BCI were considered in this work. The experiments were undertaken with the understanding and written consent of them. This study

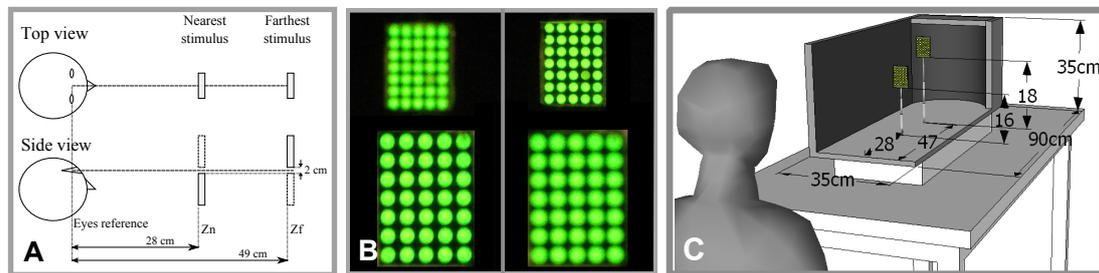


Figure 1: (a) Schematic diagram of the experimental setup; (b) photograph of the two LED arrangements when nearest and farthest stimuli were focused on; and (c) experimental setup.

was approved by the research ethics committee of the Federal University of Espirito Santo (Brazil). Two stimuli, emitted by two  $5 \times 7$  LED arrangements flickering at 5.6 Hz and 6.4 Hz located at different distances were displayed simultaneously. As shown in Figure 1(a), the nearest (6.4 Hz) and farthest (5.6 Hz) arrangements were placed at 28 and 47 cm far from the subject, respectively. In this positions non-focused objects appear considerably blurred [9]. Figure 1(b) shows the simulated appearance of the two stimuli in the subject's field of view when either the nearest or the farthest is focused. Two 10-trial experiments were performed in which the subjects were asked to focus on the nearest stimulus in the first one and focus on the farthest stimulus in the second one. During each 22-seconds trial, the subjects seated in front of a experimental box (see Figure 1(c)) and were asked to focus on the target stimulus for 17 s, then was asked to close the eyes for 5 s. The electroencephalographic (EEG) signals were recorded between 5 s and 17 s. EEG signals from 19 electrodes positioned according to the international 10-20 system were acquired at a sampling rate of 200 Hz using the BNT36 device with a cap of integrated wet electrodes. The grounding electrode was positioned on the user forehead and bi-auricular reference was adopted. Signals from electrodes O1 and O2 filtered using an (4 - 50 Hz) elliptic band-pass were used to verify the SSVEP patterns. Other channels were used to perform common average reference spatial filtering. The traditional Power Spectral Density Analysis (PSDA) method and the Canonical Correlation Analysis (CCA) method [2] were employed to compute the classification accuracy. Due to the gaze focus may occur without intention, to evaluate the robustness of the assessment and the probability of false positive errors, accuracy, Kappa coefficient, sensitivity, specificity and information transfer rate (ITR) were computed.

### 3 Results

Figure 2 shows the normalized power spectrum of the SSVEP pattern at electrode O2 of the subject 1 computed using the Fourier transformation. To emphasize the elicited SSVEP peaks, offset component was removed by subtracting the mean value of each frequency. Figures 2(a) and 2(b) show the spectral responses of all trials (gray curves) together with the average (black curve) computed over the 12 s of the EEG signals of the entire trial when the subject was asked to focus on the 6.4 Hz and 5.6 Hz stimulus, respectively. In both cases SSVEP peaks were elicited at stimulus frequency, at the second (12.8 Hz and 11.2 Hz) and at third harmonic frequencies (19.2 Hz and 16.8 Hz). The length of the analysis window is an important aspect to be considered when assuming that the background noise is a random additive signal and the potential is deterministic. Figures 2(c) and 2(d) show the normalized amplitude spectra

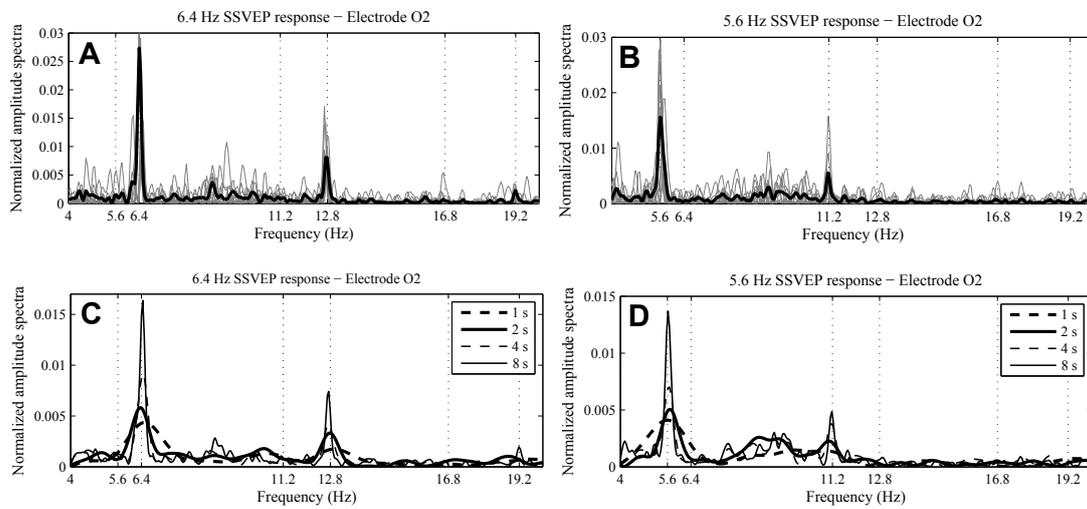


Figure 2: (a) and (b) Normalized amplitude spectra corresponding to 10 trials (gray curves) together with the average (black curves). (c) and (d) Normalized amplitude spectra corresponding to different data lengths for one subject.

computed using different time intervals (1, 2, 4 and 8 s). The signal/noise ratio can be improved by increasing data to estimate the spectra, since the energy of the deterministic signal increases quadratically with the increasing of the signal window length. Table 1 shows the parameters that were considered to evaluate the robustness of the assessment for three subjects. Values in parentheses correspond to the PSDA method applied to peaks detection. Values without parentheses correspond to the CCA method. In all cases, the performance of the CCA method was better than PSDA method.

Table 1: Evaluation parameters when the PSDA method (in parentheses) and the CCA method (without parentheses) were employed to detect SSVEP peaks.

Subject	Sensitivity	Specificity	Accuracy	Kappa	ITR
Subject 1	0,78 (0,70)	1,00 (0,68)	0,85 (0,69)	0,69 (0,37)	3,37 (1,49)
Subject 2	0,90 (0,58)	1,00 (0,87)	0,95 (0,73)	0,89 (0,45)	6,96 (2,24)
Subject 3	0,78 (0,70)	1,00 (0,68)	0,85 (0,69)	0,69 (0,37)	3,37 (1,49)
Average	0,82 (0,66)	1,00 (0,75)	0,88 (0,70)	0,76 (0,40)	4,57 (1,74)

## 4 Conclusion

The results here presented conclude that if two flickering stimuli are appropriately located, then, the focused stimulus is able to elicit distinguishable SSVEP pattern, regardless the non-focused stimulus also present in the user's field of view. Based on the results showed in Figure 2 and Table 1, which are very promissory, a SSVEP-BCI based on DOF can be developed. Although, the proposed SSVEP-BCI based on DOF uses overt orienting, it is independent of

gaze movements, because the focusing of a target is performed by an accommodation mechanism that, like a pupil contraction, is an ocular reflex response. Hence, this work becomes an alternative method for presenting SSVEP stimuli to the traditional SSVEP-BCI paradigm. The existing traditional SSVEP-BCIs are becoming robust systems and achieving high transfer rates [4]. However, many designs require reliable control of eye movements because the subject performs the selection by gazing directly at each stimulus location. On the other hand, this work also becomes a complementary method to the attention based SSVEP paradigm. SSVEP-BCIs named independent-BCIs, like BCIs based on spatial-visual selective attention [5] or non-spatial selective attention [10] use selectively covert attention to make the selection of the stimulus instead of muscular gaze shifting [1]. Nevertheless, covert orienting demands high attention, and it is associated with difficult tasks that might not become automatic, even with high levels of practice, and requires intention [7]; while overt orienting demands low attention and is associated with easy and/or well-practiced tasks.

