Proceedings of the 5th International Brain-Computer Interface Conference 2011

September 22-24 2011
Graz University of Technology, Austria

G. R. Müller-Putz, R. Scherer, M. Billinger,
A. Kreilinger, V. Kaiser, C. Neuper

Verlag der Technischen Universität Graz 2011

Laboratory of Brain-Computer Interfaces
Institut für Semantische Datenanalyse/Knowledge Discovery
Graz University of Technology, Austria

Verlag der Technischen Universität Graz
http://www.ub.tugraz.at/Verlag

ISSN: 2311-0422
ISBN (print): 978-3-85125-140-1
ISBN (e-book 2015): 978-3-85125-406-8
DOI: 10.3217/978-3-85125-140-1

PREFACE

This book contains the scientific contributions to the 5th International Brain-Computer Interface Conference (2011), held in Graz, Austria.

After the positive responses to the first four International Brain-Computer Interface Workshops and Training Courses we decided to change the style and expand the event into a Confernce, since we want to provide a platfrom 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 a poster presentation. For this purpose we encouraged the participants to submit papers, which were peer-reviewed and are published in this issue. The BCI meetings 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, Prof. Jonathan Wolpaw (Wadsworth Center, Albany, NY, USA), Nick Ramsey (UMC Utrecht, Rudolf Magnus Institute, Department of Neurology and Neurosurgery) and Dr. Donatella Mattia (Fondazione Santa Lucia, Rome, Italy) were able to accept our invitation to present keynote addresses at the Conference.

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

Andrea Kübler	Clemens Brunner
Brendan Z. Allison	Teodoro Solis-Escalante
Fabien Lotte	Moritz Grosse-Wentrup
Shin Kanoh	Donatella Mattia
Michael Tangermann	Febo Cincotti
Laura Astolfi	Luigi Bianchi
Robert Leeb	Theresa Vaughan
Benjamin Blankertz	Cuntai Guan
Ian Daly	Jeremy Hill
Peter Desain	Sonja Kleih
Tzyy-Ping Jung	Thorsten Zander
Carmen Vidaurre	Stefano Silvoni
Selina Wriessnegger	Eric Sellers
Ricardo Ron-Angevin	Dean Krusienski
Ricardo Chavarriaga	Femke Nijboer
Fabio Babiloni	Klaus-Robert Müller
José Millán	Jane Huggins
Francesco Piccione	Adrian Lee
Günther Bauernfeind	Petar Horki
Christian Breitwieser	Lisa Friedrich
Felix Darvas	

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.

We hope the content and scope of our program will contribute to a successful and constructive 5th International Brain-Computer Interface Conference 2011!

The editorial board

Group of the 4th International Brain-Computer Interface Workshop and Training Course 2008.

Contents

A Supervised Recalibration Protocol for Unbiased BCI

S. Perdikis[1], M. Tavella[1], R. Leeb[1], R. Chavarriaga[1], J. d. R. Millán[1]

[1]Chair in Non-Invasive Brain-Machine Interface, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

serafeim.perdikis@epfl.ch

Abstract

One important source of performance degradation in BCIs is bias towards one of the mental classes. Recent literature has focused on the general problem of classification accuracy drop, identifying non-stationarity as the generating factor, thus leading to several classifier adaptation approaches suggested as of today. In this work, we explicitly focus on bias elimination, demonstrating that the problem has two separate components, one related to non-stationarity and another one attributed to the nature of the feature distributions and the assumptions made by the classification methods. We propose a cued recalibration protocol including a supervised adaptation method and a novel framework for unbiased classification with a modified, unbiased Linear Discriminant Analysis classifier. Preliminary results show that our protocol can assist the subject to achieve quickly accurate and unbiased control of the BCI.

1 Introduction

Classification bias has proved to be a major problem in Brain-Computer Interfaces (BCIs), hindering user training and obstructing BCI operation, since one mental command can be heavily favored over the other, in which case the latter might often become unusable. However, bias elimination has received little attention per se so far, since in recent literature bias emergence has been only treated as part of the more general problem of accuracy degradation. Therefore, bias has been solely attributed to non-stationarity and thought to be largely eliminated by online adaptation of the classifier parameters [1, 2]. Alternatively, and although not reported in literature, the common code of practice in most labs in overcoming a biased classifier involves either a quick re-training session or "manual" adaptation of the classifier hyperplane.

Biasing effects become most prominent at the transition from the calibration (no feedback) to the online (BCI feedback) phase or in between consecutive online sessions, proving that non-stationarity accounts for a large component of the problem. Nevertheless, we discuss here other potential sources of bias and present a unified approach to tackle classification bias.

In this work, we present a novel method for supervised, adaptive estimation of Loss function parameters [3] leading to an unbiased Linear Discriminant Analysis (LDA) classifier. This supervised scheme will be applied in a cued recalibration protocol interleaved between the offline (calibration) phase and online operation of the BCI, thus achieving both classifier adaptation and explicit bias elimination for improved consecutive BCI experience.

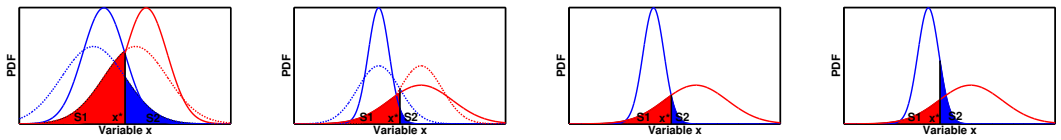
2 Methods

2.1 Motivation

Classification bias is evident from the confusion matrix of a BCI experiment where the per class accuracy may be significantly different, even for classifiers achieving high total accuracy. Such a biasing effect can occur for a variety of reasons. During online BCI operation bias can appear due to the violation of the stationarity assumption. In such case, the class distributions estimated on the training set and used to define the classifier's decision rule do not reflect any more the class distributions currently generated by the subject (Figure 1(a)), thus introducing bias and general accuracy degradation. Non-stationarity is a known problem in BCI and several adaptive

parameter estimation approaches have been proposed to eliminate its effects [1, 2], thus implicitly coping with this source of bias.

Nevertheless, classification bias can still persist, since there are additional reasons that may contribute to its appearance. These reasons include: (*i*) violation of the basic LDA assumption (when LDA is used) of identical covariance matrices for the two classes (Figure 1(b)), and (*ii*) the *inherent* bias emerging when classes are normally distributed with covariance matrices that are significantly different (Figure 1(c)). In the latter case, even a non-linear, quadratic classifier is not guaranteed to eliminate bias, since bias alleviation is not explicitly treated by Bayesian classifiers.



(a) Non-stationarity bias. (b) LDA assumption bias. (c) Inherent bias. (d) Unbiased decision rule.

Figure 1: Sources of bias for the case of a single feature and assuming equal prior probabilities without loss of generality. Dotted lines correspond to the sample distributions either estimated from the training set (a) or by assuming identical covariance matrices for the classes (b). Solid lines correspond to the actual sample distributions. Colored areas S_1, S_2 represent the error of the corresponding class. Potential bias is evident comparing the sizes of S_1, S_2.

The basic idea implemented in this work concerns a two-step algorithm, where in the first step a supervised LDA classifier adaptation technique is proposed to alleviate the non-stationarity-related bias and, in the second step, the LDA hyperplane constant term is further adjusted to eliminate the inherent- and assumption violation-related bias that might occur (Figure 1(d)).

2.2 Algorithm

Step 1 - Supervised Adaptation and Shrinkage: A supervised adaptation framework is employed, where for each incoming EEG sample $\mathbf{x}_t$ at time t the class mean vector $\boldsymbol{\mu}_i^t$ and covariance matrix $\boldsymbol{\Sigma}_i^t$ (where i the class that $\mathbf{x}_t$ belongs to) are iteratively estimated in the Maximum Likelihood (ML) approach as $\boldsymbol{\mu}_i^{t_i} = \boldsymbol{\mu}_i^{t_i-1} + \frac{1}{t_i+1}(\mathbf{x}^t - \boldsymbol{\mu}_i^{t_i-1})$ and $\boldsymbol{\Sigma}_i^{t_i} = \frac{t_i-1}{t_i}\boldsymbol{\Sigma}_i^{t_i-1} + \frac{1}{t_i}(\mathbf{x}^t - \boldsymbol{\mu}_i^{t_i-1})^T(\mathbf{x}^t - \boldsymbol{\mu}_i^{t_i-1})$ respectively, while the parameters of the other class retain their previous values. This straightforward supervised approach is possible, since data samples are acquired in a cued protocol where data labels are known. The "global" covariance matrix is also calculated at each step t as $\boldsymbol{\Sigma}^t = (t_1\boldsymbol{\Sigma}_1^t + t_2\boldsymbol{\Sigma}_2^t)/t$, $t = t_1 + t_2$, where iterators t_1, t_2 are only incremented when the sample at t belongs to the respective class i.

Parameters t_1, t_2 could generally be set to the respective sizes of the classes in the training set. However, we set this values to $t_i^0 = 1000$ in order to adapt faster to the distributions of the online session and recover quickly from any non-stationarity effect. Since for low values of t the ML estimates are known to be inaccurate and sensitive to outliers, we also employ an analytical covariance shrinkage method [4] to estimate the final covariance matrices $\widehat{\boldsymbol{\Sigma}_1^t}, \widehat{\boldsymbol{\Sigma}_2^t}, \widehat{\boldsymbol{\Sigma}}^t$. Shrinkage also avoids singularity problems. Then, we derive the conventional LDA hyperplane using a 0-1 Loss function[1] L_{0-1} at time t as $\mathbf{w_t}^T\mathbf{x} + b_{0-1,t} = 0$, where $\mathbf{w_t} = \widehat{\boldsymbol{\Sigma}^t}^{-1}(\boldsymbol{\mu}_1^t - \boldsymbol{\mu}_2^t)$ and $b_{0-1,t} = -\frac{1}{2}(\boldsymbol{\mu}_1^t - \boldsymbol{\mu}_2^t)^T\widehat{\boldsymbol{\Sigma}^t}^{-1}(\boldsymbol{\mu}_1^t - \boldsymbol{\mu}_2^t) + ln(\frac{P(\omega_1)}{P(\omega_2)})$.

Step 2 - Hyperplane Adjustment for Unbiased Classification: The main novelty of our approach lies on the introduction of a second adaptation step, aimed to alleviate the additional sources of bias based on the accurate estimation of the class distributions by the previous step. Figure 1(d) shows the basic idea consisting in finding a decision rule that "predicts" equal error rates for both classes, namely $P_e(c_1) = P_e(c_2)$.

[1] *0-1* subscript denotes that the quantity in question has been derived with a "0-1" Loss function, [3].

By constraining our problem to linear decision rules ${\mathbf{w}'_{\mathbf{t}}}^T\mathbf{x} + b_t = 0$ whose hyperplane is parallel to the one found by L_{0-1} LDA, $\mathbf{w}'_{\mathbf{t}} = \mathbf{w_t}$, the problem reduces to a single degree of freedom independently of the dimension of the feature space. Intuitively, we wish to estimate the bias term b_t of the new linear rule that will lead to theoretically equal error rates, thus operating our LDA classifier in a different point on the ROC curve than that found by conventional LDA.

The geometrical interpretation of the above demand satisfies that the hyper-volumes $P_e(c_i) = \int\int\ldots\int_{D_i} N(\mathbf{x}, \mu_i, \mathbf{\Sigma_i})\, d\mathbf{x}$, $D_i : sgn(c_i)(\mathbf{w^T x} + b) > 0$, are equal. The solution b_{unbias} is the zero of the function $f(b) = P_e(c_1) - P_e(c_2)$. This complex equation can be significantly simplified by considering a rotation $\mathbf{R}$ of the n-dimensional coordinate system of the feature space, such that the first dimension of the rotated space is parallel to the normal vector $\mathbf{w_t}$, in which case it is easy to show that $f(b) = \frac{1}{2}\,\mathrm{er}(\frac{b-m_1}{\sqrt{2\sigma_1^2}}) + \frac{1}{2}\,\mathrm{er}(\frac{b-m_2}{\sqrt{2\sigma_2^2}})$, where $m_i = (\boldsymbol{R}\mu_i)_1$, $\sigma_i^2 = (\mathbf{R}^T\mathbf{\Sigma_i}\mathbf{R})_{11}$ and er is the error function (proof omitted due to lack of space). Solving the last non-linear equation is possible by means of the Taylor approximation of er and polynomial root identification through the companion matrix formation, so that finally the only real root is a very close approximation of the desired bias term b_{unbias}.

Formally, the obtained solution b_{unbias} defines a Loss function $L_{unbias} = \begin{vmatrix} 0 & e^{b_{0-1}-b_{unbias}} \\ 1 & 0 \end{vmatrix}$ allowing our linear classifier to operate on the point that ideally produces equal error rates for both classes when the normal distribution assumptions hold, and thus zero bias. It is also worth to note that all necessary operations are simple enough to allow online implementation of the algorithm even in MATLAB for a BCI working at 16 Hz.

3 Results

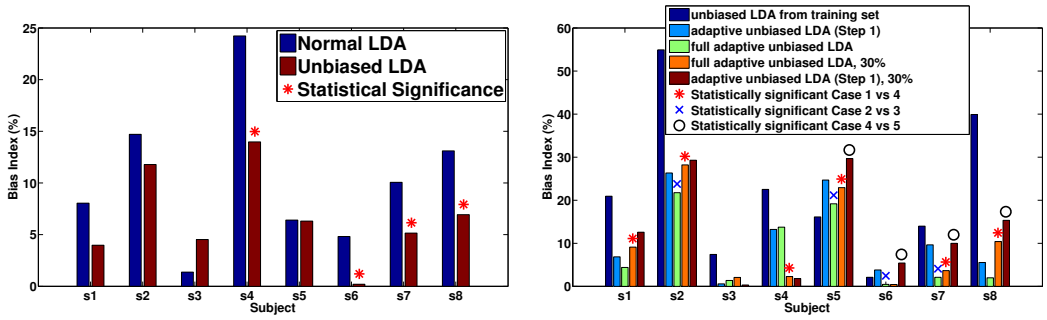
In order to evaluate the effects of the extra bias factors we have identified and the effectiveness of the unbiased LDA framework in alleviating them, we compute the *Bias Index* $BI = \left|\frac{a_1}{k_1} - \frac{a_2}{k_2}\right|$, where a_i the number of correctly classified samples and k_i the total number of samples of class i. An ideal, unbiased classifier should have $BI = 0$.

The first comparison is on a calibration session dataset of 8 subjects with a cued Motor Imagery (MI) protocol consisting of at least 120, each 5 sec long trials for each subject. BI is is calculated on the training set, thus largely excluding non-stationarity related bias. In this case we compare our unbiased LDA approach to the normal LDA.

The results illustrated in Figure 2(a) show the existence of additional bias factors, as well as the ability of the proposed method to largely eliminate them. The fact that BI does not reach 0, as theoretically predicted, as well as the exception of subject s_3, are attributed to the fact that the assumption of normally distributed features does not absolutely hold. In the same experiment total accuracy was not affected, since the maximum reported difference was found to be less than 1 % across all subjects and not statistically significant.

The ability of the proposed unbiased LDA framework to reduce bias is further demonstrated on extra (at least 120) MI trials executed the same day as the above calibration session for each subject. In this case, online feedback was driven by a Gaussian classifier and variable-length trials would end when a decision threshold on "accumulated" posterior probabilities was reached. The comparison on this dataset is done among the following variations of our unbiased LDA method: (i) unbiased LDA derived from the training set (calibration session) as above, (ii) adaptive unbiased LDA (Step 1 only) running over the whole dataset, (iii) full adaptive unbiased LDA (Steps 1 & 2) running over the whole dataset, (iv) full adaptive unbiased LDA running over the first 30 % of the dataset and stopped afterwards, and (v) adaptive unbiased LDA (Step 1 only) running over the first 30 % of the dataset and stopped afterwards. The last two cases are meant to evaluate the expected bias after the proposed re-calibration protocol has finished, where classifier (iv) would be the outcome of the proposed protocol. BI was calculated on the last 70 % of the dataset.

Figure 2(b) shows the overall potential of our framework, since for all subjects the proposed unbiased LDA (4^{th} bar) achieves a lower BI than the original unbiased LDA derived from the training set (1^{st} bar) for all subjects except s_5, who had a non-stationarity effect that made fea-



(a) Calibration session MI data.

(b) Online feedback session MI data.

Figure 2: Bias elimination through adaptive unbiased LDA on MI EEG data.

ture distributions change abruptly after adaptation was switched off. Differences are statistically significant for 5/8 subjects. Furthermore, although not shown in the figure due to space limitations, the unbiased LDA derived from the training set has a lower BI than the normal LDA for 7/8 subjects, 6 statistically significant. Statistical significance is computed at 99 % confidence interval. Additionally, it is verified that the extra "unbiasing" procedure can assist in further reducing the total bias during the protocol execution (3^{rd} bar lower than the 2^{nd} for 6/8 cases, 4 statistically significant) as well as after protocol termination (4^{th} bar lower than the 5^{th} for 6/8 cases, 4 statistically significant).

4 Discussion

The overall trends support the utility of the proposed method, while the fact that non-favourable exceptions are not statistically significant proves that in the worst-case scenarios the recalibration protocol will not inflict further bias. Concerning the total accuracy, results (not shown due to lack of space) showed that accuracy can greatly improve when non-stationarity is intense between calibration and feedback trials, otherwise there is no significant improvement. It should also be mentioned that our method can equivalently be applied in other types of features or problems.

Ongoing and future work entails online experiments under this protocol, where it can be hoped that the mutual learning procedure when the protocol directly drives the feedback can further improve performance. We will also explore an unsupervised version of the unbiased framework.

Acknowledgments

This work is supported by the European ICT Programme Project FP7-224631 (TOBI).

References

[1] J. d. R. Millán. On the Need for On-Line Learning in Brain-Computer Interfaces. In *Proceedings of the International Joint Conference on Neural Networks*, 2004.

[2] C. Vidaurre, C. Sannelli, K. R. Müller, and B. Blankertz. Machine-Learning Based Co-adaptive Calibration. *Neural Comput*, 23(3):791–816, 2011.

[3] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification*. Wiley-Interscience, 2 edition, November 2001.