## ACKNOWLEDGMENT

The authors would like to thank CAPES/Brazil and FAPES/CNPq (Process 53666038/2011) for the financial support. Also, the authors would like to acknowledge the Ophthalmologist Diogo Sperandio for their advices.

## References

- [1] Brendan Z. Allison, Dennis J. McFarland, Gerwin Schalk, Shi Dong Zheng, Melody Moore Jackson, and Jonathan R. Wolpaw. Towards an independent brain-computer interface using steady state visual evoked potentials. *Clinical Neurophysiology*, 119(2):399 – 408, 2008.
- [2] Guangyu Bin, Xiaorong Gao, Zheng Yan, Bo Hong, and Shangkai Gao. An online multi-channel ssvpe-based braincomputer interface using a canonical correlation analysis method. *Journal of Neural Engineering*, 6(4):046002, 2009.
- [3] Sheldon M. Ebenholtz. *Oculomotor Systems and Perception*. Cambridge University Press, 1 edition, 2001.
- [4] Christoph Guger, Brendan Allison, Bernhard Grosswindhager, Robert Pruckl, Christoph Hintermuller, Christoph Kapeller, Markus Bruckner, Gunther Krausz, and Gunter Edlinger. How many people could use an SSVEP BCI? *Frontiers in neuroscience*, 6, 2012.
- [5] S.P. Kelly, E.C. Lalor, C. Finucane, G. McDarby, and R.B. Reilly. Visual spatial attention control in an independent brain-computer interface. *Biomedical Engineering, IEEE Transactions on*, 52(9):1588–1596, Sept 2005.
- [6] Alex Paul Pentland. A new sense for depth of field. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, PAMI-9(4):523–531, July 1987.
- [7] Michael I. Posner and Steven E. Petersen. The attention system of the human brain. *Annual Review of Neuroscience*, 13(1):25–42, 1990. PMID: 2183676.
- [8] Samuel Sokol and Anne Moskowitz. Effect of retinal blur on the peak latency of the pattern evoked potential. *Vision Research*, 21(8):1279 – 1286, 1981.
- [9] Dhanraj Vishwanath and Erik Blaser. Retinal blur and the perception of egocentric distance. *Journal of Vision*, 10(10), 2010.
- [10] Dan Zhang, Xiaorong Gao, Shangkai Gao, A.K. Engel, and A. Maye. An independent brain-computer interface based on covert shifts of non-spatial visual attention. In *EMBC 2009. Annual International Conference of the IEEE*, pages 539–542, Sept 2009.

# Detection of Eyes Closing Activities through Alpha Wave by Variability Analysis

Denis Delisle-Rodriguez<sup>1,2\*</sup>, Javier F. Castillo-Garcia<sup>1</sup>, Teodiano Bastos-Filho<sup>1</sup>, Anselmo Frizeira-Neto<sup>1</sup>, Alberto López-Delis<sup>2</sup>

<sup>1</sup>Post-Graduate Program of Electrical Engineering, Federal University of Espirito Santo, Brazil

<sup>2</sup>Center of Medical Biophysics, University of Oriente, Santiago de Cuba, Cuba

[delisle05@gmail.com](mailto:delisle05@gmail.com), [javier.castillo@correounivalle.edu.co](mailto:javier.castillo@correounivalle.edu.co),  
[teodiano@ele.ufes.br](mailto:teodiano@ele.ufes.br), [anselmo@ele.ufes.br](mailto:anselmo@ele.ufes.br), [alberto.lopez@cbiomed.cu](mailto:alberto.lopez@cbiomed.cu)

## Abstract

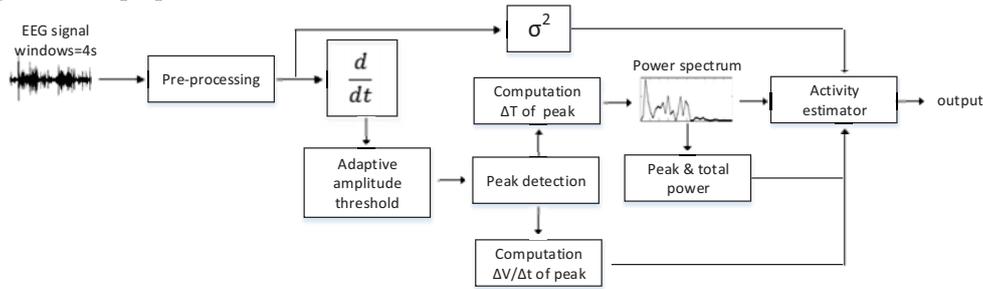
This work presents an alternative method to detect events correlated to eyes opening and closing, based on electroencephalography (EEG) measured from the occipital lobe. The goal is to propose a method based on variance and variability analysis of alpha wave to classify segments of EEG signals that contain activities originated by eyes closing. The peak values of alpha waves are determined to obtaining the variability series, which are used to compute the power spectrum through Welch periodogram. The combination of alpha wave variance and their variability, were used to obtain a coefficient that allows discriminating between events of eyes opening and closing. This approach can be used to control the switching of a brain computer interface (BCI).

## 1 Introduction

Several technologies have been developed to assist and improve the communication of people with paralysis and severe motor disability. BCI is a communication system that does not depend on the brain's normal output pathways of peripheral nerves and muscles. At UFES/Brazil we are developing a BCI system for an autonomous car based on evoked visual stimulus (Castillo *et al* 2013), which can cause visual fatigue. A good alternative is switching the BCI through a user command, that can be performed by eyes closing. This way, electroencephalography signals (EEG) have been employed, which contain information that allows the eyes closure detection. Eyes opening and eyes closing activities can be sensed on the occipital lobe through alpha wave analysis in a frequency range of 8 Hz to 13 Hz. A high energy of the alpha wave corresponds to closed eyes on awake subjects (in 90% of healthy and people with disabilities) (Alaraj and Fukami 2013). Alpha waves have been applied to operate electrical devices, however, the automatic recognition associated to eyes opened (EO) and closed (EC) is not a trivial task, because the bandwidth of alpha wave is affected by natural variation and electrical noise, and muscle artefacts. Several methods have been developed to automatic detection of alpha wave, such as: analogue filtering and smoothing (AFS), peak detection and counting, power spectrum analysis, fractal dimension, KM20-Langevin and approximated entropy (Kirkup *et al* 1998, Craig *et al* 2005, Sakai *et al* 2010, Alaraj and Fukami 2013). All aforementioned methods use a threshold value as a reference that depends on each subject and experiment conditions. The aim of this work is to propose one automatic method based on the variability information of alpha wave of EEG for recognition of eyes closing events in awake subjects, in order to activate a BCI.

## 2 Material and methods

The proposed method allows the online automatic detection of the opening and closing eyes activity from alpha wave by variability analysis in time and frequency domains. Figure 1 presents a block diagram of the proposed method.