[4] Y. Chen, A. Wiesel, Y. C. Eldar, and A. O. Hero. Shrinkage Algorithms for MMSE Covariance Estimation. *IEEE Trans Signal Process*, 58(10):5016–5029, 2010.

Spatio-Temporal Filtering for EEG Error Related Potentials

I. Iturrate[1], L. Montesano[1], R. Chavarriaga[2], J. del R. Millán[2], J. Minguez[1]

[1] Instituto de Investigación en Ingeniería de Aragón and Dpto. de Informática e Ingeniería de Sistemas, University of Zaragoza, Zaragoza, Spain
[2] Defitech Foundation Chair in Non-Invasive Brain-Machine Interface, EPFL, Lausanne, Switzerland

iturrate@unizar.es

Abstract

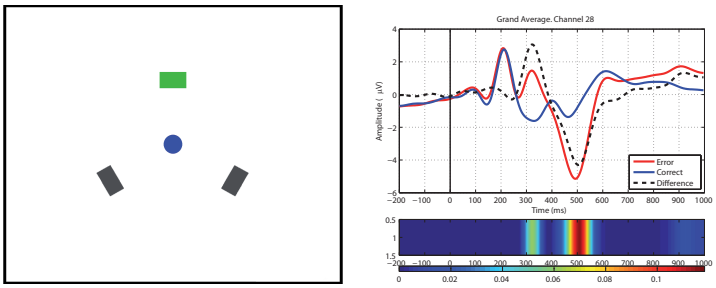
This paper presents a new filter for EEG Event-Related Potentials that relies on spatio-temporal features. The results are analyzed with error-related potentials, and compared to the original signal and with the independent component analysis spatial filter. Additionally, the obtained features are used for single-trial classification, showing that it is possible to obtain high classification accuracies with a very low number of features.

1 Introduction

Spatial filtering is a common pre-processing step in EEG-based BCIs. These filters rely on the use of different combinations of electrodes representing specific characteristics. For instance, principal component analysis (PCA) [1] is usually applied on the channel dimension, searching for a combination of electrodes that maximize the total variance of the signal; common spatial patterns (CSP) [2] filter searches for spatial patterns that maximizes the separability of the two classes; and independent component analysis (ICA) [1] extracts spatial filters based on the statistical concept of independence. Usually, CSPs are more commonly used in asynchronous signals such as $\mu - \beta$ rhythms [2], whereas ICA has demonstrated its feasibility for extracting components in event-related potentials (ERPs) [3].

Although these filters offer an automatic spatial combination of electrodes to use for classification, it is still necessary to manually choose the temporal (or frequency) information. Thus, a filter able to estimate both spatial and temporal features would be desirable. In this context, it has been studied the use of spectral information for CSPs to design spatio-spectral filters (namely CSSP and CSSSP) [4] for asynchronous signals. However, it is still an open issue on how to design spatio-temporal filters that automatically calculate both spatial and temporal combinations to separate the classes for ERPs.

In this paper, we present a spatio-temporal filter to extract discriminative information for ERPs. We tested the method on error-related potentials, signals that encode cognitive information about actions that a user has considered erroneous, and have well been studied from the neurophysiological point of view [5]. Additionally, it has been demonstrated their presence in very different situations detecting them on single-trial [6, 7]. Among the spatial filters usually used for BCI, the error responses have been successfully analyzed with independent component analysis [3]. Thus, here we analyze the filter obtained with the proposed spatio-temporal filter, and compare it with the original signal and with ICA. Additionally, we present single-trial classification performances with both ICA and the designed filter, showing that with the proposed filter it is possible to obtain high classification accuracies with a very low number of features.



(a) (b)

Figure 1: (a) Snapshot of the experiment performed. (b) Grand average, averaged for all the subjects, in channel FCz. Error, correct and error minus correct averages are shown.

2 Methods

2.1 Experimental Protocol

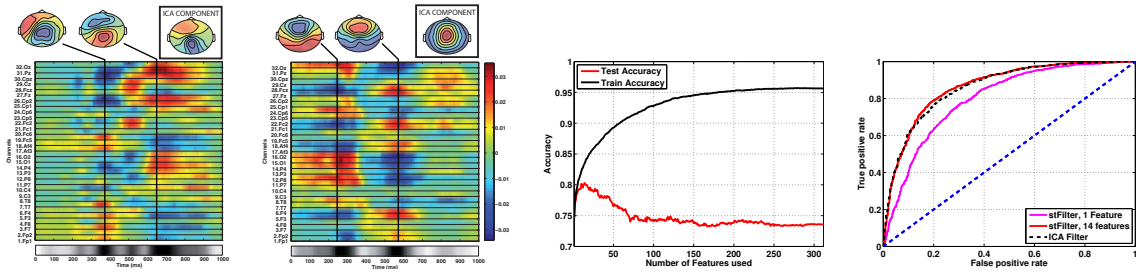
Four subjects performed the experiment, where they faced a computer screen. They were asked to monitor the actions performed by a virtual agent (in our case a blue dot), judging them as correct or incorrect actions. The experimental protocol is shown on Figure 1 (a). The blue dot performed movements to one of the squares. Each action lasted for one second with the blue dot staying on a square, and then returned to the central position for another second. For each run of the experiment, one of the squares was colored in green, indicating that a movement to that square was a correct one, whereas the other actions were considered erroneous. The probability of performing an erroneous action was of 0.2. A total of 36 runs were performed, changing the target location with each run. This leaded to 1440 and 360 correct and error responses respectively.

The data was acquired using a gTec system with 32 active electrodes with a sampling rate of 256 Hz. A power-line notch filter was applied to the signal. Then, a CAR filter and a low-pass filter with a cut-off of 10 Hz were applied. Finally, each ERP response was extracted within the time window $[0, 1000]$ ms, where the action started at t=0 ms. The grand average of the responses, averaged for all the subjects, is shown on Figure 1 (b).

2.2 Spatio-Temporal Filtering

The spatio-temporal filter (ST Filter) allowed us to obtain the spatio-temporal patterns that best separated the two classes. The main idea of the filter is the use of principal component analysis (PCA) [1] followed by feature selection. Previous to the filtering, it is needed to choose the n channels and time window of size m samples to perform the analysis. Thus, the input matrix size was $k \times n \times m$, where k represents the total number of ERP responses (1800 in our case). Then, the filter performs the following steps:

1. For each ERP response, the data is concatenated in a single feature vector of size $n \cdot m$. Thus, a 2D matrix of size $k \times (n \cdot m)$ is obtained.
2. Each feature $f \in [1 \ldots (n \cdot m)]$ is normalized within the range $[0, 1]$.
3. The normalized matrix is decorrelated using PCA, without performing any dimensionality reduction.
4. The score of each decorrelated feature is computed with the r^2 statistical measure. The features that best separate the data are retained.



(a) (b) (c) (d)

Figure 2: (a,b) ST weights of the first and second best features, and the most similar ICA component. (c) ST Filter accuracy as a function of the number of features. (d) ROC curves for ICA (dashed black) and ST Filter using 1 (magenta) and 14 (red) features.

3 Results

3.1 Filter Analysis

Here, we analyze the results obtained with the filter designed from the point of view of temporal and spatial activity. For this analysis, we used the 32 channels and the time window $[0, 1000]$ ms to extract the best two features according to the ST Filter. The weights of the PCA matrix corresponding to each of the two features were reshaped as a matrix and plotted as a color encoded image. Figures 2 (a,b) show the weights obtained for one participant.

Temporal Loadings: The cross-channel sum of weights is plotted on the lower part of Figures 2 (a,b) (dark colors indicate higher activation). This sum indicates the instants where the signal activation was higher for the specific feature. The results show that the first feature has a highest activation at approximately 350 ms and 650 ms. Note that this correspond to second and fourth peaks in the original difference grand average (see Figure 1 (b). The second feature's highest activation is present on the instants 250 ms and 550 ms (first and third peaks in the difference grand average).

Spatial Loadings: According to the previous cross-channel sum, the two scalp maps of the most representative timesteps are shown. The topographic plots show activity on the fronto-central and parietal areas, which has been suggested in the past to be related with error activity [5]. Thus, the filter seems to be able to extract useful information. Additionally, the scalp map of the most similar ICA component is also shown for each of the two features. The two features seem to be related to ICA components reflecting error processing [3]. This similarity is higher on the first feature, whereas the second feature shows a more frontal activity than the ICA component.

3.2 Single-Trial Classification

Here, we compare the ST Filter with the ICA Filter for single-trial classification with a simple LDA classifier. Next, the features used for each case are described:

- **ST Filter**: since the error potentials are known to be focused on fronto-central brain areas, we used that prior information for the input to the ST Filter: 8 fronto-central channels were chosen, within the time window $[200, 800]$ ms downsampled to 64 Hz, yielding to 312 features for each ERP response.
- **ICA Filter**: the ICA components were selected by visual inspection based on the results obtained in [3]. Depending on the subject, the components chosen varied between 2 and 3. The temporal information of each component was used within the time window $[200, 800]$ ms downsampled to 64 Hz, yielding to 39 features per component.

To validate the results, a ten-fold cross-validation strategy was applied. First, we analyzed the accuracy of the ST Filter as a function of the number of features. Figure 2 (c) (red) shows the accuracies obtained averaged for all subjects. The results show that a low number of features was enough to obtain high accuracies. The highest accuracy was obtained with 14 features. After adding more features the accuracy decreases suggesting an overfitting effect (continuous increase in the training accuracy, indicating that additional features mainly contributed to fit noise, Figure 2 (c) (black)).

Additionally, we computed the ROC curve [8] for the ICA features, and for the ST Filter when retaining 1 feature and 14 features. The Areas Under the Curve (AUC) were of 0.86, 0.80, and 0.87 respectively. This indicates that with the use of only one feature, the classifier was able to obtain high accuracies. Furthermore, with a very low number of features, the classifier obtained similar results than the ones obtained with the ICA features. However, note that ICA requires a manual selection of the components after applying the filter as such, whereas the filter proposed is completely automatic.

4 Conclusion and Future Work

In this work, we have presented a new filter for ERPs relying on spatio-temporal features. Despite the filter presents a complex spatio-temporal combination, the spatial and temporal loadings can be related to ICA and the original temporal signal. Additionally, the filter designed obtained high classification accuracies even with a low number of features without any manual feature selection.

The future work focuses on the use of the proposed filter for other protocols and EEG signals, and performing more thorough comparisons with different spatial or temporal filters.

Acknowledgments

This work has been partially supported by the Spanish Government through projects HYPER-CSD2009-00067, DPI2009-14732-C02-01 and CAI Programa Europa.

References

[1] A. Hyvärinen, J. Karhunen, and E. Oja. *Independent Component Analysis*. Wiley Interscience, 2001.

[2] J. Müller-Gerking, G. Pfurtscheller, and H. Flyvbjerg. Designing optimal spatial filters for single-trial EEG classification in a movement task. *Clin. Neurophys.*, 110(5):787–798, 1999.

[3] S. Debener, M. Ullsperger, M. Siegel, K. Fiehler, D. Y. Von Cramon, and A. K. Engel. Trial-by-trial coupling of concurrent electroencephalogram and functional magnetic resonance imaging identifies the dynamics of performance monitoring. *J. Neuroscience*, 25(50):11730, 2005.

[4] G. Dornhege, B. Blankertz, M. Krauledat, F. Losch, G. Curio, and K. Müller. Optimizing spatio-temporal filters for improving brain-computer interfacing. *Advances in Neural Information Processing Systems*, 18:315, 2006.

[5] M. Falkenstein, J. Hoormann, S. Christ, and J. Hohnsbein. ERP components on reaction errors and their functional significance: A tutorial. *Biological Psychology*, 51:87–107, 2000.

[6] R. Chavarriaga and J. d. R. Millán. Learning from EEG error-related potentials in noninvasive brain-computer interfaces. *IEEE Trans. Neural Syst. and Rehab.Eng.*, 18(4):381–388, 2010.

[7] I. Iturrate, L. Montesano, and J. Minguez. Single trial recognition of error-related potentials during observation of robot operation. In *Int.Conf IEEE Eng. in Medicine and Biology Society (EMBC)*, 2010.

[8] T. Fawcett. An introduction to ROC analysis. *Pattern recognition letters*, 27(8):861–874, 2006.

Group-Wise Stationary Subspace Analysis - A Novel Method for Studying Non-Stationarities

W. Samek[1,2,3], M. Kawanabe[2,1], C. Vidaurre[1] *

[1]Technical University of Berlin, Franklinstr. 28 / 29, 10587 Berlin, Germany
[2]Fraunhofer Institute FIRST, Kekuléstr. 7, 12489 Berlin, Germany
[3]Bernstein Center for Computational Neuroscience, Philippstr. 13, 10115 Berlin, Germany

wojciech.samek@tu-berlin.de motoaki.kawanabe@first.fraunhofer.de
carmen.vidaurre@tu-berlin.de

Abstract

In this paper we present an extension of the recently proposed Stationary Subspace Analysis (SSA). This novel method solves the problem how to group signals from different conditions and/or subjects to find stationary subspaces. The original SSA approach does not offer a natural way to group data and therefore better define the non-stationarities of interest. This drawback is solved with group-wise SSA (gwSSA) and demonstrated with a simple but illustrative example: the classification of BCI data. If not treated correctly the BCI tasks are considered as non-stationarities in SSA, which complicates its use for classification purposes. We show how, by correctly defining groups, non-stationarities of interest can be extracted. In this paper, the application is in multi-class signals, where the groups are properly defined to even improve classification performance.

1 Introduction

In Brain-Computer Interfacing (BCI) [1] one major challenge is to understand the non-stationarities in the signal of interest e.g. EEG and to develop methods which are invariant to them. The sources and time scales of non-stationarities in the signal can be very different e.g. changes in electrode impedance may occur when an electrode gets loose or the gel dries out, muscular activity or eye movements lead to artefacts in the signal and we often observe changes of task involvement as subjects get tired. Further changes in the EEG can be caused by differences between sessions e.g. no feedback in the calibration session vs. feedback in later sessions (or different kinds of feedback) or small differences in electrode positions between sessions. Several approaches were proposed to reduce the impact of non-stationarities in BCI applications. For example [2–5] use techniques for co-adaptive learning of user and machine, [6] uses extra measurement like EOG or EMG to be invariant against muscular or ocular artefacts and [7] applies covariate shift adaptation to account for changes of the features. Recently, Bünau et al. [8] proposed a novel technique called Stationary Subspace Analysis (SSA) which finds the low-dimensional projections having stationary distributions from high-dimensional observations. This method can be applied to EEG data as a preprocessing step in order to extract the stationary part of the signal as done in [9]. The authors showed that restricting the BCI to the stationary sources found by SSA can significantly increase the classification accuracy. However, SSA is a general purpose method and its usage is limited when applying it to multi-class data because the distinctive different tasks can be considered as non-stationary components of the signal (it is expected that the statistical properties of the data change with the task) and therefore disregarded. To avoid this, the data needs to be carefully

*We thank Klaus-Robert Müller and Paul von Bünau for valuable discussions. This work was supported by the German Research Foundation (GRK 1589/1), the European Union under the project TOBI (FP7-ICT-224631) and the Federal Ministry of Economics and Technology of Germany under the project THESEUS (01MQ07018). This publication only reflects the authors' views. Funding agencies are not liable for any use that may be made of the information contained herein.

preprocessed. In this paper we extend the work of Bünau et al. [9] and propose a group-wise Stationary Subspace Analysis (gwSSA) method which allows to compute the stationary subspace from different conditions and/or subjects. We analyse the emerging stationarity and non-stationarity patterns obtained from five volunteers and show that our method is better suited for multi-class data and consequently outperforms SSA.

This paper is organized as follows. In the next section we present SSA and introduce our group-wise approach. After that in Section 3 we apply it to a dataset of five subjects performing motor imagery and analyse the results in Section 4. We conclude in Section 5 with a discussion.

2 Group-Wise Stationary Subspace Analysis

Stationary Subspace Analysis (SSA) [8] is a novel method to factorize a high-dimensional multivariate time-series into its stationary and non-stationary components. Its underlying assumption is that the observed signal $\mathbf{x}(t)$ is a linear superposition of the stationary $\mathbf{s}^{\mathfrak{s}}(t)$ and non-stationary $\mathbf{s}^{\mathfrak{n}}(t)$ sources

$$\mathbf{x}(t) = A\,\mathbf{s}(t) = \begin{bmatrix} A^{\mathfrak{s}} & A^{\mathfrak{n}} \end{bmatrix} \begin{bmatrix} \mathbf{s}^{\mathfrak{s}}(t) \\ \mathbf{s}^{\mathfrak{n}}(t) \end{bmatrix}, \tag{1}$$

and A is an invertible matrix. The goal of SSA is to find a linear transformation $\hat{A}^{-1}$ that separates the $\mathfrak{s}$-sources from the $\mathfrak{n}$-sources. For that the signal $\mathbf{x}(t)$ is divided into epochs and an optimization criterion is employed to recover the sources. More precisely, SSA minimizes the distance measured as Kullback-Leibler-Divergence D_{KL}, between the distribution of the estimated $\mathfrak{s}$-sources in each epoch (described by first two moments) and the standard normal distribution.

The idea behind group-wise SSA (gwSSA) is to consider groups of epochs [1] in order to find projections which are as stationary as possible in each group. This does not necessarily imply stationarity across all epochs, however, it has the important advantage that one can combine data from many subjects to conduct group studies or one can group different conditions for a single subject, for example for classification. Since the objective function of gwSSA measures the distances between the distribution of the epoch and the mean distribution of the group, it can be written as

$$L(R) = \sum_{i=1}^{M} \sum_{j=1}^{N_i} D_{\mathrm{KL}} \left[\mathcal{N}(\hat{\boldsymbol{\mu}}_{ij}^{\mathfrak{s}}, \hat{\Sigma}_{ij}^{\mathfrak{s}}) \;||\; \mathcal{N}(\overline{\boldsymbol{\mu}}_{j}^{\mathfrak{s}}, \overline{\Sigma}_{j}^{\mathfrak{s}}) \right], \tag{2}$$

where M is the number of groups, N_i is the number of epochs in group i, $\mathcal{N}(\hat{\boldsymbol{\mu}}_{ij}^{\mathfrak{s}}, \hat{\Sigma}_{ij}^{\mathfrak{s}})$ is the distribution of epoch j in group i and $\mathcal{N}(\overline{\boldsymbol{\mu}}_{j}^{\mathfrak{s}}, \overline{\Sigma}_{j}^{\mathfrak{s}})$ is the average distribution in group i. As with SSA it is possible minimize the objective function by conjugate gradient descend in the space of antisymmetric matrices (see [8] for details).