**Figure 1:** Representation of the proposed method to detect eyes opening or eyes closing.

The electroencephalographic (EEG) signals are captured and analyzed in sliding windows of 4s, shifted for 1s. First, EEG signals are pre-processed by a Common Average Reference (CAR) filter and a bandpass elliptic filter (bandwidth: 8-13 Hz, 5<sup>th</sup> order), to reduce the common interference and select the principal information of the alpha wave, respectively. Second, the variance and first derivative of alpha wave are computed, and adaptive amplitude thresholds are calculated on the first derivative signal to detect local peak values. The thresholds are computed on windows of 78 ms with overlapping 90% using the equations (1) and (2). The local peak values higher than a threshold are taken as reference to detect the peak values of the alpha wave, which are employed to obtain the temporal series of the variability. After, these peak values are used to compute the transition velocity ( $\Delta V/\Delta t$ ) and power spectrum (by Welch periodogram, overlapping of 50%, and Blackman window) to determine the peak and total power of the variability. Finally, the variance and variability information of the alpha wave are employed to compute a coefficient to detect the events (eyes opening or eyes closing). This coefficient can be calculated by equation (3).

$$RMS_k = \frac{1}{N} \sum_{i=1}^N y_i^2 \quad (1)$$

$$AT_k = RMS_k - 0.4 \times \sigma_k \quad (2)$$

where  $RMS$  is the root mean square of the window of 4 s,  $y$  is the first derivative of alpha wave,  $N$  is the total samples,  $\sigma$  is the standard deviation of the first derivative,  $AT$  is the amplitude threshold,  $0 < k < W-1$ , and  $W$  is the total windows.

$$coefficient = \frac{P_{max}}{\left(v_t \times \frac{1}{f_{max}}\right)^2} + \frac{P_{total}}{\sigma^2} \quad (3)$$

where  $P_{max}$  and  $P_{total}$  are the peak and total value of power spectrum of the variability series,  $f_{max}$  is the frequency that correspond to  $P_{max}$ ,  $\sigma^2$  and  $V_i$  are the variance and transition velocity between peak values (from minimum to maximum value) of the alpha wave, respectively.

### 2.1 Experimental protocol

The experimental protocol allows the acquisition of EEG signals in two events: opened eyes and closed eyes. Eight healthy subjects participated in the experiment and provided written informed consent. Each subject is sit near to a visual (virtual animated eye) and auditory (tone) synchronized stimuli that guide the oscillation period necessary to open (remaining 10 seconds) and close (remaining 10 seconds) the eyes during 100 seconds. Five records of EEG signals were acquired to each subject. EEG signals

were acquired from Emotiv EPOC neuroheadset using fourteen electrodes on the scalp, placed according to the international 10/20 system: AF3, AF4, F3, F4, F7, F8, FC5, FC6, P7, P8, T7, T8, O1, O2. However, only the data from occipital alpha waves (8-13Hz) on O1 and O2 sites were used. The occipital locations O1 and O2 were chosen for two reasons: (1) alpha activity is larger in the occipital region as it is directly linked to the visual perception; and (2) there is less artefact in this region, such as ocular muscle activity, compared to the frontal scalp regions. EEG channels were sampled with 128 Hz at 1.95µV, the least significant bit voltage resolution.

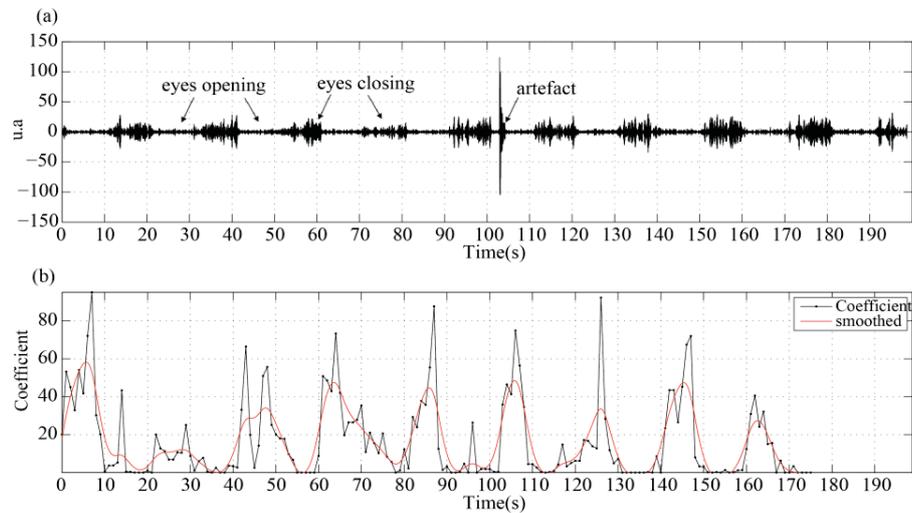
### 3 Results and discussion

Table 1 shows the behavior of the variables considered to estimate the possible event of eyes in the analyzed segment of the eight subjects.

Events	Transition velocity (u.a/s)	Variance	Total power (s <sup>2</sup> )	Peak power (PP) (s <sup>2</sup> )	Frequency of PP (Hz)
Eyes opening	161.07 ± 73.42	14.72 ± 17.95	20.66 ± 8.40	305.64 ± 117.77	1.59 ± 1.25
	114.81 ± 40.38	5.72 ± 6.41	16.36 ± 8.71	237.05 ± 110.6	1.59 ± 1.20
	257.62 ± 64.08	26.59 ± 13.31	16.07 ± 7.29	258.40 ± 116.4	1.64 ± 1.08
	109.94 ± 50.17	5.83 ± 6.96	17.28 ± 10.63	261.05 ± 121.84	1.29 ± 0.98
	127.37 ± 38.56	6.38 ± 7.85	17.45 ± 6.93	265.23 ± 108.82	1.81 ± 1.14
	172.23 ± 51.75	11.73 ± 7.98	17.59 ± 12.69	259.76 ± 128.14	1.64 ± 1.18
	133.76 ± 41.35	8.73 ± 4.55	19.62 ± 8.80	302.89 ± 128.07	1.75 ± 1.25
	60.65 ± 17.59	1.63 ± 1.46	21.13 ± 8.39	316.54 ± 141.63	1.68 ± 0.98
Eyes closing	536.56 ± 171.53	101.33 ± 52.16	4.50 ± 2.31	64.23 ± 30.72	1.37 ± 1.09
	453.54 ± 145.30	70.63 ± 33.87	4.80 ± 3.16	71.42 ± 45.04	1.35 ± 0.96
	539.61 ± 190.13	110.66 ± 66.42	6.26 ± 3.88	88.48 ± 51.88	1.55 ± 0.92
	364.95 ± 90.74	46.94 ± 19.14	3.98 ± 2.69	60.52 ± 41.72	1.71 ± 1.08
	186.69 ± 64.81	12.98 ± 9.59	9.40 ± 5.15	129.27 ± 60.59	1.64 ± 1.38
	269.64 ± 106.80	27.76 ± 22.58	8.25 ± 5.01	117.60 ± 75.87	1.55 ± 1.14
	276.07 ± 78.28	32.15 ± 15.20	9.92 ± 4.91	141.21 ± 68.00	1.68 ± 1.10
	114.24 ± 37.43	4.93 ± 2.99	8.49 ± 4.53	116.40 ± 62.74	1.73 ± 1.16
<i>p</i> -value	0.0078	0.0078	0.0078	0.0078	0.2343
<i>d</i> effect	1.68	1.43	5.25	5.91	0.34

**Table 1:** Performance of the variables (mean ± standard deviation) for the proposed method to detect eyes opening and eyes closing.