3 Data

The data used in this paper consists of two calibration (i.e. without feedback) recordings from five healthy participants. The volunteers performed motor imagery of two limbs, specifically 'left hand' and 'foot'. The cues were presented either visually (with an arrow appearing in the centre of the screen) or auditory (a voice announcing the task to be performed), resulting in two different datasets for each user. In this experiment, the training data was the calibration with visual stimuli and the testing data, the calibration with auditory stimuli. The preprocessing parameters (frequency band and time interval) were subject-optimized in the training test, which consisted of 132 trials (equally distributed for each class). The testing set contained the same number of trials as the training dataset. The data was recorded with a multichannel system of 85 electrodes densely

[1] Epochs can be grouped according to different criteria e.g. each subject and/or condition can represent a group.

Methods / Subjects	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5
Without SSA	90.9	**80.0**	73.3	70.8	94.2
SSA (epoch = 1 trial)	90.9	60.0	**82.5**	70.8	82.5
SSA (epoch = 1 trial per class)	87.8	75.8	77.5	74.1	93.3
gwSSA (groups = classes)	**91.7**	78.3	80.0	**77.5**	**97.5**

Table 1: Comparison of classification accuracies for five subjects performing motor imagery. The SSA-based methods are applied as preprocessing step i.e. the bandpass-filtered EEG data are projected to the stationary subspace before performing CSP. The target dimensionality is selected via 5-fold cross-validation on the training data. From the results we see that our group-wise SSA method performs better than the other approaches, especially on subject 3, 4 and 5. The direct application of SSA (without trial concatenation) performs very poorly on subject 2 and 5, probably because it removes the differences between the classes.

covering the motor cortex at 1000 Hz. After filtering, it was down-sampled to 100 Hz. The features are extracted using log-band power on CSP filtered channels. The CSP filters were computed in the training test (three filters were selected per class). Finally, the classifier was Linear Discriminant Analysis (LDA). In the case of applying SSA or gwSSA, the band-pass filtered EEG data of the training set was used to feed the algorithm. The data was projected in the resulting stationary dimensions and after that the same feature extraction method as explained above was applied. SSA and gwSSA were restarted 50 times in order to avoid local minima and the dimensionality of the stationary subspace was selected via 5-fold cross-validation on the training set.

4 Results

In this section we compare our group-wise SSA to three baselines, namely SSA which takes every trial as an epoch, SSA which combines trials of opposite classes (as done in [9]) and no SSA. We apply gwSSA to each subject using two groups i.e. one group consist of trials which are labeled as 'left hand' whereas all 'foot' trials are in the other group. This is to assure that differences between both classes are not treated as non-stationarities and ignored. In Table 1 we see that gwSSA greatly improves the classification accuracies of subject 3, 4 and 5 while leaving the other subjects more or less at baseline level. Although SSA with trial combination also leads to higher classification accuracies in those subjects, it performs worse than gwSSA. We conjecture that this is because our method solves the problem in a principled way and does not require heuristics such as concatenation of trials with opposing labels. As expected the direct application of SSA (without trial combination) performs poorly as important differences between both classes may be treated as non-stationarities. In order to understand why group-wise SSA improves classification performance, we consider the changes which occur between training and test phase. In Figure 1 we plot the mean differences in power between the training and test features. Since we provided the visual cue only in the training phase, we observe large changes in the occipital areas. Although our SSA approach is computed on training data only, it greatly reduces these changes providing more stationary features. The most significant noise reduction can be observed for subjects who improve the most. By analysing the changes in power for each trial we can identify the locations of non-stationarities. These locations vary between subjects and can partially explain the results difference between subjects 1, 2 and 3, 4, 5. The first two subjects do not benefit from applying SSA as non-stationarities are located in areas which are important for discrimination (i.e. important information may be discarded), whereas in the other group the changes lie in the occipital area.

5 Conclusion

We presented an extension of the SSA algorithm which can identify stationary brain sources from groups of subjects and/or conditions. We showed that our group-wise SSA method can be naturally

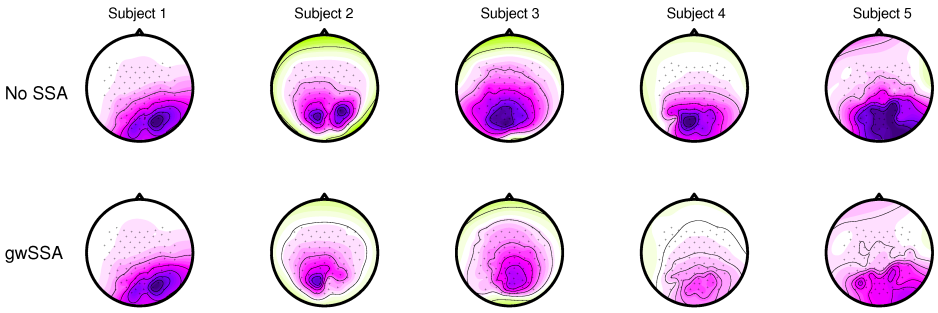


Figure 1: Scalp plots showing the mean difference in power between the training and the test data. When no preprocessing is applied there is a significant change between training and test features, most likely because the cue is showed with different types of stimuli (visual and auditory). Our approach reduces the shift, thus makes the signal more stationary but preserves the class related information contained in the signal. The effect is especially large for the subjects 3 and 4.

applied to multi-class signals and does not rely on heuristics to optimize its use. One important drawback of the heuristic approach applied in [9] is that it cannot be applied when the classes are not equally distributed. However, group-wise SSA solves the grouping problem in a principled way. Furthermore, it outperforms the standard SSA approach and the no SSA baseline for 3 out of 5 users. Finally, gwSSA can also be applied to group studies, because it optimizes stationarity in groups irrespectively of the grouping criterion. This means that stationary properties can be found in neuro-scientific data by simply defining the group of subjects as the target. In the future we want to use this tool to show how to analyze EEG data in group studies.

References

[1] G. Dornhege, J. del R. Millán, T. Hinterberger, D. McFarland, and K. R. Müller, editors. *Toward Brain-Computer Interfacing*. MIT Press, Cambridge, MA, 2007.

[2] C. Vidaurre, C. Sannelli, K. R. Müller, and B. Blankertz. Machine-learning based co-adaptive calibration. *Neural Comput.*, 2010. in press.

[3] A. Buttfield, P. W. Ferrez, and J. del R. Millán. Online classifier adaptation in high frequency EEG. In *Proceedings of the 3rd International Brain-Computer Interface Workshop*, 2006.

[4] J. Q. Gan. Self-adapting BCI based on unsupervised learning. In *3rd International Workshop on Brain-Computer Interfaces*, pages 50–51, 2006.

[5] S. Lu, C. Guan, and H. Zhang. Unsupervised brain computer interface based on intersubject information and online adaptation. *IEEE Trans Neural Syst Rehabil Eng*, 17(2):135–145, 2009.

[6] B. Blankertz, M. Kawanabe, R. Tomioka, F. U. Hohlefeld, V. Nikulin, and K. R. Müller. Invariant common spatial patterns: alleviating nonstationarities in brain-computer interfacing. In *Ad. in NIPS 20*, pages 113–120, 2008.

[7] Y. Li, H. Kambara, Y. Koike, and M. Sugiyama. Application of covariate shift adaptation techniques in brain-computer interfaces. *IEEE Trans. Biomed. Eng.*, 57(6):1318–24, 2010.

[8] P. von Bünau, F. C. Meinecke, F. Király, and K. R. Müller. Finding stationary subspaces in multivariate time series. *Physical Review Letters*, 103:214101, 2009.

[9] P. von Bünau, F. C. Meinecke, S. Scholler, and K. R. Müller. Finding stationary brain sources in EEG data. In *Proceedings of 32nd Conference of EMBS*, 2010.

Interactive Hierarchical Brain-Computer Interfacing: Uncertainty-Based Interaction between Humans and Robots

M. Chung[1], M. Bryan[1], W. Cheung[1], R. Scherer[2], R. P. N. Rao[1]

[1]Neural Systems Lab, Computer Science & Engineering, University of Washington, Seattle, USA
[2]Institute for Knowledge Discovery, BCI-Lab, Graz University of Technology, Graz, Austria

Abstract

Current non-invasive brain-computer interfaces such as those based on electroencephalography (EEG) [1] suffer from the problem of low signal-to-noise ratio, making fine-grained moment-by-moment control tedious and exhausting for users. To address this problem, we have previously proposed an adaptive hierarchical approach to brain-computer interfacing: users teach the BCI system new skills on-the-fly and these skills are later invoked directly as high-level commands, relieving the user of tedious lower-level control. However, the high-level commands learned from user demonstrations are often not reliable due to incomplete or insufficient data. In this paper, we address the unreliability of such learned high-level commands by proposing an interactive hierarchical BCI. The proposed approach utilizes an uncertainty metric in the learning algorithm to determine whether the learned high-level command is reliable enough to be performed in the present context. The BCI system interacts with the user to make the best decision at each stage. We illustrate the approach using an interactive hierarchical BCI for controlling a simulated wheeled robot. In a study involving two human subjects controlling the robot in a simulated home environment, each subject successfully used the system to complete a sequence of five different navigational tasks. Our results suggest that interactive hierarchical BCIs can provide a scalable and robust way of controlling complex robotic devices in real-world environments.

1 Introduction

Moment-by-moment control of EEG-based brain-computer interfaces (BCIs) over long periods of time can be a significant cognitive load and can exhaust the user. Therefore, EEG signals have often been used to select a task that can be semi-autonomously performed by an application (e.g., control of a humanoid robot in [2]). Recently, we introduced an adaptive hierarchical approach to BCIs, which combines the flexibility of fine-grained moment-by-moment control with the ease of high-level learned commands [3]. By incorporating learning algorithms, high-level commands (tasks that can be semiautonomously performed) can be learned from a user's lower-level control demonstrations. To illustrate the approach, we have focused on navigation problems, i.e., we teach robots new skills such as "Go to location A" by navigating the robot using the BCI to the desired destination. In order to create an internal representation, the machine needs a number of training trials. BCI control, however, is slow compared to manual control and thus data collection is very time consuming. As a result, training data is scarce, which makes accurate predictions difficult and the execution of high-level commands unreliable.

To overcome this problem, we introduce an interactive approach to hierarchical BCIs. We propose a system which possesses the ability to interact with the user whenever user guidance is required to successfully complete a task: the system's behavior relies on an uncertainty metric. The measure of uncertainty is obtained using a Gaussian Process (GP) model [4] for learning the high-level commands. When the uncertainty in a given region is too high, the BCI asks the user for further guidance rather than continuing to execute the unreliable and potentially dangerous high-level command. When the uncertainty in a region is low, after the user's guidance, the

BCI takes control from the user to complete the issued high-level command. The uncertainty-based interaction in hierarchical BCIs not only helps the system cope with unreliable high-level commands but also provides a smooth interface that combines the strength of the user and of high-level control.

We present results from user studies involving two human subjects who successfully completed five consecutive navigational tasks assigned to them. The users taught, controlled, and interacted with a wheeled robot in a simulated home environment. Our results provide a proof-of-concept demonstration that interactive hierarchical BCIs may provide a robust and flexible approach to controlling complex robotic devices.

2 System Architecture

The hierarchical BCI consists of three closely coupled systems: a SSVEP-BCI, the learning framework and the wheeled robot. Several enhancements were included compared to the system proposed in [3]. First, for the SSVEP-based BCI we use a support vector machine (SVM) classifier instead of selecting SSVEPs based on the absolute power in the spectrum. Second, we replaced the simulated Fujitsu HOAP-2 humanoid robot with a simulated K-Team Khepera wheeled robot. Since our experiments are focused heavily on the navigational aspect of mobile robots, a wheeled robot is more suitable for this purpose. The wheeled robot is pre-programmed with basic navigational functionalities such as driving forward, turning left, turning right, etc., and basic obstacle avoidance functionality. Third, the GP algorithm was used for learning high-level commands from logged position data during user demonstrations. During execution of a high-level command, the robot queries the user for navigation direction based on its uncertainty in its current position. When GP model is used, one obtains both a predicted mean value as well as the variance of the prediction. This variance can be related to the "confidence" in the learned model: high variance implies low confidence in the predicted navigational command and vice versa. For the current implementation, we used a simple threshold (determined empirically) to decide whether the robot should ask the user for guidance or take control from the user to finish a commanded tasks. Last, hierarchical adaptive menu structure [3] was updated to incorporate changes made in the other components.

3 Experimental Procedure

Two able-bodied male subjects (ages: 19, 21) participated in the study, which was approved by the University of Washington Institutional Review Board. All subjects gave written informed consent. Subjects did not have any experience with using our interactive hierarchical BCI system. However, they had been test subjects in SSVEP-based BCI experiments in the past.

The experiment was conducted over three days. On the first day, the subjects' SSVEP responses were characterized and the subjects were allowed to familiarize themselves with the SSVEP-BCI control. On the second day, subjects used the entire system to perform two assigned navigation tasks. The first task was teaching the robot how to navigate from the starting location, room A (upper left corner), to room B (upper right corner), which we called "skill1" (see Figure 1). The second task was teaching the robot how to navigate from its current location, room B, to room C (lower left corner), which we called "skill2". With skills 1,2 the subject provided only one training example. On the third day, subjects were assigned three different navigation tasks. The third task was navigating the robot to room B from its current location, room C, by using learned skill1. The fourth task was teaching the robot how to navigate from its current location, room B, to room D (lower right corner), which we called "skill3". The last task was navigating the robot to room C from its current location, room D, by using skill2.

Note that a skill represents a goal and a set of trajectories to reach that goal. This implies that a skill can be robust with respect to starting position. For instance, once a user defines skill1 with a goal of room B and a starting position A, they add additional information to that skill on

the third day to teach the robot to navigate to B from a third position, C.

To graph the confidence metric of learned high-level commands and the interactive behavior of the system, we collected navigational traces and and plotted confidence maps computed from the learned GP during the performance of the assigned tasks (Figure 1).

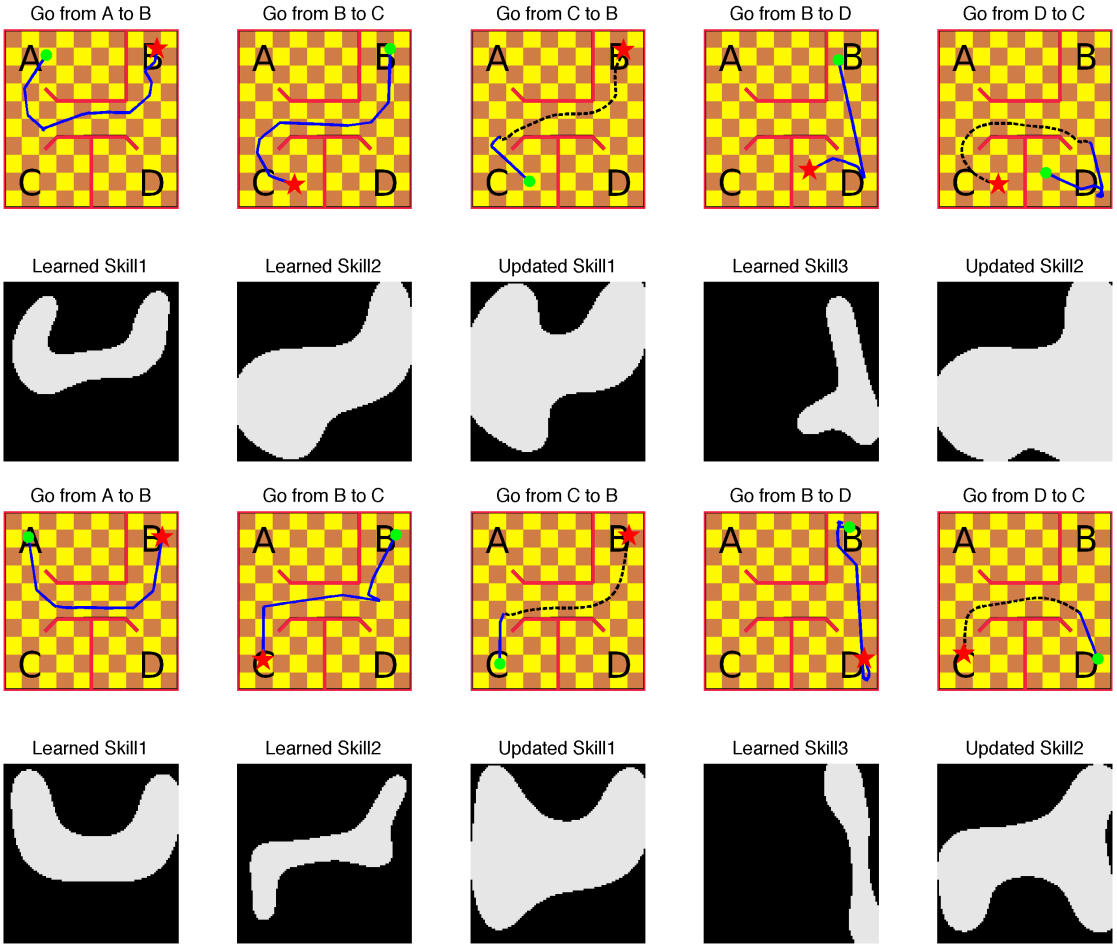


Figure 1: **Navigational traces and learning in the hierarchical BCI.** Columns correspond to assigned tasks. The first two rows are results from the first user and last two rows are results from the second. The first and third rows indicate actual navigational traces from the user. Solid lines are paths navigated by the user and dotted lines are paths navigated by the autonomous system. Dots mean starting positions and stars mean end positions. The second and fourth rows are confidence maps of the high-level skills learned during each task. Bright regions are high-confidence areas and dark regions are low-confidence areas.

4 Results

Both subjects were able to use the interactive hierarchical BCI to complete the five assigned tasks. The trace plots provide several interesting insights. First, the navigation traces from the autonomous system during the third and fifth tasks resemble the initial demonstrations provided by each user in the first and second assignment. This means the system is adaptive and can generalize based on the demonstrations provided by the user. Second, the generalization of performance is highly dependent on pre-chosen confidence thresholds on the variance of GPs. This is partially

because the small quantity of training data implies relatively large changes when new data is added. (Usually, training data includes more than one or two datasets.) Since the goal of this system is to learn new high-level skills on-the-fly, the BCI system starts learning after the very first example. The generalization process is determined by the parameters chosen to determine the high-confidence regions (i.e. white regions).