Wilcoxon Signed-Ranks test for paired samples and effect size methods were employed to evaluate the performance of the proposed method (Haidous *et al* 2013). For each event of eyes (eyes opening or eyes closing), it can be observed that there are statistical significance (*p*-value<0.05) and high effect size (*d* effect>0.8) in the variables that contain information of the variability (transition velocity, variance, total power, peak power), as well as the variance. It is worth to connect that the last subject presented a low difference in the variance. Figure 2 shows the coefficient computed throughout the signal. In these figures it can be seen that artefacts present during segments of open eyes do not affect the coefficient value.



**Figure 2:** Representation of the coefficient computed to detect eyes opening or eyes closing.

## 4 Conclusion

The preliminary results suggest that the combination of the variance and information variability can be used to detect events of eyes opening and eyes closing through EEG signals measured on the occipital lobe. Several works have reported that power spectral analysis of alpha wave is not a good option to detect events of eyes closing and eyes opening. Thus, this work presents the possibility of using the power spectral technique to analyze the variability information of the alpha wave. As future works, the proposed method should be improved to detect the peak values detection on the alpha wave. This approach here proposed can be also used to control the switching of BCI systems.

## References

- Kirkup L, Searle A, Craig A, McIsaac P and Moses P. (1997) *EEG based system for rapid on-of switching without prior learning*, Med. Biol. Eng. Comput., pp. 504–9.
- Kirkup L, Searle A, Craig A, McIsaac P and Larsen G. (1998). *Three methods compared for detecting the onset of alpha wave synchronization following eye closure*, Physiol. Meas., vol. 19, 213–224.
- Craig A, Tran Y, Craig D and Thuraisingham R. (2005). *Improving correct switching rates in a hands free environmental control system*, Journal of Neural Engineering, pp. L9-L13.
- Sakai M, Wei D, Kong W, Dai G, and Hu H. (2010). *Detection of change in alpha wave following eyes closure based on KM2O-Langevin equation*, International Journal of Bioelectromagnetism, 12(2), 89–93.
- Alaraj M and Fukami T (2013) *Quantification of subject wakefulness state during routine EEG examination*, International Journal of Innovative Computing, Information and Control, 9 (8), 3211–3223.
- Castillo J, Muller S, Caicedo E, Ferreira A and Bastos T (2014) *Proposal of a Brain Computer Interface to Command an Autonomous Car*. In: 5th IEEE Biosignals and Biorobotics Conference (BRC 2014), Salvador. Proc. of the 5th IEEE Biosignals and Biorobotics Conference, pp. 6.
- Haidous, N (2013) Robustness and Power of the Kornbrot Rank Difference, Signed Ranks, and Dependent Samples T-test. American Journal of Applied Mathematics and Statistics 1(5), 99–102

# Survey Results Regarding BCI Terminology

Brendan Allison<sup>1,2,\*</sup>, Jane Huggins<sup>3</sup> and Jaime Pineda<sup>1</sup>

<sup>1</sup>Cognitive Science Department, University of California, San Diego, USA

<sup>2</sup>Electrical and Computer Engineering Department, Old Dominion University, Norfolk, USA

<sup>3</sup>Department of Physical Medicine and Rehabilitation and Department of Biomedical Engineering, University of Michigan, Ann Arbor, USA

ballison@ucsd.edu, janeh@umich.edu, pineda@ucsd.edu

## Abstract

Different groups and projects have discussed official definitions and terms within the BCI community. We presented a survey at the Fifth BCI Meeting in 2013 that included a section on terminology. This paper discusses results to questions about the term “BCI” and its essential features, and terms such as “BCI Illiteracy”, “BCI inefficiency” or “BCI Proficiency”. While most respondents agreed on the term “BCI”, replies otherwise reflected significant disagreement. These survey results may facilitate discussion and understanding of different viewpoints, and encourage consensus on key terms and definitions.

## 1 Introduction

The term “Brain-Computer Interface” or BCI has been in use since the 1970s (Vidal, 1973). Numerous definitions of the term have been presented in the published literature and elsewhere. In addition, many other terms have been used to describe similar systems, such as “Direct Brain Interface,” “Brain-Machine Interface,” and “Brain Interface”. As BCI research gains attention in academic, commercial, medical, and other sectors, it seems increasingly important that we agree what a BCI is. A standardized definition is an obvious prerequisite for a mature field, and should not be especially daunting for a field dominated by scientists, engineers, and clinicians.

Other BCI researchers have agreed that a standardized definition is important. At the Fourth International BCI Meeting at the Asilomar Conference Grounds in Pacific Grove, California in 2010, over 65% of survey respondents felt that a standard definition of a BCI is needed within two years, and 79% felt that one is needed within five years (Nijboer et al., 2011). At the Fifth International BCI Meeting in the same location in June 2013 (organized by the Program Organization Committee for the Fifth International BCI Meeting), attendees unanimously voted to establish a BCI Society. This Society has actively sought to explore a BCI definition through online discussions, workshops at the BNCI Horizon 2020 retreat in Hallstadt, Austria in March 2014 (organized by the Graz University of Technology through the BNCI Horizon 2020 project), and other mechanisms. EU-funded projects, including two Coordination and Support actions called Future BNCI (2010-2011) and BNCI Horizon 2020 (2013-2015) have also been charged with developing, publicizing, and entrenching a standard BCI definition. This paper presents some results from a survey we conducted at the Fifth International BCI Meeting focused on two terminological issues – the definitions of “BCI” and “BCI Illiteracy”.

---

\* To whom correspondence should be addressed.

## 2 BCI: Term and Key Features

Our survey contained seven sections, and section three was titled “Terminology.” The first question asked: “Which term should be used to label BCIs?” 123 people answered this question. 93 people chose “Brain Computer Interface (BCI),” 24 people chose “Brain Machine Interface (BMI),” and the last two choices, “Direct Neural Interface” and “Brain-Neuronal Computer Interaction” each got three votes. Hence, consistent with our prior survey, results indicate a strong preference for the term BCI, with BMI a distant second and very little support for other terms (Nijboer et al., 2011).

The next question asked: “Which of the following elements is essential in a BCI? In other words: devices that do not have these elements are not BCIs. (Please check all that apply)” Table 1 summarizes the results.

Characteristic	# of respondents	% of respondents
<b>Must detect brain activity directly (before signal goes through peripheral nerves and muscles)</b>	101	77.10
<b>Must classify brain activity</b>	80	61.07
<b>Provides feedback to the user based (at least partly) on brain activity</b>	76	58.02
<b>Provides feedback in real-time, or near real-time</b>	76	58.02
<b>User must make voluntary choice to send each message or command</b>	43	32.82

**Table 1:** Essential components of a BCI, according to survey respondents.

These results reflect strong disagreement over the critical features of a BCI. These results echo the 2010 survey, which also found that respondents had widely divergent views. Other efforts to summarize different views of BCI components have also been presented (See Fig. 1). This figure was developed through discussion among the team that proposed and later executed the Future BNCI project, and reflects the efforts of several BCI experts to define critical features to foster discussion. This figure is consistent with the newer results presented above. Most respondents felt that a BCI must directly detect brain activity, but generally disagreed with the “intentional” component below.

A BCI requires four components:

- 1) Direct
- 2) Realtime
- 3) Realtime feedback
- 4) Intentional



**Figure 1:** The four components of a BCI, as defined in the proposal for the Future BNCI project, submitted in 2009. The project was active from Jan 2010 to Dec 2012 and extensively discussed the definition of a BCI and how to encourage an “official” definition within the field.