5 Conclusion

BCIs for robotic control have in the past faced a trade-off between cognitive load and flexibility. More robotic autonomy implied coarse-grained control and less flexibility, while fine-grained control provided greater flexibility but higher cognitive load. We propose a hierarchical approach which overcomes this tradeoff by combining the advantages of these two approaches. We introduced interactive components to increase the robustness of hierarchical BCIs against unreliable high-level commands.

Our results from the user studies using EEG-based interactive hierarchical BCIs confirms that (1) users can use the interactive hierarchical BCI to train a robot in a simulated environment, allowing learned skills to be translated to high-level commands, (2) the high cognitive load associated with fine-grained control can be alleviated by storing user-taught skills in a learned model for long-term use, allowing the learned skill to be selected and executed as a high-level command, (3) a probabilistic learning model (e.g., GPs) can be used to mediate the switch between high-level autonomous control and low-level user control, safeguarding against potentially catastrophic accidents, and (4) the hierarchical architecture allows the user to simultaneously control multiple devices, opening the door to multi-tasking BCIs. Our ongoing efforts are focused on 1) testing the approach with a larger number of subjects. 2) investigating its applicability to other challenging problems such as controlling a robotic arm with grasping capabilities. 3) exploring other types of brain responses (P300 and imagery) to achieve more natural and intuitive control.

Acknowledgments

This work was supported by the National Science Foundation (0622252 & 0930908), the Office of Naval Research (ONR), and the ICT Collaborative Project BrainAble (247447).

References

[1] J. del R. Millán, F. Renkens, J. Mouriño, and W. Gerstner. Noninvasive brain-actuated control of a mobile robot by human EEG. *Biomedical Engineering, IEEE Transactions on*, 51(6):1026–1033, 2004.

[2] C. J. Bell, P. Shenoy, R. Chalodhorn, and R. P. N. Rao. Control of a humanoid robot by a noninvasive brain–computer interface in humans. *Journal of Neural Engineering*, 5:214, 2008.

[3] M. Chung, W. Cheung, R. Scherer, and R. P. N. Rao. Towards hierarchical BCIs for robotic control. *Neural Engineering, 2011. NER'11. 5th International IEEE/EMBS Conference on (to appear)*, 2011.

[4] C. E. Rasmussen. Gaussian processes in machine learning. *Advanced Lectures on Machine Learning*, pages 63–71, 2004.

A Fast Feature Selection Method for High-Dimensional MEG BCI Data

M. Spüler[1], W. Rosenstiel[1], M. Bogdan[1,2]

[1]Wilhelm-Schickard-Institute for Computer Science, University of Tübingen, Sand 13, 72076 Tübingen, Germany
[2]Computer Engineering, University of Leipzig, Johannisgasse 26, 04103 Leipzig, Germany

spueler@informatik.uni-tuebingen.de

Abstract

Magnetoencephalography (MEG) is a rarely used technique for BCI, which benefits are good signal quality and high spatial resolution. The latter can be seen as a drawback, when it comes to selecting the most important features or sensors. To increase accuracy and reduce classification time feature selection is an important step in signal classification. Due to the large amount of MEG sensors (≥ 275 possible) feature selection is very time consuming, which is why we present a fast feature selection method for high-dimensional data. It gives similar results as established feature selection methods and we show that its time-complexity grows linearly with the number of samples and is $O(n \log n)$ for the number of features.

1 Introduction

A Brain-Computer Interface (BCI) enables a user to communicate or control a computer by brain activity only. Brain activity can be measured by different methods like EEG or MEG. Using MEG has the benefit of recording from more sensors, which also results in having more dimensions in the feature space. Depending on the method of feature extraction being used, there can be an additional growth of features. Considering a MEG with $n = 275$ channels a spectral estimation by an autoregressive model in the range of 7–21 Hz in 2 Hz bins would result in $n \cdot 7 = 1925$ features. Using connectivity measures like coherence or phase synchronisation [1] results in $\frac{n \cdot (n-1)}{2} = 37675$ features. When dealing with such an amount of features it is important to have a method for feature selection, which reduces the full set of features to a small set including only the features that are most important for classification [2]. Another demand results from the limited computation power and time available. In most BCI experiments data is recorded in a first session and used for training a classifier, which is tested in an online-session some minutes later. In this case we need a fast feature selection method that can deal with a large number of features very quickly.

2 Methods

2.1 Data

For analysis and comparison of different feature selection methods we used the dataset from [3]. It consists of MEG data measured from 10 subjects. Each subject participated in 2 sessions and performed 7 different imagination tasks without feedback (51 trials per task and session). Since it was shown in [3] that imagined right hand movement and subtraction were the two best tasks to classify, we focused on data from these two tasks for evaluation. For feature extraction we used spectral coherence [1] resulting in 37675 features for all 275 channels.

2.2 r^2 Ranking

In a BCI context r^2 values are often used for performance estimation [4]. They can be seen as measure of correlation between a feature and the class membership. In other words: they give the proportion of variance for a feature, that is explained by the class membership.

Assume our data consists of n trials $x_i, i = 1, ..., n$ with each trial having m features $f_{ij}, j = 1, ..., m$ and a class label $y_i \in \{1, -1\}$. If n_1 trials are in class 1 and n_{-1} trials are in class -1, the r^2 value for feature j is calculated by:

$$r^2(j) = \left(\frac{(\sum_{i,y_i=1} f_{ij})^2}{n_1} + \frac{(\sum_{i,y_i=-1} f_{ij})^2}{n_{-1}} - \frac{(\sum_{i=1}^{n} f_{ij})^2}{n}\right) \cdot \left(\sum_{i,y_i=1} f_{ij}^2 + \sum_{i,y_i=-1} f_{ij}^2 - \frac{(\sum_{i=1}^{n} f_{ij})^2}{n}\right)^{-1}$$

The result is a value between 0 and 1 where 0 stands for no correlation and 1 for perfect correlation (although perfect class seperability can be reached with $r^2(j) < 1$). When using r^2 values for feature selection, the r^2 values for all features are calculated, sorted in descending order and the features with the highest r^2 values are chosen for classification.

2.2.1 Incremental, Decremental Update of r^2 Values

R^2 values allow the possibility for incremental or decremental update. Having calculated the r^2 values for n samples, the solution for $n + 1$ or $n - 1$ samples can be calculated in constant time. By keeping in memory the following 6 intermediate results $(\sum_{i,y_i=c} f_{ij})^2$, $\sum_{i,y_i=c} f_{ij}^2$ and n_c for $c \in \{1, -1\}$ each incremental and decremental step can be calculated in constant time (20 numerical operations) by updating these intermediate results. Therefore the time complexity for calculating r^2 values grows linearly with the number of samples. Since the r^2 value for one feature is independent from all other features, each r^2 value can be calculated separately and the time complexity is also linear in respect to the number of features. Due to the ranking, where the r^2 values are sorted in descending order, the total time complexity for r^2-ranking is $O(n \log n)$.

The possibility of incremental and decremental update can be used for faster implementation of feature selection in cross-validations (CVs) or leave-one-out estimations (LOOEs). To ensure valid results a feature selection has to be performed on every training set each fold, which means performing n feature selections on a n-fold CV. By using decremental update it is possible to calculate the r^2 values with all trials only once and store the intermediate results in memory. In each fold the decremental update is used to remove the test trials from the full solution. The result is the same as the r^2 values would be calculated from the training set, but especially for LOOEs this is a big speedup, since the feature selection only needs to be done once and the results for each fold only need to be updated decrementally.

2.3 Performance Evaluation

To evaluate performance of r^2-ranking we tested some other feature selection methods: Recursive Feature Elimination (RFE) [5], Particle Swarm Optimization (PSO) [6], Principal Component Analysis (PCA) and Fast Correlation-Based Filter (FCBF) [7].

RFE is an iterative method that can be used with a Support Vector Machine (SVM). In each iteration a SVM is trained, a ranking criterion (feature weights in the linear case) is calculated and the feature with the smallest ranking criterion is removed. The RFE continues until the designated number of features is reached. To speed up the process we also used a modified version of the RFE in this paper (called mRFE later), which removes the 25 features with the lowest ranking criterion if the current number of features is greater than the designated number of features plus 100. After that only one feature is removed each iteration.

To compare computation time we used an Intel Dual Core E2180 at 2 GHz with 4 GB RAM running 64-bit Linux and 64-bit Matlab. Apart from LibSVM [8] only Matlab functions were used. We tested different numbers of starting features ranging from 250 to 37675 and selected the 100

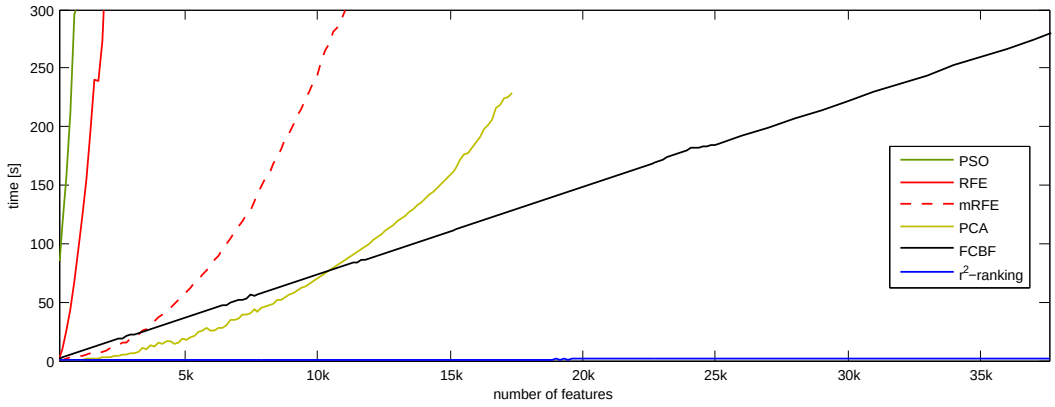


Figure 1: Averaged computation time to select the 100 best features for r^2-ranking, FCBF, mRFE, RFE, PCA and PSO with 204 trials. PCA could not be tested with more than 17350 features because the machine that was used for evaluation ran out of memory.

	optimal number of features			100 features	
method	# features	accuracy	time	accuracy	time
no feature selection	16110	80.2 %	-	-	-
FCBF	73	83.9 %	127 s	-	-
RFE	550	87.6 %	6986 s	83.9 %	7019 s
mRFE	550	87.5 %	608 s	84.4 %	261 s
r^2-ranking	330	86.5 %	0.79 s	83.8 %	0.79 s
PCA	50	83.1 %	146 s	78.5 %	146 s
PSO	9565	83.1 %	35758 s	-	-

Table 1: Comparison of different feature selection methods: number of features, accuracy and computation time needed for feature selection.

best features. For each method and each number of features we performed 10 runs with 204 trials and averaged the computation time.

To compare classification accuracies we used the first session of the data for feature selection and training and the second session to test the classifier. For classification we used LibSVM with a RBF-kernel. To test accuracy we excluded the outer MEG-sensors from the data, since they are supposed to have very little class related information but instead being more contaminated by artefacts than the inner sensors. This reduces the number of features to 16110.

3 Results

The results for computation time can be seen in Figure 1. It shows that r^2-ranking is faster than the other methods, especially for very high-dimensional data. Due to memory limitations it was not possible to calculate PCA with more than 17350 features on our machine.

The classification accuracy, number of features and computation time averaged over 10 subjects can be seen in Table 1 for the optimal number of features as well as a fixed number of 100 features. Since FCBF and PSO do not reduce to a specified amount of features, there are no results with 100 features.

While PCA, RFE and r^2-ranking show significantly better accuracy than without feature selection ($p < 0.01$, paired t-test), FCBF and PSO are not significantly better ($p > 0.1$). In addition there is no significant difference between r^2-ranking and RFE ($p > 0.25$).

To test the benefit of the decremental update we performed a LOOE with 204 trials and 37675 features. While the r^2-ranking without decremental update took 160.2 seconds, with decremental

update it finished in 2.8 seconds. For comparison we also tried a 10-fold CV without decremental update that took 9.5 seconds. All times reported here only relate to the computation time needed for the feature selection, not for the whole LOOE or CV.

4 Discussion

The results show that r^2-ranking is faster than the other tested methods. In terms of accuracy there is no significant difference between r^2-ranking and RFE, but r^2-ranking yields significantly better results than without feature selection, thus making it a viable option for feature selection in high-dimensional data.

Another advantage of r^2-ranking is the possibility for incremental and decremental updates. Feature selection in CVs or LOOEs usually is the most time consuming step in the whole process. While the decremental update has proven useful to further speedup CVs or LOOEs, the incremental update could be used for an online feature selection, that updates the feature ranking every time a new trial is available without the need to know the previous trials. While PCA already has shown memory problems, the same problem arises for RFE if the number of trials becomes larger. Although the memory requirements for r^2-ranking are lower in general, the possibility for incremental update could help to further aid this problem.

5 Conclusion

We showed on high-dimensional data that r^2-ranking is superior to other feature selection methods concerning computation time. In terms of accuracy there is no significant difference to established feature selection methods like RFE. The possibility of incremental and decremental update also shows other interesting areas for application, like faster cross-validation or online feature selection.

References

[1] E. Gysels and P. Celka. Phase synchronization for the recognition of mental tasks in a brain-computer interface. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 12(4):406 –415, 12 2004.

[2] M. Bensch, M. Bogdan, and W. Rosenstiel. Phase synchronization in MEG for brain-computer interfaces. In *In Proceedings of the 3rd Int. Brain-Computer Interface Workshop*, pages 18–19, Graz, 09 2006.

[3] M. Bensch, J. Mellinger, M. Bogdan, and W. Rosenstiel. A multiclass BCI using MEG. In *Proceedings of the 4th Int. Brain-Computer Interface Workshop*, pages 191–196, Graz, 09 2008.

[4] H. Sheikh, D. J. McFarland, W. A. Sarnacki, and J. R. Wolpaw. Electroencephalographic (EEG)-based communication: EEG control versus system performance in humans. *Neuroscience Letters*, 345(2):89 – 92, 2003.

[5] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik. Gene selection for cancer classification using support vector machines. *Machine Learning*, 46:389–422, 2002.

[6] C. J. Tu, L. Y. Chuang, J. Y. Chang, and C. H. Yang. Feature Selection using PSO-SVM. *IAENG International Journal of Computer Science*, 33:111–116, 2007.

[7] L. Yu and H. Liu. Feature selection for high-dimensional data: a fast correlation-based filter solution. In *Proc. 20th Int'l Conf.Machine Learning*, pages 856–863, 2003.

[8] C. C. Chang and C. J. Lin. *LIBSVM: a library for support vector machines*, 2001. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.

Toward Incorporating Anatomical Information in BCI Designs

E. Larson[1], A. K. C. Lee[1]

[1]Institute for Learning and Brain Sciences & Department of Speech and Hearing Sciences, University of Washington, Seattle, WA

akclee@uw.edu

Abstract

This paper aims to provide a proof of concept that anatomical information can be leveraged in BCI designs. Testing with simulated brain data, we used anatomical data from subjects to build classifiers that perform well in low-SNR conditions without subject-specific training.

1 Introduction

Recently, it has been postulated that source localization could be incorporated in BCI designs [1], or perhaps more generally, that we could use neuroscience information to improve BCI. We examined whether anatomical information - when coupled with EEG co-registration, source localization estimates, and surface-based inter-subject morphing - could be utilized to build BCI classifiers. We used subject structural brain scans to construct anatomical BCI classifiers without subject-specific training. In tests using offline-simulated brain activity, the anatomical classifier performed favorably in low-SNR, low-trial-count conditions compared to a standard BCI classifier.

2 Methods

Nine healthy subjects participated[1] in a previous study using EEG [2]. As detailed below, for each subject we used anatomical scans to construct a source space for cortical brain signals, and co-registered anatomical and EEG sensor locations. We used this process to simulate cortical activity in different brain areas with controllable SNR, to transform brain activity from one subject to another, and to perform detection of the corresponding EEG signals against a noisy background.

2.1 Anatomical Information and Co-Registration

We took three structural magnetic resonance imaging scans per subject to highlight tissue contrast.[2] After reconstructing the cortical surface and outer/inner skull surfaces using FreeSurfer [4], we constructed boundary element models (BEMs) to set up a source space in the brain consisting of ~ 7000 dipole sources oriented normally to the cortical surface along the segmented grey/white matter boundary. We also recorded locations of cardinal landmarks (nasion, left/right periauriculars), EEG cap electrodes, and extra points on the subject's scalp for anatomical co-registration. Using the linear colocation approach [5], we then calculated the forward solution gain matrix $\mathbf{G}$ for each subject that mapped cortical dipole currents $\mathbf{j}$ to EEG sensor activity $\mathbf{x}$, allowing us to simulate EEG sensor activity that would have been observed in response to arbitrary brain signals.

[1]Each subject gave informed consent according to procedures approved by Massachusetts General Hospital.

[2]Using a 1.5-T Avanto scanner, $\sim$ 8 min/scan. Multi-echo magnetization-prepared rapid gradient echo scan settings: flip angle = 7 °; $TR = 2.73\,s$; $TE = (1.64 + 1.86 \times n)$ ms, $n = 0 - 3$; acc. factor = 2, voxel size = 1 mm isotropic; fast low-angle (5 ° and 30 °) shot scan: $TR = 20$ ms, $TE = (1.85 + 2 \times n)$ ms, $n = 0 - 6$; acc. factor = 2, voxel size = 1 mm isotropic. This provided different tissue contrasts for constructing the BEMs [3].