### 3 “BCI Illiteracy”

One of the most consistent challenges in BCI research has been dubbed “BCI Illiteracy” (Kübler and Müller, 2007). That is, a minority of users cannot use any particular BCI system, and may be as high as 20% (among healthy users) with some BCI approaches (Allison and Neuper, 2010). This problem has persisted since the earliest days of BCI research, across a wide variety of different types of sensors, mental activities and corresponding changes in brain activity, signal processing parameters, and other BCI features. This problem can be especially daunting for home use with less sophisticated systems. Several publications have argued that reliability is one of the most serious problems in the BCI community (e.g., Huggins et al., 2011, Wolpaw and Wolpaw, 2012; Allison et al., 2013).

At many BCI conferences and other discussion forums, many people have expressed concern about the term “BCI Illiteracy,” arguing that it is vague or implies that the problem should be blamed on the user. “We asked: Sometimes, a BCI system does not work for a particular user. Different terms have been used in the literature to describe this problem. What term do you prefer?” Of the 107 responses, 32 chose “BCI Illiteracy,” 33 chose “BCI Inefficiency,” 13 chose “BCI apraxia” and 29 chose “Poor BCI Proficiency.” These results indicate considerable disagreement over the best term to describe this phenomenon. Many respondents commented on this question, emphasizing that (as several people replied), “It is not the user’s fault!” We authors strongly agree and hope we can agree on a term that is descriptive and avoids blaming the user.

### 4 Summary and Comments

As with our 2010 survey, we have some concerns about our survey that could improve future survey efforts. Respondents were anonymous, which could lead to problems like multiple votes per person or “joke” voting. Some respondents did not answer all questions, and hence a shorter and more focused survey might have produced different results. Similarly, a longer survey might have elucidated more topics. Although we made a strong effort to present questions fairly, some biases in the questions and answer choices were present nonetheless. For example, the question about which term should describe BCIs explicitly used that term, and the question was asked at a conference with “BCI” in the name. New surveys might find new results among different venues and respondents.

These results show that respondents generally agree that the term “BCI” is preferable to other options such as “BMI”. However, there remains little agreement on what constitutes a BCI. Similarly, the community has divided views on which term should be used when someone can’t use a BCI. Part of the problem stems from the acronym itself – the term could refer to a very wide variety of systems. Characteristics such as feedback or closed-loop operation are not obvious from the acronym. Notably, many other terms within the BCI community, such as “synchronous” or “SSVEP” are not controversial, perhaps largely because these terms inherently specify defining characteristics.

We are concerned about the lack of accord on important terminological issues. The term “BCI” might be misused to gain attention, grant funding, or sales. Clear terms and definitions could help educate students, foster efficient discussion and collaboration, provide clear information to patients and other end users, and identify relevant publications. Thus, we encourage efforts to develop standard terms and definitions, including the emerging BCI Society (Ramsey et al., 2014) and BNCI Horizon 2020 Project. For example, the BNCI Horizon 2020 retreat included new surveys and six focus groups that discussed the BCI definition. We also hope to encourage effective interaction by better defining essential characteristics of a BCI, which can help people consider resulting examples of what is, or is not, a BCI. Our complete survey included follow-up questions that presented examples of systems that included different characteristics and asked respondents to consider whether

each example is a BCI. Such efforts might encourage more focused surveys, discussions in person and online, and other efforts to find consensus. The nascent BCI Society has also been discussing the BCI definition, and may consider actions such as forming a panel with an appropriate range of different people to develop official terms and definitions for the Society. We hope to work with our colleagues to reach consensus acceptable to most (if not all) of us in the BCI community.

## References

- Allison BZ, Neuper C. (2010). *Could anyone use a BCI?* In: Tan DS, Nijholt A, editors. *Applying our minds to human-computer interaction*. Springer, p. 35–54.
- Allison BZ, Dunne S, Leeb R, Millan J, and Nijholt A. (2013). *Towards Practical Brain-Computer Interfaces: Bridging the Gap from Research to Real-World Applications*. Springer Verlag, Berlin Heidelberg.
- Huggins J, Wren P, Gruis K. (2011) *What would brain-computer interface users want? Opinions and priorities of potential users with amyotrophic lateral sclerosis*. *Amyotrophic Lateral Sclerosis* (0):1–8.
- Kübler A, Müller K-R (2007). *Toward brain-computer interfacing*. In: *An Introduction to Brain-Computer Interfacing*. MIT Press, Boston, pp 1–25.
- Nijboer F, Clausen J, Allison BZ, and Haselager P. (2011). *The Asilomar survey: researchers' opinions on ethical issues related to brain-computer interfacing*. *Neuroethics*, 1-38.
- Ramsey, N., Allison, B.Z., Donchin, E., Sellers, E., Huggins, J., Millán, J Del R., Guger, C., Gao, S., and Wolpaw, J. (2014). *The Emerging Brain-Computer Interface Society*. Poster accepted for presentation at the Sixth International BCI Conference, Graz, Austria, Sep. 2014.
- Vidal JJ. (1973). *Towards direct brain-computer communication*. *Annu Rev Biophys Bioeng* 2:157–180.
- Wolpaw JR, Wolpaw EW. (2012). *Brain – Computer Interfaces Principles and Practice*. Oxford: Oxford University Press.

# The effect of task based motivation on BCI performance: A preliminary outlook.

Kelly E. Sheets, David Ryan, and Eric W. Sellers

East Tennessee State University

brownke@etsu.edu, ryand1@golmail.etsu.edu,  
sellers@etsu.edu

## Abstract

BCI technology can provide a non-muscular method of communication. In the P300-based BCI, users focus their attention on a specific item within an array of flashing items. After the BCI selects an item, it is presented on a computer monitor as feedback to the subject. In the BCI literature, only a few studies have examined motivation. The current study examined the effect of an intrinsic motivation induction on BCI performance in a non-disabled population. Participants were randomly sorted into two groups. Before participating in the experiment, one group (i.e., Motivation) was read a vignette describing the significance of BCI research for individuals with amyotrophic lateral sclerosis (ALS); the other group (i.e., non-Motivation) was not. The Motivation group performed at 89.7% accuracy in a copy-spelling task and the non-Motivation group performed at 80.6% accuracy. Further, data showed that BCI use causes a mild increase in sleepiness. Future studies should look into the role of motivation and sleepiness on BCI performance within the ALS population.

## Introduction

A brain-computer interface (BCI) using the P300 event-related potential offers an alternative method of communication. Originally introduced by Farwell and Donchin (1988), users are shown a matrix that randomly flashes alphanumeric characters. The participant's task is to focus attention on the item they wish to select. When the desired item flashes, a discriminable ERP is produced. Several paradigm and stimulus presentation modifications have improved speed and accuracy of the P300 BCI (e.g., Sellers, Arbel, and Donchin, 2013). However, few studies have examined psychological and situational factors that may affect BCI use. The current study examines the construct of motivation.

Motivation as a psychological construct is commonly divided into intrinsic and extrinsic categories. Intrinsic motivation implies that an individual pursues a goal for the enjoyment of doing so, while extrinsically motivated individuals pursue a goal in order to receive some sort of outside compensation for completing a task (Kruglanski et al, 1975; Gillet et al, 2011; Reiss; 2012).

Goldstein et al. (2006) used a monetary reward to show that P300 amplitude increase was positively correlated with the amount of money that could be earned. Kleih et al (2010) reported similar findings using a BCI task; P300 amplitude was higher for participants that reported higher a higher level of motivation. Monetary rewards can be problematic because they provide an external reward separate from simply succeeding at the task. Moreover, payment has been shown to decrease intrinsic motivation (Anderson et al, 1976).

In a clinical population, extrinsic motivation may not be as relevant. It is reasonable to propose that clinical use of the BCI in certain populations may be more intrinsic in nature. Kleih and Kluber

(2013) examined the effects of intrinsic motivation using verbal and visual prompts to explain the importance of BCI research for a clinical population. They found a significant effect of grouping, participants in the motivation condition reported higher levels of motivation; however, the groups did not differ in copy spelling accuracy.