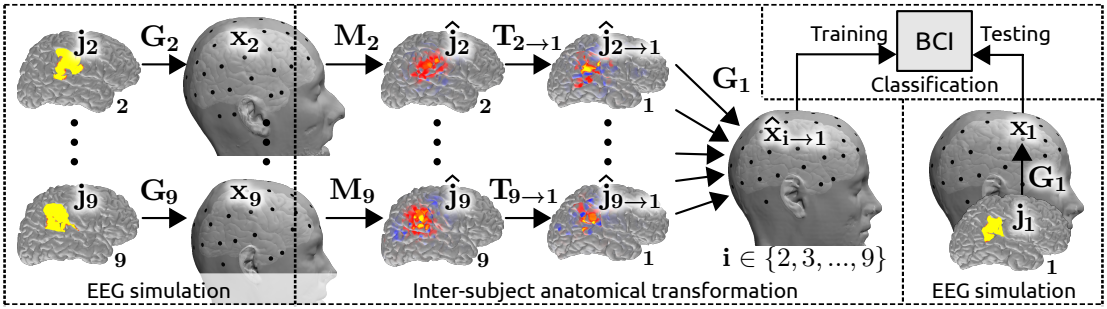


Figure 1: This diagram illustrates the brain/EEG signal simulation, and how the anatomical BCI classifier (here for Subject 1) pools data from other subjects. First, to calculate the observed EEG signals (left and lower right), each subject's vector of brain activity $\mathbf{j}$ is multiplied by a gain matrix $\mathbf{G}$ to calculate the corresponding observed EEG signals $\mathbf{x}$. By leveraging anatomical information, EEG activity can be mapped between subjects (middle): in the top row, Subject 2's original (unobserved) brain activity $\mathbf{j_2}$ is estimated from the EEG activity $\mathbf{x_2}$ using the inverse operator $\mathbf{M_2}$. The inter-brain transformation matrix $\mathbf{T_{2\to 1}}$, calculated using surface-based spherical mapping [6], then estimates Subject 2's equivalent activity on Subject 1's brain $\mathbf{\hat{j}_{2\to 1}}$, which is transformed to EEG activity $\mathbf{\hat{x}_{2\to 1}}$ using the forward solution $\mathbf{G_1}$. The BCI classifier (upper right) then trains using these estimated data and tests on the simulated EEG data $\mathbf{x_1}$ for Subject 1.

2.2 Brain Activity Simulation

Using the gain matrix $\mathbf{G}$ for each subject, we simulated EEG sensor activity that would be observed for different cortical currents. Using standard anatomical parcellations of cortex [6], 156 distinct anatomical labels (encompassing the cortical surface) were morphed from an average brain to the brain of each subject using surface-based spherical morphing [6]. For each subject, we constructed 156 cortical current sources (one for each label) represented as binary-valued $\sim 7000 \times 1$ vectors $\mathbf{j}$ (one for sources inside the label, zero otherwise). Prior to calculating the resulting observed EEG signals $\mathbf{x}$, we added a white Gaussian noise vector $\mathcal{N}$. By choosing the standard deviation of $\mathcal{N}$, this simulation allowed us to control the SNR across the resulting observed EEG sensors $\mathbf{x}$.

2.3 Inter-Subject Anatomical Transformation

To translate EEG measurements between subjects, we first calculated an inverse solution $\mathbf{M}$ for each subject, which estimates the equivalent brain activity $\hat{\mathbf{j}}$ that was the source of the observed EEG activity $\mathbf{x}$ (as $\hat{\mathbf{j}} = \mathbf{Mx}$). $\mathbf{M}$ is a pseudoinverse of the gain matrix $\mathbf{G}$, and determining it from the forward solution $\mathbf{G}$ is an ill-posed problem (with fewer EEG sensors than brain sources). Although there are several possible solutions to this inverse problem, we used the minimum-norm solution, which preserves linearity and assumes currents in the brain are small and distributed [7]. We then calculated an inter-subject transformation matrix $\mathbf{T_{k\to i}}$ that maps brain activity from Subject $\mathbf{k}$ to Subject $\mathbf{i}$ using a standard surface-based morphing procedure that preserves sulcal-gyral alignment [6]. EEG sensor activity can then be remapped as $\mathbf{\hat{x}_{k\to i}} = \mathbf{G_i T_{k\to i} M_k x_k}$.

2.4 Classification

To test classification, we simulated a signal/no-signal task for each brain region. For each subject, using N_T trials with signal (and N_T without), we tested a standard classifier by partitioning trials into one set of $2N_T - 2$ for training, and two trials (signal and no-signal) for testing, and permuted over partitions. Due to variability from training classifiers using few trials, we averaged 20 iterations of this procedure to better estimate the performance for each subject. For speed and

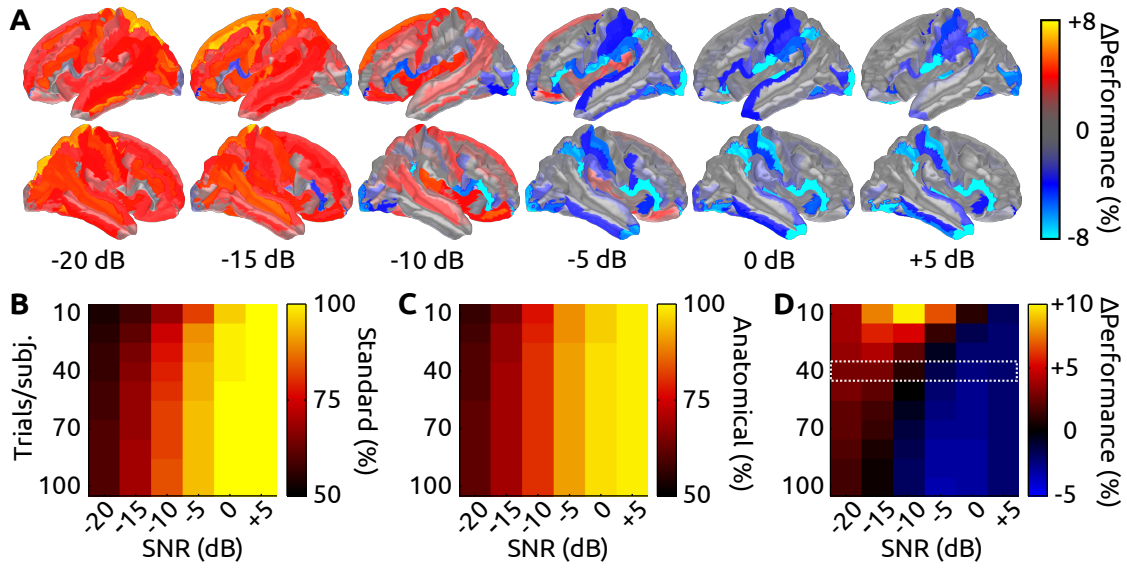


Figure 2: Anatomical classifiers outperform standard classifiers under low-SNR, low-trial conditions. *A*: The performance difference between the anatomical and standard methods for 156 tested regions are mapped onto their corresponding locations on the left (top) and right (bottom) hemispheres (at 40 trials per subject used in training; positive, hot colors represent anatomical performing better than the standard classifier). *B, C, D:* Raw performance for the (*B*) standard and (*C*) anatomical classifiers and (*D*) their difference averaged across all brain regions tested as a function of the number of trials per subject used for training and the SNR of the signal.

simplicity, we used regularized linear discriminant analysis [8] (LDA; $\lambda = 0.05$, optimized for the standard classifier), although support vector machine tests yielded comparable results.

To construct the anatomical classifier for each Subject $\mathbf{i}$, we trained a classifier (LDA, $\lambda = 0.05$) using EEG data only from the other 8 subjects (Figure 1). Here, we transformed the $2N_T$ EEG signals from each Subject $\mathbf{k}$ to Subject $\mathbf{i}$'s EEG sensors using the aforementioned transformation $\hat{\mathbf{x}}_{\mathbf{k}\to\mathbf{i}} = \mathbf{G_i T_{k\to i} M_k x_k}$. Classification performance was obtained by testing the $2N_T$ trials of EEG signals $\mathbf{x_i}$ from Subject $\mathbf{i}$ using a classifier built using the $2N_T$ trials of $\hat{\mathbf{x}}_{\mathbf{k}\to\mathbf{i}}$ transformed from each of the other 8 subjects. This procedure was repeated over 20 iterations to determine the anatomical BCI performance as a function of anatomical location, trial count, and SNR.

3 Results

We compared the performance of the standard BCI classifier and the anatomical BCI classifier across 156 different brain regions (Figure 2 *A*) and averaged across brain regions (Figure 2 *B–C*). We found that, while the standard classifier performed well in many cases, the anatomical BCI classifier outperformed the standard BCI classifier under low-SNR, low-trial-count conditions. For example, when the number of with-signal trials per subject was 20 at -10 dB SNR, the anatomical classifier (77.6 %) outperformed the standard classifier (72.9 %) by 4.7 % (averaged across anatomical regions and subjects). Using these methods, we did not find significant differences in overall brain region, hemisphere, or sulcul-gyral folding (not shown).

4 Discussion

This study provides a potential first step toward the incorporation of anatomical information in BCI designs. We found through our simulations that under low-SNR, low-trial-count conditions,

the anatomical BCI classifier outperformed the standard classifier, suggesting that anatomical information could be leveraged to improve BCI designs. This leaves open several important questions. The observed improvement is likely due to denoising by pooling data across subjects. In this study, the underlying activity pattern across subjects was the same by construction, and it remains to be seen the extent to which activity colocation can be leveraged in realistic BCI settings. However, the success of cross-subject neuroimaging studies suggests it should be possible to utilize common activation patterns.

While the performance of standard BCI designs was superior under high-SNR and high-trial-count conditions, the anatomical method provides several advantages. First, by utilizing an existing pool of data from N subjects, this method could provide for the immediate (non-individually trained) BCI deployment for Subject $N + 1$ given the structural and co-registration information for that subject. Second, although we did not explore the possibility here, it might be possible to achieve even higher performance by building a classifier in the brain (source) space after doing inter-subject brain activity transformation $\hat{\mathbf{j}}_{\mathbf{i}\rightarrow\mathbf{k}} = \mathbf{T}_{\mathbf{i}\rightarrow\mathbf{k}}\mathbf{j}_{\mathbf{i}}$ (as opposed to projecting this transformed brain activity to the EEG sensors) - it is possible that neuroscience insights into functional mapping could be incorporated as priors for future BCI classification designs in the brain space.

5 Conclusion

This study suggests that anatomical information could be leveraged to improve BCI designs, and future work is needed to test this technique using experimental data. In principle, such anatomically based BCI designs could provide the potential to build classifiers that incorporate neuroscience information as priors, and, given subject anatomical information and EEG co-registration, this type of across-subject BCI generalization may allow deployment of out-of-the-box BCI designs.

References

[1] S. Haufe, R. Tomioka, T. Dickhaus, C. Sannelli, B. Blankertz, G. Nolte, and K. R. Müller. Large-scale EEG/MEG source localization with spatial flexibility. *NeuroImage*, 54(2):851–859, January 2011.

[2] E. Larson and A. K. C. Lee. Cortical dynamics during endogenous redirection of auditory spatial attention. In *First International Conference on Cognitive Hearing Science for Communication*, Linkoping, Sweden, June 2011.

[3] B. Fischl, D. H. Salat, A. J. W. van der Kouwe, N. Makris, F. Ségonne, B. T. Quinn, and A. M. Dale. Sequence-independent segmentation of magnetic resonance images. *Neuroimage*, 23 Suppl 1:S69–84, December 2004.

[4] A. M. Dale, B. Fischl, and M. I. Sereno. Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage*, 9(2):179–94, January 1999.

[5] M. S. Hämäläinen and J. Sarvas. Realistic conductivity geometry model of the human head for interpretation of neuromagnetic data. *IEEE Transactions in Biomedical Engineering*, 36:165–171, 1989.

[6] B. Fischl, M. I. Sereno, R. B. H. Tootell, and A. M. Dale. High-resolution intersubject averaging and a coordinate system for the cortical surface. *Human Brain Mapping*, 8(4):272–84, December 1999.

[7] M. S. Hämäläinen and R. J. Ilmoniemi. Interpreting magnetic fields of the brain: minimum norm estimates. *Medical & Biological Engineering & Computing*, 32(1):35–42, January 1994. PMID: 8182960.

[8] J. H. Friedman. Regularized discriminant analysis. *Journal of the American Statistical Association*, 84(405):165–175, 1989.

Movement Trajectory Estimation in Three-Dimensions from Magnetoencephalographic Signals

H. G. Yeom[1,2], J. S. Kim[1,3], C. K. Chung[1,2,3]

[1]MEG center, Seoul National University Hospital, Seoul, Korea
[2]Interdisciplinary Program in Neuroscience, Seoul National University, Seoul, Korea
[3]Departments of Neurosurgery, Seoul National University Hospital, Seoul, Korea

chungc@snu.ac.kr

Abstract

Brain-Computer Interfaces (BCIs) are very useful technologies and it has attracted much attention for recent decades. However, previous BCI studies have been focused on indirect methods. In this paper, we show that movement trajectory can be estimated from non-invasively recorded brain signals, magnetoencephalography (MEG), with high correlation coefficient between real movement and estimated signals.

1 Introduction

Brain-Computer Interfaces (BCIs) are very useful technologies which help disabled people to express their thought or to control robot arm. There have been numerous BCI studies for decades. Several features of brain activity can be utilized for BCI systems such as slow cortical potentials (SCP), sensorimotor rhythms (SMR), P300, steady-state visually evoked potentials (SSVEP). However, SCP is too slow to be used as interface inputs and P300 and SSVEP require display equipments. Details are described in reference [1].

Thus to control machine with brain activity, SMRs are generally used. However, previous BCI methods have been focused on indirect methods. It means that subject should imagine movements which different from real application. For example, imaginations of hand movements are required to shift cursor [2]. Fortunately, recent studies show the possibility controlling robot arm according to movement trajectory. Research team at University of Pittsburgh announced their success to control robot arm in 3-dimensions with monkeys by invasive methods [3]. After this research, several studies tried to reconstruct movement trajectory non-invasively [4, 5]. However, research on reconstruction of trajectory is still in the early stage.

In this paper, we show our results on estimation of 3-dimensional movements from magnetoencephalography (MEG) signals for human with high correlation between real movement and estimated signals.

2 Methods

2.1 Measurements

Six right-handed subjects (age: 21–38, 3 males and 3 females) were instructed to move his index fingertip according to stereoscopic image. To show 3D image, we used anaglyph technique. An image with a ball on the center was presented during 4 sec. After the time, another ball and stick image was displayed on the corner, one of left-up-front, right-up-front, left-down-front, right-down-front, to instruct directions and this sequence was repeated. 1 session of the experiment consisted of 30 trials for each direction (total 120 trials) and 2 sessoins for each subject were measured. The

sequence of the experiment is illustrated in Figure 1. During the experiments, the Brain signals were measured with a 306 channel whole-head MEG system (VectorView^{TM}, Elekta Neuromag Oy, Helsinki, Finland) which arranged in triplets of two planar gradiometers and one magnetometer. The sampling frequency was 600.615 Hz and to increase signal to noise ratio (SNR), SSS filter (Maxwell filter) was applied [6]. To record movement trajectory, 3-axis accelerometer (KXM52, Kionix) was used and the sensor signals were simultaneously recorded with MEG signals.

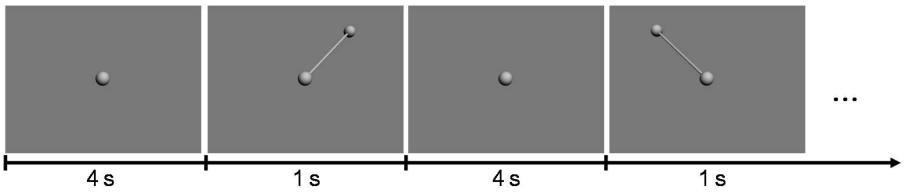


Figure 1: Sequence of different direction movement tasks.

2.2 Analysis

204 channel planar gradiometer signals were analyzed from 1 s before to 2 s after the cue onset and baseline was from 1 s to 0 s before the cue onset. First, we examined time-frequency representation to determine frequency band which related to our experiments. Figure 2 illustrates the time-frequency representation of whole planar gradiometer signals for session 1 of subject 1. The figure displays results from 1 s before to 2 s after cue onset and from 0 Hz to 50 Hz at a frequency resolution of 1 Hz. The figure shows that event-related synchronization (ERS) are occurred in 0–8 Hz and event-related desynchronization (ERD) are shown in 9–14, 22–27 Hz.

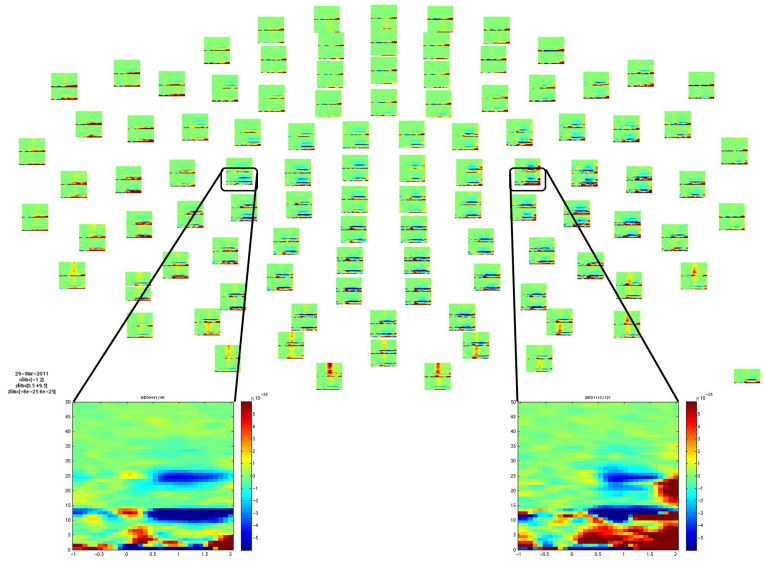


Figure 2: Time-frequency representation for session 1 of subject 1 on planar gradiometer.

2.3 Estimation of Movements

Purpose of this study is to estimate movement trajectory from MEG signals. Thus, MEG signals will be used as input signals and movement trajectory from accelerometer will be desired output. Movement trajectory was calculated from accelerometer signals by integrating the signals twice and MEG signals were filtered with 0.5–4, 4–8, 9–14, 22–27 Hz frequency band. The reason that we

separated 0.5–8 Hz into 0.5–4 and 4–8 Hz is that 0.5–4 Hz is delta wave and 4–8 Hz is theta wave. To prevent over fitting, the filtered signals were downsampled with 20 ms interval like reference [5]. After the downsampling, 5-fold cross validation were employed and then positions of x, y, z were estimated by following equations:

$$\mathrm{x(t)} = \sum_{i=1}^{n}\sum_{j=0}^{10} \mathrm{W}_{ij}^{x} \times \mathrm{J}_i(\mathrm{t}-\mathrm{j}) + \mathrm{W}_0^{x} \tag{1}$$

$$\mathrm{y(t)} = \sum_{i=1}^{n}\sum_{j=0}^{10} \mathrm{W}_{ij}^{y} \times \mathrm{J}_i(\mathrm{t}-\mathrm{j}) + \mathrm{W}_0^{y} \tag{2}$$

$$\mathrm{z(t)} = \sum_{i=1}^{n}\sum_{j=0}^{10} \mathrm{W}_{ij}^{z} \times \mathrm{J}_i(\mathrm{t}-\mathrm{j}) + \mathrm{W}_0^{z} \tag{3}$$

where x(t), y(t), z(t) are estimated fingertip position at time t. W_{ij} is a weight value matrix obtained from regression methods and J_i is a MEG signal of channel i. n is channel numbers and j is time lag (20 ms interval) and W_0 is a constant. W_{ij} and W_0 were obtained by multiple linear regression. Finally, mean correlation coefficient (r) was computed across folds to assess decoding performance.

3 Results

The best results were obtained when the band-pass frequencies were 0.5–4 Hz. Other cases in the frequency band 4–8 or 9–14 or 22–27 had very low correlation. This result corresponds to previous studies [7]. Figure 3 shows estimated movement trajectory averaged for all trials during each session for each subject. Different colors express different direction tasks, blue for left-up-front, red for right-up-front, green for left-down-front, black for right-down-front. Averaged correlation coefficient for all sessions of each subject are illustrated on Figure 4 Correlation coefficient of all trials and subjects were 0.612 for x-direction, 0.6 for y-direction, 0.517 for z-direction.

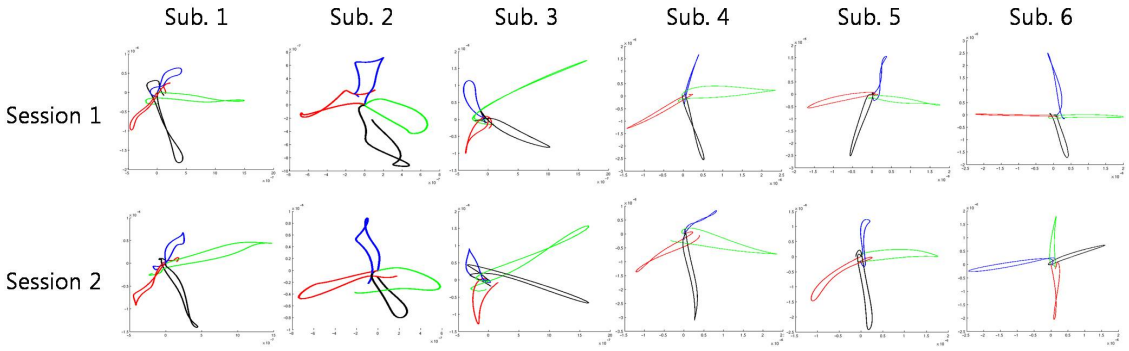


Figure 3: Estimated movement trajectory averaged for each subject and each session.

4 Discussion

We estimated movement trajectory from different frequency band and the results were best in frequency 0.5–4 Hz corresponding to previous studies. Furthermore, recent study suggests that low frequency brain signals relates to movement [8]. We used different testing data set from training data for reconstruction of movement trajectory as discribed in Section 2.3. Also, we showed reconstructed movement trajectory for 6 subjects. Therefore, the results are reliable.

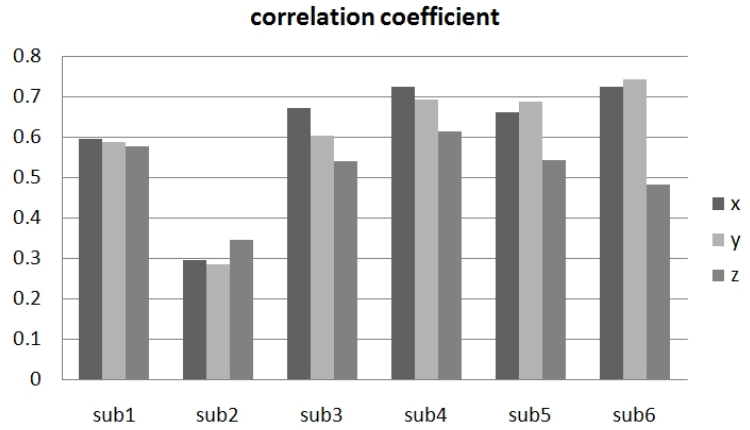


Figure 4: Averaged correlation coefficient for all sessions of each subject.

5 Conclusion

There were few previous studies to reconstruct movement from non-invasively recorded brain signals. In this paper our results have shown that not only movement trajectory in 3 dimension can be estimated by non-invasive methods but also correlation values between estimated signals and real movement signals were higher than existing invasive studies.

References

[1] G. Dornhege, J. R. Millán, T. Hinterberger, D. J. McFarland, and K. R. Müller. *Toward Brain-Computer Interfacing*. MIT press, 2007.

[2] 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 Transactions on Biomedical Engineering*, 55:2452–2462, 2008.

[3] M. Velliste, S. Perel, M. C. Spalding, A. S. Whitford, and A. B. Schwartz. Cortical control of a prosthetic arm for self-feeding. *Nature*, 453:1098–1101, 2008.

[4] T. J. Bradberry, R. J. Gentili, and J. L. Contreras-Vidal. Reconstructing three-dimensional hand movements from noninvasive electroencephalographic signals. *The Journal of Neuroscience*, 30:3432–3437, 2010.

[5] A. Toda, H. Imamizu, M. Kawato, and M. Sato. Reconstruction of two-dimensional movement trajectories from selected magnetoencephalography cortical currents by combined sparse Bayesian methods. *NeuroImage*, 54:892–905, 2011.

[6] S. Taulu, J. Simola, and M. Kajola. Applications of the signal space separation method. *IEEE Transactions on Signal Processing*, 53:3359–3372, 2005.

[7] S. Waldert, T. Pistohl, C. Braun, T. Ball, A. Aertsen, and C. Mehring. A review on directional information in neural signals for brain-machine interfaces. *Journal of Physiology - Paris*, 103:244–254, 2009.

[8] M. Bourguignon, X. De Tiege, M. O. de Beeck, B. Pirotte, P. Van Bogaert, S. Goldman, R. Hari, and V. Jousmaki. Functional motor-cortex mapping using corticokinematic coherence. *Neuroimage*, 15:1475–1479, 2011.

Feature Extraction for Brain-Computer Interface (BCI) Based on the Functional Causality Analysis of Brain Singals

J. H. Lim[1,2], H. J. Hwang[1,2], Y. J. Jung[1,2], C. H. Im[2]

[1]Department of Biomedical Engineering, Yonsei University, Wonju, Korea
[2]Department of Biomedical Engineering, Hanyang University, Seoul, Korea

gobbili1@hanmail.net han-jeong@yonsei.ac.kr microbme@gmail.com ich@hanyang.ac.kr

Abstract

The aim of this paper was verify whether the features derived from functional causality analysis of human brain would be promising feature candidates that can enhance classification accuracy of motor imagery based brain-computer interface (BCI). To this end, we classified left and right hand motor imagery tasks using the features derived from directed transfer function (DTF) method as well as the conventional features derived from power spectral density (PSD) and phase locking value (PLV) analyses. As a result, classification accuracy for feature sets including DTF features was higher than that for the opposite case, thereby demonstrating that the DTF features could be utilized as one of the promising feature candidates capable of enhancing the performance of BCI systems.

1 Introduction

In electroencephalography (EEG)-based BCI studies, one of the widely used mental tasks is motor imagery defined as mental simulation of a kinesthetic movement [1] because the brain activities modulated by motor imagery are readily reproducible and show consistent EEG patterns on the sensorimotor area [2]. In particular, it has been well established that left and right hand motor imagery can modulate mu and beta rhythm activities in the sensorimotor cortex, and thereby resulting in event-related desynchronization (ERD) of mu and beta bands in the contralateral sensorimotor areas [3]. Thanks to the contralateral localization character in mu and beta rhythms, the features derived from power spectral density (PSD) analysis in these bands has been traditionally used as a main feature set that is capable of discriminating different motor imagery tasks [4].

In order to increase classification accuracy of motor imagery task-based BCI systems, some recent studies introduced a new feature based on the functional connectivity between different brain regions, for which phase locking value (PLV) has been adopted [4, 5]. These two previous studies showed consistent results in that the ability of discriminating different motor imagery tasks can be enhanced when PLV features were utilized along with PSD features. These results demonstrated that the features derived from PLV analysis could increase the overall performance of BCI systems, particularly when they were combined with PSD features.

In the meanwhile, there have been some indicators that are capable of measuring directional connectivity or causality, which represents the information flow among different brain areas, and the directed transfer function (DTF) analysis has been one of the most widely used causality measures [6]. Recently, DTF analysis was used to investigate the flow of information processing in brain during real and imagined motor tasks, and similar propagation patterns were found in the both tasks [7]. This result was in line with those of PSD [2] and PLV [8] analyses, in that brain responses modulated by real motor tasks are similar to those modulated by imagined motor tasks. These findings indirectly indicate that the features derived from DTF analysis would also be one of the useful feature candidates to enhance the classification accuracy of motor imagery tasks.

In the present study, our hypothesis was that motor imagery tasks might be classified more accurately when DTF features were utilized along with PSD and PLV features. To verity our hypothesis, we used the *BCI Competition IV Data set 2a*, including EEG data acquired during left and right hand motor imagery [9]. The seven feature sets, consisting PSD, PLV, DTF, and the combination of these features, were extracted from the EEG data in 6–30 Hz frequency band. The sequential feature selection (SFS) and support vector machine (SVM) algorithms were used for the feature selection and classification, respectively. The classification accuracy for the feature sets derived from DTF analysis was compared with those derived from PSD and PLV analyses to confirm whether or not DTF features are a promising feature candidate to enhance the performance of BCI systems.

2 Methods

2.1 Data Description

We used the *BCI Competition IV Data set 2a* acquired from nine subjects with 22 channels, while they were performing four different motor imagery tasks, namely the imagination of movement of the left hand, right hand, both feet, and tongue. We reduced the complexity of the problem by changing the original 4-class problem into a new 2-class classification problem. EEG signals recorded during left and right hand motor imagery tasks were used for this study and the number of trials is 144 each. For the data analysis, the 4.0 s time segment of each trial was extracted from the beginning of a cue signal to notify the starting point of performing motor imagery tasks.

2.2 Feature Extraction

The raw EEG signals were filtered at a bandpass of 6–30 Hz including mu and beta bands because these frequency bands are associated with motor imagery tasks [3], and then DTF features were extracted along with PSD and PLV features from each 4 s epoch. To make the same number of features for PSD and PLV, we used slightly different electrode configurations and frequency spans in extracting these features. However, all selected electrode positions in the different electrode configurations were around sensorimotor cortex reflecting motor imagery tasks.

2.2.1 PSD Feature Set

PSD features were obtained from 14 channels (Fz, FCz, FC3, FC4, C1, C2, C5, C6, CPz, CP3, CP4, P1, P2, POz). For each channel, the PSD features were extracted by using fast Fourier transformation (FFT) in 16 frequency bands, namely 12 nonoverlapping narrow bands (width 2 Hz) and 4 broad bands (6–13, 13–20, 20–30, and 6–30 Hz). As the numbers of channels and frequency bands were 14 and 16 respectively, 224 PSD features were extracted for each trial. The PSD features were not logarithmized.

2.2.2 PLV Feature Set

PLV features were extracted from 8 electrode locations (FCZ, FC3, FC4, C1, C2, CPz, CP3, CP4), and thereby constructing 28 different electrode pairs. For all pairs of electrodes, PLV were calculated in 8 equally divided frequency bands in which the span of frequency band was 3 Hz [4]. Eventually, 224 PLV features were extracted per each trial. In the case of extracting PLV features, we did not utilize broad frequency bands so as to make the number of PLV features identical to PSD features.

2.2.3 DTF Feature Set

The same electrode locations and frequency bands were utilized in extracting DTF features as PLV feature extraction. However, the number of DTF features was twice than that of PLV features

because two features were extracted from a pair of electrodes [7]. Consequently, 448 DTF features were extracted for each trial.

2.2.4 The Combined Feature Sets

The additional four feature sets were created by the combination of PSD, PLV, and DTF feature sets, which were *PSD&PLV*, *PSD&DTF*, *PLV&DTF*, and *PSD&PLV&DTF*. We finally obtained seven different feature sets labeled as *PSD,PLV, DTF*, *PSD&PLV*, *PSD&DTF*, *PLV&DTF*, and *PSD&PLV&DTF*.

2.3 Feature Selection

To choose the best feature subsets in terms of classification accuracy, the SFS algorithm was applied to each constructed feature set [10]. The SFS algorithm was set to find fourteen optimal feature subsets for each feature set according to the number of features ranging from 2 to 15. The reason why we selected many best feature subsets was to avoid biased results by averaging classification accuracies for the best feature subsets.

2.4 Classification

For the classification, we used the SVM algorithm that has been widely used in BCI research [11]. Also, 10×10 fold cross validation method was introduced to avoid biased results in estimating classification accuracy for each best feature subset of seven feature sets. The classification accuracy for each feature set was finally obtained by averaging classification accuracies calculated from fourteen optimal feature subsets.

3 Result

Figure 1 shows the averaged classification accuracies for the feature sets including DTF features (*DTF*, *PSD&DTF*, *PLV&DTF*, and *PSD&PLV&DTF*) and those not including DTF features *PSD,PLV*, *PSD&PLV*). Classification accuracy for the feature sets including DTF features was higher than those for non-DTF feature sets in most cases. The percentage of the DTF features among the selected features was about 32 %. The enhancement of the classification accuracy was statistically significant (paired t-test; $p = 0.0022$).

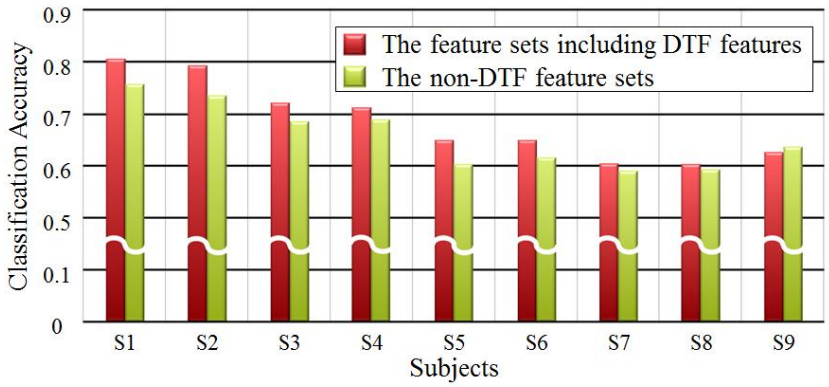


Figure 1: Comparison of averaged classification accuracies for the feature sets with and without DTF features.

From these results, we could confirm that DTF features would be a superior feature candidate to increase classification accuracy of motor imagery tasks when using PSD and PLV features together.

4 Conclusions

To the best of our knowledge, DTF analysis has never been used for extracting BCI features for discriminating motor imagery tasks. In the present study, we explored whether the features derived from DTF analysis could be used as a useful feature candidate to enhance classification accuracy. Our results have shown that the classification accuracy for the feature sets including DTF features were higher than that not including DTF features, demonstrating that DTF would provide promising feature candidates that can increase classification accuracy of motor-imagery-based BCI.

5 Acknowledgments

This research was supported by Public welfare&Safety research program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology (No. 2010-0020807).

References

[1] J. Decety and D. H. Ingvar. Brain structures participating in mental simulation of motor behavior: A neuropsychological interpretation. *Acta Psychologica*, 73(1):13 – 34, 1990.

[2] G. Pfurtscheller and C. Neuper. Motor imagery activates primary sensorimotor area in humans. *Neurosci.Lett.*, 239:65 – 68, 1997.

[3] M. Lotze, P. Montoya, M. Erb, E. Hülsmann, H. Flor, U. Klose, N. Birbaumer, and W. Grodd. Activation of cortical and cerebellar motor areas during executed and imagined hand movements: An fMRI study. *J. Cognitive Neuroscience*, 11:491–501, September 1999.

[4] C. Brunner, R. Scherer, B. Graimann, G. Supp, and G. Pfurtscheller. Online control of a brain computer interface using phase synchronization. *IEEE Trans. Biomed. Eng.*, 53:2501 –2506, 2006.

[5] E. Gysels and P. Celka. Phase synchronization for the recognition of mental tasks in a brain-computer interface. *IEEE Trans. Neural. Syst. Rehabil. Eng.*, 12:406 –415, 2004.

[6] P. J. Franaszczuk, G. K. Bergey, and M. J. Kaminski. Analysis of mesial temporal seizure onset and propagation using the directed transfer function method. *Electroen. Clin. Neuro.*, 91:413–427, 1994.

[7] J. Ginter, K. J. Blinowska, M. Kaminski, and R. Kus. Propagation of brain electrical activity during real and imagined motor task by directed transfer function. In *proc. 2nd IEEE Int. EMBS Conf. Neural Eng.*, 2005.

[8] M. L. Stavrinou, L. Moraru, L. Cimponeriu, S. D. Penna, and A. Bezerianos. Evaluation of cortical connectivity during real and imagined rhythmic finger tapping. *Brain Topogr*, 19:137–145, 2007.

[9] C. Brunner, M. Naeem, R. Leeb, B. Graimann, and G. Pfurtscheller. Spatial filtering and selection of optimized components in four class motor imagery EEG data using independent components analysis. *Pattern Recogn. Lett.*, 28:957–964, 2007.

[10] P. Pudil, J. Novovicova, and J. Kittler. Floating search methods in feature selection. *Pattern Recogn. Lett.*, 15:1119 – 1125, 1994.

[11] A. F. Cabrera, D. Farina, and K. Dremstrup. Comparison of feature selection and classification methods for a brain computer interface driven by non-motor imagery. *Med. Biol. Eng. Comput.*, 48:123–132, 2010.