In the current study, we focused on intrinsic motivation. It is likely that individuals who need or desire assistive communication will be more intrinsically motivated. Thus, we focused on task-based motivation, operationally defined as ‘willingness to participate in a task’ (Appel & Gilabert, 2002). In order to manipulate task-based motivation, we designed an information vignette in order to elicit a reaction from the participants that may increase their intrinsic desire to perform well. This is novel in that it does not provide extrinsic motivation or mention a performance standard; rather, the manipulation is in the form of a personal choice. Our hypothesis is that copy-spelling performance will be higher in the Motivation group (Mot) than in the non-Motivational (nMot) vignette group. Further, we hypothesize that those in the Mot group would have higher P300 amplitude than those in the nMot group.

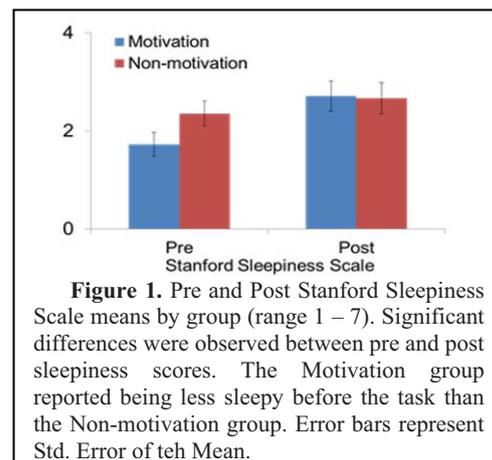
## Methods

Thirty-six non-disabled students (11 male, 25 female) from East Tennessee State University participated in the study. Students were enrolled in psychology courses and received extra course credit for participating. The study was approved by the East Tennessee State University Institutional Review Board.

EEG was recorded with a 32-channel tin electrode cap (Electro-Cap International, Inc.). All channels were referenced to the right mastoid and grounded to the left mastoid. Impedance was reduced to below 10.0 kOhm before recording. Two Guger Technologies g.USBamps were used to record EEG data, which were digitized at 256 Hz, and bandpass filtered from 0.5 to 30 Hz. Stepwise linear discriminant analysis was used to classify ERP responses. Only electrodes Fz, Cz, P3, Pz, P4, PO7, PO8, and Oz were used for online BCI operation (Krusiński et al., 2006).

Upon arrival, students were randomly assigned to one of two categories: Mot (n=18) or nMot (n=18). The Mot participants were told that it was essential that their full attention and effort could be applied to the task. They were given the option of leaving without penalty (i.e., they would be awarded full extra credit if they felt as though they could not perform at top capacity). It was reasoned that those who agreed to participate even after being informed that they could leave without penalty would be more motivated to perform the task than those performing the task without prompting. The Mot participants were then read a vignette describing the importance of BCI research for severely disabled people who may eventually depend on BCI technology. The nMot group was not given the opportunity to leave upon arrival nor where they read the vignette.

Prior to completing a copy-spelling BCI task, all participants completed the Online Motivation Questionnaire (OMQ) (Boekaerts, 2002) and the Stanford Sleepiness Scale (SSS; Hoddes et al, 1973). The OMQ measures task-based motivation in a learning environment. The BCI task was novel to all of the participants; thus, we assumed it would be a good measure of motivation. The SSS was selected

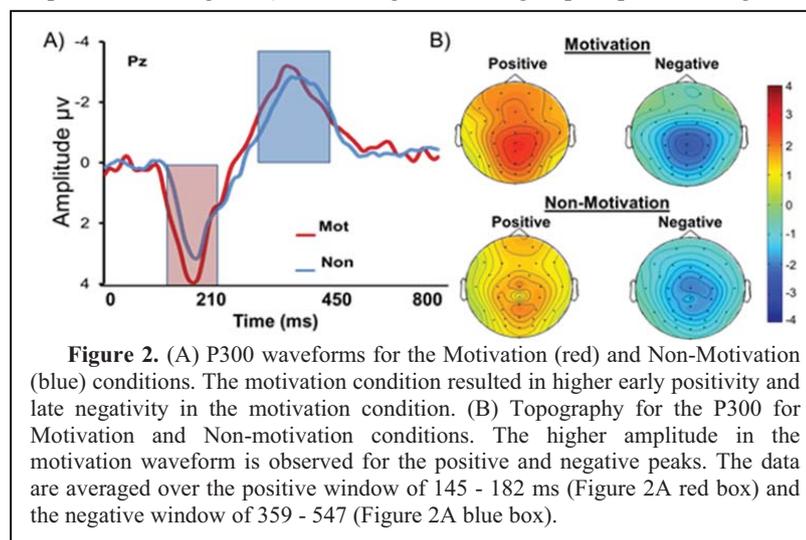


because BCI use may increase sleepiness, and there may be an interaction between motivation and sleepiness. After the participants completed the OMQ and the SSS, they were provided experimental instructions and completed the BCI task. Each participant copy spelled three randomly selected six-letter words. These data were used for calibration (SWLDA was used to derive the classification coefficients). After calibration, they copy spelled four randomly selected six-letter words and were provided with online feedback as to the accuracy of the BCI classification. When the BCI task was complete, each participant completed the SSS for a second time (i.e., post task).

## Results

A 2x2 mixed model ANOVA (group (Mot vs. nMot) x SSS (pre vs. post score)) was used to examine whether sleepiness scores varied across group or time. Only the main effect of SSS score was significant ( $F(1, 33) = 7.25, p = .004$ ; see Figure 1), indicating that both groups reported being more sleepy after the BCI session. However, mean scores were generally high; mean ratings were “awake but relaxed” after the session. Independent samples t-tests were used to examine motivation and accuracy. Motivation was statistically similar for the nMot and Mot groups ( $t < 1, ns$ ). Accuracy was also statistically similar across the nMot and Mot groups ( $t(27; adjusted) = 1.84, p = .072$ ); however, accuracy in the Mot condition (Mean=89.7, SD=11.2) was higher than the nMot condition (Mean=80.6 SD=17.2).

The waveforms for the Mot condition showed higher early positivity and later negativity than the nMot condition. Although not statistically significantly different, the differences were in the hypothesized direction. Figure 2A shows Pz waveforms for the Motivation and non-Motivation groups. Figure 1B shows topographies that correspond to the shaded boxes in panel 2A.



## Discussion

This study examined the effect of an intrinsically based motivation induction on BCI performance and waveform morphology. The study revealed several noteworthy findings. Most importantly, the group that was exposed to the motivation induction was 89.7% accurate, whereas the group that was not exposed to the motivation induction was only 80.6% accurate. Although the difference in accuracy was not statistically significant, we suggest that the current findings have high practical significance. For example, based on a Monte Carlo simulation conducted by Sellers et. al. (2006), to produce an

error corrected message containing 100 characters would require 130 selections at 90% accuracy and 170 selections at 80% accuracy. To put this finding into context, participants in the motivation condition would complete their message 12.45 minutes more quickly than participants in the non-motivation condition would complete their message. This is in spite of the fact that the motivation questionnaire did not reveal statistical significance. A possible explanation for this finding is that our motivation induction was not effective. However, this is unlikely because the ratings for both groups were near ceiling. Thus, the most likely explanation of this finding is that the questionnaire was not sensitive enough to detect existing differences in motivation. The waveform data support the ceiling effect explanation; higher early positivity and higher late negativity was observed for the motivation group.

The results also suggest that participants are sleepier after BCI use. Despite being sleepier, pre-experiment mean ratings were reported as: “Functioning at high levels, but not at peak; able to concentrate,” and mean post-experiment ratings were: “awake but relaxed; responsive but not fully alert.” It is recommended that future studies examine the effects of motivation and sleepiness in an ALS population. Presumably, people with ALS would have more intrinsic motivation to perform well. At the same time, people with ALS may become tired more quickly than non-ALS participants, which could cause frustration when sleepiness begins to affect accuracy.