A New Dual-Frequency Stimulation Method for SSVEP-Based Brain-Computer Interface

H. J. Hwnag[1,2], J. H. Lim[1,2], C. H. Im[2]

[1]Department of Biomedical Engineering, Yonsei University, Wonju, Korea
[2]Department of Biomedical Engineering, Hanyang University, Seoul, Korea

han-jeong@yonsei.ac.kr gobbili1@hanmail.net ich@hanyang.ac.kr

Abstract

The aim of this study was to generate more visual stimuli with limited numbers of stimulating frequencies in steady state visual evoked potential (SSVEP)-based brain-computer interface (BCI) system. To this end, a new dual-frequency stimulation method was devised by modifying the conventional checkerboard pattern stimuli and consequently ten different visual stimuli could be generated from the combination of only four different frequencies. Six participants took part in the preliminary experimental studies, which verify the feasibility of the proposed stimulation method. EEG signals were acquired from only one electrode mounted on Oz position while the participants were focusing their eyes on the ten visual stimuli. The power spectrum analysis was applied to the recorded EEG data to observe SSVEP responses evoked by each visual stimulus. From the analysis results, it was confirmed that each visual stimulus could evoke a distinct SSVEP peak, demonstrating that the suggested dual-frequency stimulation method can be an effective way to increase the number of visual stimuli in SSVEP-based BCI systems.

1 Introduction

Steady-state visual evoked potential (SSVEP) is a brain response induced by the repetitive presentation of a visual stimulus flickering or reversing at a certain frequency. Recently, SSVEP has received increased attention in brain-computer interface (BCI) field because it shows high information transfer rate (ITR), it requires a fewer number of electrodes, and training period is short as compared to those based on the other brain signals [1, 2], such as mu rhythm, slow cortical potential (SCP), and auditory steady state response (ASSR).

In previous SSVEP-based BCI studies, a fairly broad frequency band ranging from 1 to 60 Hz have been generally used for eliciting SSVEP responses [3–8]. The available frequencies for stimulation, however, are considerably restricted by several reasons, especially when computer monitor is used to present the visual stimuli. First, all available stimulating frequencies cannot always evoke high SSVEP responses because it depends on a person and the various properties of visual stimuli, such as color, size, and contrast [2]. Second, it has been typically avoided to use two frequencies F1 and F2 simultaneously when F1 is a multiple of F2 or vice versa because of its harmonic characteristic [3]. Third, the frequencies in alpha band (8–13 Hz) should be carefully selected due to a considerable amount of false positives along with harmonic responses [2]. Furthermore, in the case of using a computer monitor as a rendering device, the stimulating frequencies should be the subharmonics of the monitor refresh ratio to get accurate SSVEP responses [2]. Therefore, one of the challenging issues in SSVEP-based BCI research has been to make as many available frequencies as possible, which can elicit distinct brain responses.

In the present study, we propose a new dual-frequency stimulation method to increase the number of visual stimuli by using a limited number of stimulating frequencies. Six subjects

participated in our experiment in which they were asked to gaze a checkerboard pattern modulated with two different stimulating frequencies.

2 Methods

2.1 Subjects

Six subjects (five males and one female) aged between 21 and 27 years took part in the present study, all of which had normal or corrected-to-normal vision. None of them had a previous history of neurological, psychiatric or other severe disease known to adversely affect EEG recording. A fully detailed summary of the experimental procedures and protocols were explained to each participant before the experiment. All subjects signed consent forms and received proper reimbursement for their participation. The study was reviewed and approved by the Institutional Review Board (IRB) committee of Yonsei University, South Korea.

2.2 Visual Stimulation

Figure 1 illustrates the visual stimulus used in this study, which was generated using Cogent 2000 and Cogent Graphics (http://www.vislab.ucl.ac.uk/cogent.php). In the case of using a traditional checkerboard pattern to induce a SSVEP response, two patterns colored in black and white are reversed with a constant frequency at the same time. For the proposed dual-frequency stimulation method, the checkerboard pattern was made to be individually alternated at two different frequencies for each white and black pattern as described in Figure 1. The checkerboard pattern was placed on the center of a gray (RGB: 132, 132, 132) background and was presented to the subjects with a visual angle of approximately $3.58° \times 3.58°$. A 19-inch LCD monitor configured with a resolution of 640×480 pixels was used for the presentation of the visual stimuli and the refresh rate of the monitor was set at 60 Hz.

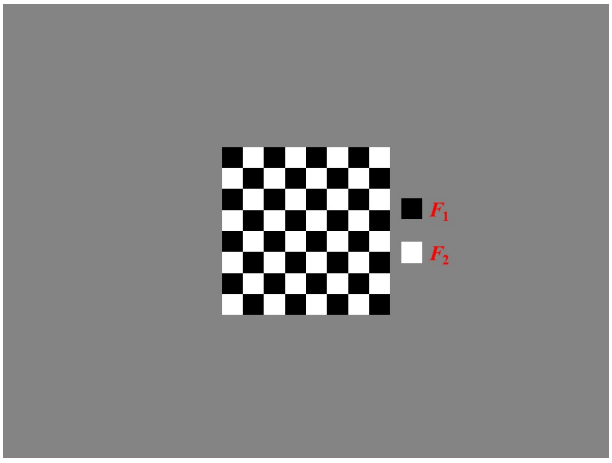


Figure 1: The visual stimulus used in this study. Each white and black pattern was reversed with different frequencies.

Ten visual stimuli were generated by using only four slightly different frequencies, i.e. 3, 3.33, 3.75, and 4.285 Hz. Four visual stimuli were obtained by using the same frequencies for each white and black pattern, i.e. 3–3 Hz, 3.33–3.33 Hz, 3.75–3.75 Hz, and 4.285–4.285 Hz. These four visual stimuli were actually identical to traditional checkerboard stimuli inducing SSVEP responses at 6, 6.6, 7.5, and 8.57 Hz, respectively. The other six visual stimuli were generated by using the combination of the four fundamental frequencies, i.e. 3–3.33 Hz, 3–3.75 Hz, 3–4.285 Hz, 3.33–3.75 Hz, 3.33–4.285 Hz, and 3.75–4.285 Hz.

2.3 EEG Data Collection

Each of the ten visual stimuli was presented to the subjects during 30 s with a short break. While the subjects were focusing on each visual stimulus, EEG signals were recorded with one electrode (Oz) attached on the subjects' scalp according to international 10-20 system using a multi-channel EEG acquisition system (WEEG-32, Laxtha Inc., Daejeon, Korea) in a dimly lit, soundproof room. The EEG channel was referenced to an electrode behind the right mastoid and a ground electrode was placed behind the left mastoid. The sampling rate was set at 512 Hz in the experiment.

2.4 Power Spectrum Analysis

To observe SSVEP responses, a recorded 30 s EEG signal for each stimulus was divided into segments of 4 seconds with 50 % overlapping between the segments which resulted in 14 time segments. The power spectrum of each 4-second epoch was calculated by fast Fourier transformation (FFT) and then the average of 14 spectra was computed to smooth out random peaks in the frequency domain.

3 Result

Figure 2 shows SSVEP peaks elicited by the dual-frequency visual stimuli for subject WH. It was clearly observed in Figure 2 that the SSVEP peaks were evoked at different frequencies for each stimulus, and thereby demonstrating that only four different frequencies can generate ten discriminable selections. The SSVEP responses for the other five subjects exhibited the same tendency with those for subject WH, and the results for all participants will be presented in the upcoming conference.

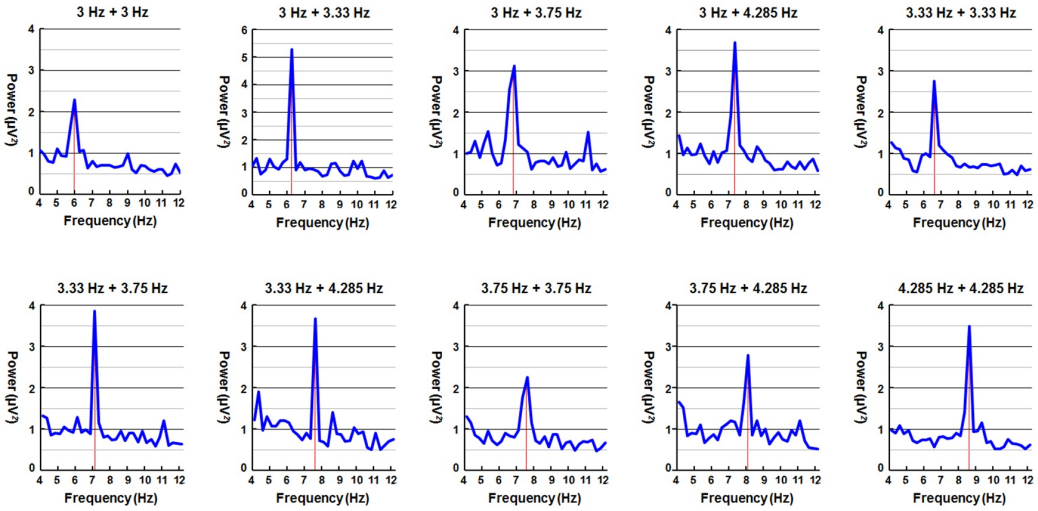


Figure 2: SSVEP responses derived from the proposed dual-frequency stimulation method for subject WH. SSVEP peaks were observed at the sum frequencies of two fundamental stimulating frequencies for every stimulus. Note that the range of y-axis for the result of 3 Hz + 3.33 Hz is from 0 to 6 while the others ranged from 0 to 4.

4 Discussion

As seen in the Figure 2, all SSVEP peaks were induced at the sum frequencies of two stimulating frequencies for each stimulus. These results have a thread of connection with those of

the previous studies based on dual-frequency stimulation methods [4, 5]. However, it is worthy noticing that only one distinct SSVEP response for each stimulus was observed using the suggested dual-frequency stimulation method; while SSVEP responses at two main frequencies used for stimulation were shown along with those at their harmonics in the previous studies [4–7]. In the present study, one electrode site (Oz) was used for observing SSVEP responses. Since it has been demonstrated that the performance of SSVEP-based BCI systems can be enhanced with additional electrode positions (i.e. O1 and O2) [8], we will apply this fact in realizing an online BCI system based on the suggested dual-frequency stimulation method.

5 Conclusions

In the present study, a new dual-frequency stimulation method was proposed in order to increase the number of selections with a limited number of stimulating frequencies. Only four different frequencies were used to generate ten visual stimuli which comprised of white and black patterns reversing at different frequencies. From the preliminary experimental results, it was shown that each checkerboard pattern could produce a discriminable spectral peak, and thereby demonstrating that the suggested stimulation method could be one of the useful solutions to make the best use of available frequencies in SSVEP-based BCI systems.

6 Acknowledgments

This research was supported by Public welfare&Safety research program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology (No. 2010-0020807).

References

[1] F. B. Vialatte, M. Maurice, J. Dauwels, and A. Cichocki. Steady-state visually evoked potentials: focus on essential paradigms and future perspectives. *Prog. Neurobiol.*, 90:418 – 438, 2010.

[2] D. Zhu, J. Bieger, G. G. Molina, and R. M. Aarts. A survey of stimulation methods used in SSVEP-based BCIs. *Comput. Intell. Neurosci.*, 2010:1:1–1:12, 2010.

[3] H. Bakardjian, T. Tanaka, and A. Cichocki. Optimization of SSVEP brain responses with application to eight-command brain-computer interface. *Neurosci. Lett.*, 469:34 – 38, 2010.

[4] M. Cheng, X. Gao, S. Gao, and D. Xu. Multiple color stimulus induced steady state visual evoked potentials. In *Proceedings of the 23rd Annual EMBS International Conference*, 2001.

[5] T. M. S. Mukesh, V. Jaganathan, and M. R. Reddy. A novel multiple frequency stimulation method for steady state VEP based brain-computer interfaces. *Physiol. Meas.*, 27:61–71, 2006.

[6] K. K. Shyu, P. L. Lee, Y. J. Liu, and J. J. Sie. Dual-frequency steady-state visual evoked potential for brain-computer interface. *Neurosci. Lett.*, 483:28–31, 2010.

[7] B. Z. Allison, B. Graimann, T. Leuth, and A. Graser. An SSVEP brain-computer interface (BCI) with simultaneous attention to two targets. In *the 37th annual meeting of the Society for Neuroscience*, 2007.

[8] C. Brunner, B. Z. Allison, C. Altstatter, and C. Neuper. A comparison of three brain-computer interfaces based on event-related dysynchronization, steady state visual evoked potentials, or a hybrid approach using both signals. *J. Neural Eng.*, 8(2):025010, 2011.

Automatic Frequency Band Selection for BCIs with ERDS Difference Maps

M. Billinger[1], V. Kaiser[1], C. Neuper[1,2], C. Brunner[1]

[1]Institute for Knowledge Discovery, BCI Lab, Graz University of Technology, Austria
[2]Department of Psychology, University of Graz, Austria

martin.billinger@tugraz.at

Abstract

We present a new fully automated algorithm for frequency band selection in band power based brain-computer interfaces (BCIs). The algorithm performs a single pass and only requires information on the training data in the form of time-frequency ERDS maps, which visualize event-related desynchronization (ERD) and event-related synchronization (ERS). Frequency bands are selected by segmentation of ERDS difference images, similar to the way a human expert would manually select bands by inspecting these maps. We could show that bands selected by this approach perform almost as well as manually selected bands in an offline BCI experiment with data recorded from 18 users.

1 Introduction

Many brain-computer interfaces (BCIs) rely on a variety of different techniques to extract information from brain signals such as the electroencephalogram (EEG), which is often used in non-invasive BCIs. The instantaneous power of selected frequency bands (band power) has been successfully used to discriminate between different tasks in motor imagery (MI) based BCIs [1]. Band power can be estimated using filter banks [1] or autoregressive estimates of the spectrum [2].

Usually, frequency bands that contain discriminative information relevant to the BCI paradigm must be selected. This can be done using knowledge of the underlying physiological processes, or after analyzing a BCI user's specific task-related brain patterns. Selecting bands corresponding to fixed μ or β frequency bands may not be optimal for every user due to individually varying frequency components [1]. Automatic feature selection algorithms such as distinction sensitive learning vector quantization (DSLVQ) [1] or sequential floating forward selection (SFFS) [3] optimize bands for individual persons but do not have any knowledge about the nature of the features, which may lead to suboptimal or redundant band selection. Manual band selection by an expert can account for differences between individuals, as well as relevant physiological background.

An algorithm to select a frequency band for the common spatial patterns (CSP) approach is briefly mentioned in [4]. This method selects a single frequency band for all channels, based on a correlation score. We present a new algorithm based on image segmentation methods from the field of computer vision, which mimics an expert inspecting time-frequency maps of event-related desynchronization (ERD) and event-related synchronization (ERS) to select frequency bands. Multiple bands are selected for each channel separately, so that they discriminate between two different BCI tasks. In Section 2, we provide a detailed description of the algorithm and present possible extensions. In Section 3, we compare the performance of the algorithm to a human expert performing band selection on the same data set.

2 Methods

2.1 Manual Band Selection

ERDS maps show changes in the power spectrum (with respect to a reference interval) related to a recurring event such as MI. Power changes are averaged over many trials, and only if the change is significantly higher or lower than zero, the map is shaded in color. An increase in power is known as ERS (typically represented by blue color), whereas a power decrease is called ERD (represented by red color) [5].

After calculating ERDS maps corresponding to each MI task (class), an experts inspects these maps and selects those frequency bands where the maps differ most in a task-related way. However, the selected bands depend on the expert's subjective interpretation of the ERDS maps.

2.2 Automatic Band Selection

ERDS maps do not only produce visually appealing and informative images. On a lower level, the mean power change and associated confidence interval are obtained for each pixel in the map. Using this information from maps of two different classes, we can construct ERDS difference maps. The shading of these difference maps show where and how much the ERDS maps of the classes differ. Pixels with overlapping confidence intervals do not contain significant class differences and remain blank. One such difference map is created for each EEG channel.

Simply plotting ERDS difference maps would already provide a rough overview of which frequency bands could be selected to cover the most prominent class differences. Still, this would have to be done by a human supervisor. Thus, we attempt to automate this process.

Automation is based on the idea of eliminating differences caused by noise and fitting frequency bands to the remaining patches of differences. This is accomplished by first constructing a significance bitmap, in which 1 and 0 encode significant and non-significant differences at the corresponding point in the ERDS difference map. Groups of adjacent significant pixels are referred to as regions. Regions that are smaller than a threshold area A_{th} are considered as noise and can be removed by area-opening [6]. For each remaining region, a frequency band is created that ranges from this region's lowest to highest frequencies. Finally, overlapping frequency bands are merged. Figure 1 shows a summary of the algorithm. Note that the algorithm treats each channel individually. Thus, different bands may be selected for each channel.

Two parameters have direct impact on band selection: the Type I error probability of pixels in the ERDS map (α) and the area threshold for discarding small regions (A_{th}). Note that these parameters are not independent. A lower α leads to smaller regions, which causes more regions to be discarded with constant A_{th}.

2.3 Evaluation of the Algorithm

We used MI data from 18 persons to compare automatic with manual band selection. The data sets used are BCI Competition IV data sets 2A and 2B [7], each recorded from 9 different persons. From data set 2A, only three bipolar channels (C3, Cz, and C4) and two MI tasks (left versus right hand) were used to match the data available in data set 2B. All data available for each participant was separated into training and testing sets. Trials with artifact markers were removed. On average, the training set from participants of data set A consisted of 129 ± 9 trials, and 205 ± 29 trials from data set B. The testing set contained 131 ± 12 trials from data set A, and 360 ± 38 trials from data set B.

To test manual band selection, ERDS maps from the training data were inspected by an expert, who selected frequency bands for each participant that would allow classification of the MI tasks using band power features and linear discriminant analysis (LDA). The algorithm for automatic band selection was applied to the training data with parameters $\alpha = 0.01$ and $A_{tr} = 2$. If no bands were found, a wide frequency range from 5–30 Hz for every channel was selected.