## References

- Anderson, R., Manoogian, S.T., & Reznick, J. (1976). The undermining and enhancing of intrinsic motivation in preschool children. *J Pers Soc Psychol*, 34(5), 915-922.
- Appel, C., & Gilabert, R. (2002). Motivation and task performance in a task-based web-based tandem project. *Recall: Journal Of Eurocall*, 14(1), 16-31.
- Boekaerts, M. (2002). The on-line motivation questionnaire: A self-report instrument to assess students' context sensitivity. In P. R. Pintrich & M. L. Maehr (Eds.) *Advances in Motivation and Achievement, Volume 12: New Directions in Measures and Methods*. (pp. 77-120). New York: JAI / Elsevier Science.
- Farwell, L. A., & Donchin, E. (1988). Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials. *Electroen Clin Neuro*, 70(6), 510-523.
- Gillet, N., Vallerand, R. J., & Lafrenière, M. K. (2012). Intrinsic and extrinsic school motivation as a function of age: The mediating role of autonomy support. *Social Psychology Of Education*, 15(1), 77-95.
- Goldstein, R. Z., Cottone, L. A., Jia, Z., Maloney, T., Volkow, N. D., & Squires, N. K. (2006). The effect of graded monetary reward on cognitive event-related potentials and behavior in young healthy adults. *Int J Psychophysiol*, 62(2), 272-279.
- Hoddes, E. (1973). Quantification of sleepiness: A new approach. *Psychophysiology*, 10(4), 431-436.
- Kleih SC and Kübler A (2013) Empathy, motivation, and P300-BCI performance. *Front. Hum. Neurosci.* 7:642.
- Kleih, S. C., Nijboer, F. F., Halder, S. S., & Kübler, A. A. (2010). Motivation modulates the P300 amplitude during brain-computer interface use. *Clinical Neurophysiol*, 121(7), 1023-1031.
- Kruglanski, A. W., Riter, A., Arazi, D., Agassi, R., Montegio, J., Peri, I., & Peretz, M. (1975). Effect of task-intrinsic rewards upon extrinsic and intrinsic motivation. *J Pers Soc Psychol*, 31(4), 699.
- Reiss, S. (2012). Intrinsic and extrinsic motivation. *Teaching of Psychology*, 39(2), 152-156.
- Sellers, E.W., Arbel, Y., Donchin, E. (2012). P300 event-related potentials and related activity in the EEG. In J.R. Wolpaw and E.W. Winter Wolpaw (Eds.). Oxford University Press Inc. NY. (pp. 215-226).

# The Effects of Motivation on Task Performance Using a Brain-Computer Interface

Samantha A. Sprague<sup>1</sup>, David B. Ryan<sup>1</sup>, and Eric W. Sellers<sup>1</sup>

<sup>1</sup>East Tennessee State University, Johnson City, TN  
spragues@goldmail.etsu.edu, ryand1@goldmail.etsu.edu,  
sellers@mail.etsu.edu

## Abstract

Non-invasive brain-computer interfaces (BCIs) provide communication that is independent of muscle control. Thus, BCIs are a viable communication option for severely disabled people. Most BCI research has focused on improving signal processing methods and paradigm manipulations. Examining psychosocial factors may be particularly important for disabled people who have several co-morbidities. The current studies hypothesize that participants will have higher motivation in a free-spelling paradigm than in a copy-spelling paradigm. In Experiment 1, non-disabled participants completed copy- and free-spelling tasks. Motivation was measured three times during each session. In Experiment 2, ALS participants completed the same BCI tasks and motivation was measured at the beginning and end of the session using a visual analogue scale and the Questionnaire for Current Motivation for BCI2000. Significant differences in motivation were not observed in either sample. In Experiment 1, the participants that completed the free-spelling task first had significantly higher accuracy on the subsequent copy-spelling task. No differences were observed in performance accuracy between the two tasks in Experiment 2.

## Introduction

Noninvasive brain-computer interfaces (BCIs) use electroencephalogram (EEG) to provide non-muscular communication. A paucity of BCI research has considered how psychosocial factors affect task performance. Recent research suggests a need for studies that focus on specific qualities of BCI users, such as motivation and depression [Kleih et al., 2010; Nijboer, et al., 2010]. Examination of these factors may increase BCI performance by learning more about individual users, treating them as active participants, and addressing their specific needs. In the current studies, participants completed two tasks: copy- and free-spelling. After each task their motivation to perform the task was assessed with a visual analogue scale and one of two Likert scale measures. It is hypothesized that free spelling will lead to higher motivation and accuracy. This hypothesis is driven by the finding that autonomy in a task (such as free-spelling) is related to an increase in intrinsic motivation (Ryan & Deci, 2000). Learning about how and why a particular person may be motivated or unmotivated can help determine what tasks may lead to higher BCI performance by fully engaging participants in the task.

## Experiment 1

**Participants.** Non-disabled participants (n=21) were recruited from the East Tennessee State University (ETSU) subject pool. All participants provided informed consent and the study was approved by the ETSU Institutional Review Board.

**Materials.** Guger Technologies g.USBamps were used to record EEG data, which were digitized at 256 Hz, and bandpass filtered from 0.5 to 30 Hz. Stimulus presentation, EEG data collection, and online processing were controlled via BCI2000 [Schalk, et al., 2004]. Stepwise-linear discriminant analysis (SWLDA) was used for classification. Eight electrode locations were used: Fz, Cz, Pz, Oz, PO7, PO8, P3, P4.

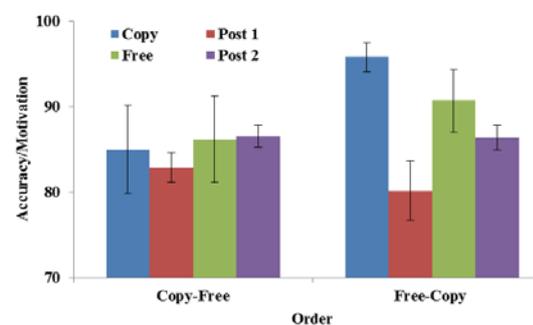
**Methods.** Participants were instructed to attend to a target item, by mentally saying or counting the item, when it flashed. Stimuli were presented in an 8x9 (72-item) matrix. Stimulus duration was 62.5ms with an SOA of 125ms (ISI=62.5ms). Copy- and free-spelling tasks were counter balanced for each subject. Before the copy- and free-spelling tasks, each participant copy spelled 18-characters without feedback. These data were then used for calibration. The resulting classifiers were used for online copy- and free-spelling tasks. The words used for calibration and in the online copy-spelling task were randomly selected from a database of 6,000 words.

The copy-spelling task consisted of three six-letter words presented as one string with a space in between each word (20 total selections). In the free-spelling task, participants constructed their own sentence. The sentence ranged from 20-24 characters. Before the free spelling task, each participant wrote his or her sentence on a sheet of paper (to calculate accuracy). They were also instructed to correct selection mistakes using the BACKSPACE character (denoted “Bs”). If the participant had not completed the sentence before the limit of 24 character selections the task was terminated by the experimenter in order to keep the number of character selections on the free-spelling task between 20 and 24 characters in length.

In addition to the BCI tasks, participants also completed two surveys: the Stanford Sleepiness Scale (SSS), which was used to measure fatigue, [Hoddes et al., 1973] and the On-Line Motivation Questionnaire (OLMQ), which was used to measure motivation [Boekaerts, 2002]. The SSS and OLMQ were administered three times throughout the session. The OLMQ consists of pre- and post-task components. The pre-task OLMQ was completed before the first BCI task, the post OLMQ was completed after each of the two BCI tasks. The SSS was completed before the first BCI task and after each of the two BCI tasks.

## Results

Performance accuracy for the copy-spelling condition was calculated by dividing the number of correct selections by the number of total selections. Performance accuracy for the free-spelling condition was calculated by comparing the message completed through the BCI to the intended message. Mean accuracy in the copy- and free-spelling conditions did not show a significant difference (91.19% and 88.71%, ( $t < 1$ )). No significant differences in OLMQ scores were observed at any of the three time points (pre, post 1, and post 2). Similarly, SSS scores were not significantly different across the three time points. For BCI accuracy there was an order effect. The participants who completed free-spelling first had significantly higher accuracy in copy-spelling task than the participants in the copy-free order ( $F(1, 19) = 5.617, p = .029$ ). Figure 1 shows the relationship between motivation and accuracy for the copy- and free-spelling tasks and post 1 and post 2 motivation.



**Figure 1:** Average performance accuracy for the copy-, and free-spelling tasks, post 1, and post 2 motivation scores organized by task order.

## Experiment 2

**Participants.** Participants with ALS ( $n=6$ ) were recruited based on participation in previous studies. All participants provided informed consent and the study was approved by the ETSU Institutional Review Board. ALSFRS-r scores were recorded for all participants indicating level of functionality ranging from 0 (completely locked-in) to 48 (fully functional). ALSFRS-r scores were as follows: 1, 5, 21, 30, 33, 33.

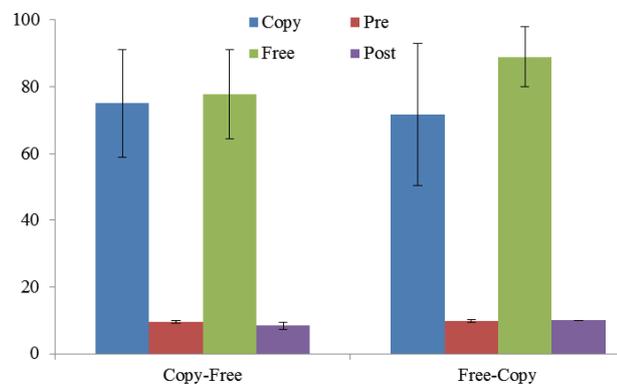
**Materials.** The materials were identical to Experiment 1 except that only one 16-channel amplifier was used and the cap contained 16 instead of 32 electrodes.

**Methods.** Participants wore a 16-channel electrode EEG cap and attend to a screen that presented a 6x6 item matrix. In contrast to Experiment 1, stimulus duration was 187.5ms and ISI was 62.5ms. In general, we have used longer stimulus durations when testing people with ALS [Townsend et al., 2010]. All other methods were the same as those used with the non-disabled participants.

Participants with ALS also completed two survey instruments: visual analogue scales were used to measure motivation and mood and the Questionnaire for Current Motivation for BCI2000 (QCMBCI2000) was used to measure motivation [Kleih, Nijboer, Halder, & Kübler, 2010]. The researchers found the QCMBCI2000 after the first experiment was completed and decided to use it as a measure of motivation for the second experiment since it has been previously used to measure motivation to perform a BCI task in people with ALS. The visual analogue scales for motivation and mood and the QCMBCI2000 were given to participants with ALS twice; once at the beginning of the session and once at the end of the session. At the conclusion of the session, participants were asked a qualitative question describing their motivation for participating in the study (i.e., “What was your motivation for participating in this study?”). This served as additional data to confirm participants’ responses to the surveys. The majority of participants described their main motivation as helping others to communicate.

## Results

Mean accuracy in the copy- and free-spelling conditions did not show a significant difference (73.33% and 83.33%, respectively ( $t<1$ )). The visual analog scale for motivation was not statistically different at pretest (9.60) and posttest (9.19;  $p=.52$ ). Similarly, the visual analog scale for mood did not show a significant difference at pretest (8.75) and posttest (8.47;  $p=.64$ ). Figure 2 shows the relationship between accuracy and motivation organized by the order in which the tasks were performed (i.e., copy first/free second or free first/copy second). No significant differences were observed in the four subscales of the QCMBCI2000.



**Figure 2:** Average performance accuracy for the copy-, and free-spelling tasks, pre- and post motivation scores organized by task order.

## Discussion

This study examined the hypothesis that participants will be more motivated in a free-spelling paradigm than in a copy-spelling paradigm, which would provide higher accuracy in a free-spelling

task. In Experiment 1, no differences were observed in motivation as measured by the OLMQ, or in fatigue as measured by the SSS. In terms of BCI performance, copy-spelling accuracy was higher for participants who completed the free spelling task first. The researchers speculate that this increase in accuracy on the copy-spelling task is due to a practice effect combined with a reduced workload. In other words, the free-spelling task requires the participant to perform four tasks: 1) pay attention to the flash of the character to be selected, 2) remember their entire statement, 3) know their position in that statement, and 4) decide the next character selection by evaluating the feedback (i.e., correct errors made by the system). In contrast, during the subsequent copy-spelling task there is only one task, pay attention to the target character. Free spelling acts as intensive training improving the allocation of attentional resources for the BCI task. When copy spelling follows this training, the participant has honed in their task related skills and the subsequent task becomes easier.

In Experiment 2, no differences were observed in motivation, mood, or BCI accuracy regardless of the order in which the tasks were performed. Unfortunately, a limitation of this study was that motivation was only measured twice. Nonetheless, people with ALS may have a higher personal investment than nondisabled participants, which results in uniformly high scores on the measures of motivation and mood. Therefore, regardless of BCI task, the inherent intrinsic motivation of a person with ALS does not vary with task demands. Based on the findings of this study, additional research that investigates specific tasks to improve attentional allocation may prove to be fruitful.

## Acknowledgements

This research is supported by NIH/NIDCD (R33 DC010470-03) and NIH/NIBIB & NINDS (EB00856).

## References

- Boekaerts, M. (2002). The on-line motivation questionnaire: A self-report instrument to assess students' context sensitivity. In P. R. Pintrich & M. L. Maehr (Eds.) *Advances in Motivation and Achievement, Volume 12: New Directions in Measures and Methods*. (pp. 77-120). Elsevier Science New York.
- Hoddes E, Zarcone V, Smythe H, Phillips R, Dement WC. Quantification of sleepiness: a new approach. *Psychophysiology*,10: 431-6, 1973.
- Kleih SC, Nijboer F, Halder S, Kubler A. Motivation modulates the P300 amplitude during brain-computer interface use. *Clin Neurophysiol*, 121: 1023-31, 2010.
- Nijboer, F., Birbaumer, N., & Kübler, A. (2010). The influence of psychological state and motivation on brain-computer interface performance in patients with amyotrophic lateral sclerosis – a longitudinal study. *Frontiers in Neuroscience*, 4(55), 1-13. doi:10.3389/fnins.2010.00055
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55, 68-78. doi:10.1037//0003-066X.55.1.68
- Schalk, G., McFarland, D. J., Hinterberger, T., Birbaumer, N., & Wolpaw, J. R. (2004). BCI2000: a general-purpose brain-computer interface (BCI) system. *IEEE Trans Biomed Eng*, 51, 1034-1043.
- Townsend, G., LaPallo, B. K., Boulay, C. B., Krusienski, D. J., Frye, G. E., Hauser, C. K., Sellers, E. W. (2010). A novel P300-based brain-computer interface stimulus presentation paradigm: Moving beyond rows and columns. *Clinical Neurophysiology*, 121, 1109-1120. doi:10.1016/j.clinph.2010.01.030