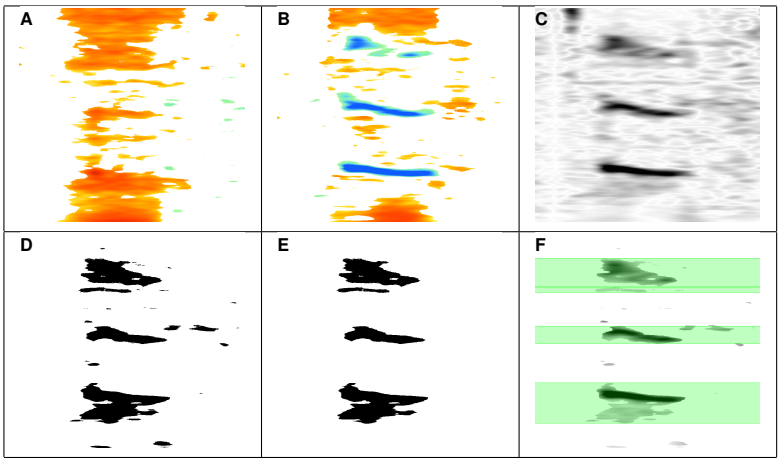


Figure 1: Processing steps of the band selection algorithm. (A) and (B) ERDS maps of each of the two MI classes, (C) ERDS difference map, (D) significance bitmap (E) significance bitmap after area-opening, (F) ERDS difference map with significance information and selected frequency bands. Axes are not labelled to emphasize that the algorithm treats these maps simply as images.

An LDA classifier was trained on the selected bands and subsequently tested with the same bands on the testing set. Classification accuracy was calculated from the continuous classifier output for each trial. The 0.9 quantiles of classification accuracy are reported as robust and comparable measures of classification accuracy for each participant.

2.4 Possible Extensions of the Algorithm

Extension of the band selection algorithm to an arbitrary number of N classes is straightforward by using multiple classifiers and assuming that each classifier has its own set of bands. Consider a three-class classification problem with classes A, B, and C. A pairwise classification scheme requires training three classifiers A–B, A–C, and B–C. Thus, our algorithm can be applied to each combination of classes to select the bands for each classifier. In a one versus the rest classification scheme, classification for class A versus the combined classes B and C is performed. To select bands for the A–BC classifier, our algorithm must be applied to the difference of the ERDS maps from class A and the combined ERDS maps from classes B and C.

If we wanted to select the same frequency bands for multiple channels, the difference maps of individual channels have to be merged into a single map. With this modification applied, and the selection process restricted to select only the single most important frequency band, we expect our approach to be suitable also for CSP-based classification.

Furthermore, our algorithm does not necessarily rely on ERDS difference maps. Any measure in time-frequency space that provides information about class discrimination and significance can be subject to band selection. The most obvious example is using the difference of power maps directly instead of ERDS maps.

3 Results

Automatic band selection failed for 3 out of 18 participants (A2, A6, and B3). For these participants, broadband features were used instead, as described above. Table 1 lists classification accuracies for each participant for both automatic and manual band selection. Average accuracy was $68.1 \pm 13.5\,\%$ with automatic and $70.5 \pm 14.5\,\%$ with manual band selection. According to a paired t-test, the difference between both methods was not significant ($p = 0.198$).

Participant	A1	A2	A3	A4	A5	A6	A7	A8	A9
Automatic	73.8	*57.0*	70.8	64.4	54.8	*59.3*	66.4	91.0	89.2
Manual	83.7	58.5	65.0	58.3	58.5	60.2	54.3	94.8	88.5
Participant	B1	B2	B3	B4	B5	B6	B7	B8	B9
Automatic	54.6	53.2	*53.2*	94.1	85.8	67.3	61.4	64.3	65.8
Manual	63.6	53.2	54.6	95.4	84.5	65.0	78.1	82.7	70.8

Table 1: Classification accuracies for all participants with automatic and manual band selection. The 0.9 quantiles of classification accuracy (in %) are listed in the table. Note that A1–A9 are different persons than B1–B9.

4 Discussion and Conclusion

The main disadvantage of applying feature selection algorithms like SFFS or DSLVQ to the band selection problem is that these algorithms do not have knowledge about the nature of the features. When applied to band selection, a subset of pre-defined frequency bands is selected. To allow these methods enough flexibility and accuracy in band selection, a sufficiently large number of bands has to be provided, which can dramatically increase computational requirements.

The band selection method we present in this paper does not suffer from these disadvantages. Instead of selecting bands from a pre-defined set, band limits are fitted to statistically significant class differences in the time-frequency domain. This approach closely resembles the way a human expert selects frequency bands and yields comparable results. Segmentation of significant regions is very efficient. The computationally limiting factor is the calculation of ERDS maps.

Although automatic band selection performed slightly worse than manual selection, the null hypothesis that both methods perform equal could not be rejected. Thus, we conclude that the difference between the two methods (if there is any) is too small to expose a significant effect in the available data.

Depending on the application, a small loss of classification accuracy may be preferrable over the need of a human expert to interact with the BCI setup process. This may be the case especially with BCIs for home use, where the whole BCI process should run as automated as possible.

References

[1] G. Pfurtscheller and C. Neuper. Motor imagery and direct brain-computer communication. *Proceedings of the IEEE*, 89:1123–1134, 2001.

[2] D. J. McFarland and J. R. Wolpaw. Sensorimotor rhythm-based brain-computer interface (BCI): model order selection for autoregressive spectral analysis. *Journal of Neural Engineering*, 5:155–162, 2008.

[3] P. Pudil, J. Novovičová, and J. Kittler. Floating search methods in feature selection. *Pattern Recognition Letters*, 15:1119–1125, 1994.

[4] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K. R. Müller. Optimizing spatial filters for robust EEG single-trial analysis. *IEEE Signal Processing Magazine*, 25:41–56, 2008.

[5] B. Graimann, J. E. Huggins, S. P. Levine, and G. Pfurtscheller. Visualization of significant ERD/ERS patterns in multichannel EEG and ECoG datas. *Clinical Neurophysiology*, 113:43–47, 2002.

[6] S. T. Acton and N. Ray. Biomedical image analysis: Segmentation. *Synthesis Lectures on Image, Video, and Multimedia Processing*, 4(1):1–108, 2009.

[7] BCI Competition 2008 - Graz data sets 2A and 2B. *http://www.bbci.de/competition/iv/.*

Freeze the BCI Until the User is Ready: a Pilot Study of a BCI Inhibitor

L. George[1], L. Bonnet[1], A. Lécuyer[1]

[1]INRIA, Campus Universitaire de Beaulieu, F-35042 Rennes Cedex, France

{laurent.f.george, laurent.bonnet, anatole.lecuyer}@inria.fr

Abstract

In this paper we introduce the concept of Brain-Computer Interface (BCI) inhibitor, which is meant to standby the BCI until the user is ready, in order to improve the overall performance and usability of the system. BCI inhibitor can be defined as a system that monitors user's state and inhibits BCI interaction until specific requirements (e.g. brain activity pattern, user attention level) are met. In this pilot study, a hybrid BCI is designed and composed of a classic synchronous BCI system based on motor imagery and a BCI inhibitor. The BCI inhibitor initiates the control period of the BCI when requirements in terms of brain activity are reached (i.e. stability in the beta band). Preliminary results with four participants suggest that BCI inhibitor system can improve BCI performance.

1 Introduction

There are several ways to improve performance of Electroencephalography (EEG) based Brain-Computer Interface (BCI) systems: improving classification methods, developing new signal processing algorithms or increasing EEG hardware efficiency are the most common directions taken [1]. One recent approach intends to combine different paradigms into a hybrid system [2]. For example, the Error Related Potential can be detected during a BCI-based interaction to correct the interaction and increase overall performance [3].

However relatively few studies try to improve the performance of BCI systems by focusing on interaction techniques and usage. An approach is to use a brain switch [2, 4] which can be described as a preliminary BCI that allows the user to activate the BCI interaction at will. This system allows reducing false positive in asynchronous BCI context as confirmed in [5].

In this paper we propose an implicit and complementary system which we have named BCI inhibitor. The main objective is to increase performance of a BCI by using a paradigm that will activate the BCI system only if the best conditions are met, with respect to the current state of the user. Thus BCI inhibitor will avoid classifying EEG features when the system detects that the current user's state will undoubtedly lead to an erroneous result. We evaluated our solution through a pilot study involving motor imagery tasks.

2 Concept

BCI inhibitor can be defined as a system that pauses the BCI until specific conditions (e.g. optimum condition in terms of user's state) are met. BCI inhibitor is in relation with the recent idea of a brain switch, but in the case of BCI inhibitor it is not the user that will intend to explicitly activate the BCI. Instead the BCI inhibitor monitors user state to assess the readiness of the user. For this reason BCI inhibitor can be viewed as an implicit (or passive) counterpart to the brain-switch. We expect that the combination of a classical BCI with a BCI inhibitor will result in an improvement of the overall performance, along with more comfort for the user. Recently, Panicker et al. described a P300 speller BCI enhanced with a constant flickering [6]. The P300-based BCI was paused when no Steady State Visual Evoked Potential (SSVEP) response was detected. This system can be seen as a BCI inhibitor in which the inhibition condition is "user is not looking at

the screen" and the inhibition signal is a SSVEP measured by EEG. Several other sensor channels like electromyography (EMG) or electrooculography (EOG) could be chosen as inhibitor signal. However the monitoring of the brain activity through EEG seems to be particularly adapted and could provide relevant information to the BCI inhibitor system. Numerous EEG markers appear to be useful to inhibit a BCI: Error potentials, rhythms associated to attention level etc. Another approach is to use features directly correlated to the BCI control signal. This will allow inhibiting the system until some specific conditions about the control signal are met. The usage of BCI inhibitor seems to be relevant to both asynchronous and synchronous BCI. In the asynchronous context, we can for example think of a hybrid BCI consisted of a Brain switch, followed by a BCI inhibitor that ensures that the user is in the desired state before enabling the control BCI that actually drives the interaction. In the context of synchronous BCI an inhibitor can be used between phases to put the system in standby until specific conditions are satisfied.

To evaluate the inhibitor process based on this last idea, we designed a hybrid BCI by combining a synchronous motor imagery based BCI that uses beta ERS posterior to feet movement ("beta rebound"), with a BCI inhibitor that checks whether the signal in the beta frequency band does not show any burst of activity before starting the control period.

3 Method

Participants: Four participants took part in the experiment, respectively aged of 27, 24, 28 and 26. They had never used any motor imagery based BCI system before.

Setup: EEG signals were recorded using a g.USBAmp (G.Tec) amplifier, sampled at 512 Hz. The setup was made of 7 electrodes, positioned according to the 10-20 system: a ground electrode (located on AFz position) and a reference electrode (located on the left earlobe), along with five measurement electrodes on Cz, C1, C2, FCz and Cpz. This EEG setup allowed us to record the EEG activity related to motor imagery of the feet [7]. The application is based on the Virtual Reality application "Use the force" presented by Lotte et al. [7]. A virtual spaceship is displayed on the screen. The goal is to lift the spaceship by doing motor-related tasks: real and imaginary feet movements. Whenever a burst is detected in the beta activity related to feet movement, the application raises the spaceship proportionally. Instructions can be displayed, asking the participant to either stand still, move or stop.

Procedure: The experiment was divided into 2 parts: a baseline and the series of motor imagery trials, using real movements or imaginary movements. The baseline consisted of a 25 sec period, where the participants were asked to stand still and relaxed, eyes opened. No feedback was provided during the baseline which was done once, at startup. The trial sequence is inspired by the startup sequence of an athletics run: Ready, Steady, Go. During one trial, the participants were instructed to relax ("Ready") for a certain period of time, then waited for 1 sec ("Steady"). Finally the "Move" instruction is displayed during 3 sec, followed by "Stop" during 3 sec. The movement of the feet done during the "Move" phase was instructed to be either real or imaginary. The "Stop" instructed the user to stop doing movement which should induced a beta rebound. The BCI inhibitor was either activated or deactivated during these trials without telling the participants. The "Ready" phase lasted 3 sec without inhibitor. Once the inhibitor was activated, the duration of the "Ready" phase varied from a minimum of 0.5 sec to a maximum of 10 sec. Participants were asked to start with 6 real movement sessions, followed by 6 imaginary movement sessions. The BCI inhibitor was activated on half of the sessions, randomized to eliminate an order effect. Each session was made of 10 trials with 4 sec between trials. The whole experiment (setup, trials, and questionnaire) lasted about 1 hour.

Signal Processing: EEG acquisition and online processing were conducted using the open-source software OpenViBE [8]. The EEG signal is band-pass filtered in 2–40 Hz band. Then, a Laplacian spatial filter centered on Cz is computed. The signal was then filtered in the Beta

band (16–24 Hz). A band power technique was applied to compute the power of the Beta band. We distinguish then two signals processed: the Control Signal (CS) that is used to control the spaceship, and the Inhibitor Signal (IS) used by the BCI inhibitor to decide to either launch or not launch the BCI-based interaction.

We define the **Control Signal (CS)** as the beta band power extracted on a 1 s window every 100 ms. The last 4 features were averaged with a moving window to produce a smooth control signal. To detect the post movement Beta ERS, the CS was compared to a threshold: $\mathsf{Th1} = \mathsf{baseline}_{\mathsf{mean}} + 3 * \mathsf{baseline}_{\mathsf{std}}$; where $\mathsf{baseline}_{\mathsf{mean}}$ and $\mathsf{baseline}_{\mathsf{std}}$ correspond respectively to the average and standard deviation of the control signal during the baseline phase.

We define the **Inhibitor Signal (IS)** as the beta band power extracted on a 2 s window every 500 ms. The IS was compared to a threshold $\mathsf{Th2} = \mathsf{baseline}_{\mathsf{mean}} + 1 * \mathsf{baseline}_{\mathsf{std}}$. If the computed control signal stayed below $\mathsf{Th2}$ 99 % of the time then the inhibition was deactivated and the BCI started. The maximum time of the inhibition was 10 s, after this the BCI started anyway.

4 Results

Table 1 shows the participants' performance for each condition (inhibitor on vs. inhibitor off, real movements vs. imaginary movements) and the duration of the *"Ready" phase*. To assess the participants' performance we counted the number of false positives (FP) and the number of true positives (TP). A false positive occurs if the value of CS went at least once above Th1 during a "Move" phase. A true positive occurs if CS signal went at least once above Th1 during a "Stop" phase. What happened during the other phases was not taken into account. We also computed the Hit-False (HF) difference which is equal to the number of TP minus the number of FP.

	Task	Inhibitor	Duration of the *"Ready Phase" (sec)*	FP	TP	HF
Subject 1	Real	on	1.85 ± 1.03	9/30	29/30	20
		off	3.00 ± 0.00	3/30	30/30	27
	Imaginary	on	1.30 ± 0.63	6/30	14/30	8
		off	3.00 ± 0.00	2/30	12/30	10
Subject 2	Real	on	3.89 ± 2.94	19/30	28/30	9
		off	3.00 ± 0.00	24/30	29/30	5
	Imaginary	on	1.25 ± 0.65	15/30	18/30	3
		off	3.00 ± 0.00	18/30	14/30	-4
Subject 3	Real	on	6.56 ± 3.26	14/30	30/30	16
		off	3.00 ± 0.00	21/30	30/30	9
	Imaginary	on	4.37 ± 3.18	8/30	17/30	9
		off	3.00 ± 0.00	19/30	22/30	3
Subject 4	Real	on	3.04 ± 2.88	8/30	30/30	22
		off	3.00 ± 0.00	13/30	28/30	15
	Imaginary	on	2.65 ± 2.86	8/30	16/30	8
		off	3.00 ± 0.00	16/30	14/30	-2
Average	Real	on	3.84	12.5/30	29.25/30	16.75
		off	3.00	15.25/30	29.25/30	14.00
	Imaginary	on	2.39	9.25/30	16.25/30	7.0
		off	3.00	13.75/30	15.5/30	1.75

Table 1: Performance achieved with the motor imagery BCI with and without BCI inhibitor in real and imaginary condition and mean duration of the *"Ready" phase*. Last row provides the average values over participants.

5 Discussion

Results suggest that BCI inhibitor works and provides an effect on the BCI behavior. It is materialized by different inhibition times for each user. Even if we have to take some cautions considering the limited number of participants it seems that the BCI inhibitor is able to improve the system performance: the average Hit-False difference over subjects is higher when the inhibitor was enabled for real movement condition (16.75 vs. 14.0) and imaginary condition (7.0 vs. 1.75). This result is mainly due to the reduction of false positive (e.g. subjects 2, 3 and 4).

After each session, participants were asked whether they felt a difference between the two conditions (with and without inhibitor) and to quantify it on a scale between 1 (not at all) and 7 (very). Suprisingly the participants reported only a small difference (2.2 ± 0.92), which suggests that this BCI inhibitor tends to be transparent to the users' perception.

The participants could also provide additional comments about each condition. Subject 1 was the only one who identified a different timing between the two conditions. During sessions with inhibitor enabled, he reported to be "caught by surprise" due to very short inhibition delay. This suggests that setting a higher minimum inhibition's duration could help to avoid such an effect.

6 Conclusion

In this paper we have introduced the concept of BCI inhibitor which can be defined as a system that pauses the BCI until some specific conditions are met. We presented a pilot study in the case of a synchronous motor imagery based BCI. Preliminary results with four participants suggest that inhibition process can be used to improve system performance. These hypotheses should be confirmed with more subjects. Future work should also address the use of BCI inhibitor for other paradigms such as P300 and SSVEP. Exploring adapted inhibitor signals (e.g. EEG markers correlated to level of attention) seems to be particularly relevant in these cases.

Acknowledgments

This work was supported by the French National Research Agency within the OpenViBE2 project (ANR-09-CORD-017). The authors would also like to thank Fabien Lotte(INRIA Bordeaux Sud-Ouest) for its helpful remarks.

References

[1] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi. A review of classification algorithms for EEG-based brain–computer interfaces. *Journal of Neural Engineering*, 4, 2007.

[2] G. Pfurtscheller, B. Z. Allison, C. Brunner, G. Bauernfeind, T. Solis-Escalante, R. Scherer, T. O. Zander, G. Müller-Putz, C. Neuper, and N. Birbaumer. The hybrid BCI. *Front. Neurosci.*, 2010.

[3] P. W. Ferrez and J. R. Millán. Simultaneous real-time detection of motor imagery and error-related potentials for improved BCI accuracy. In *4th international BCI Workshop*, 2008.

[4] S. G. Mason and G. E. Birch. A brain-controlled switch for asynchronous control applications. *Biomedical Engineering, IEEE Transactions on*, 47(10):1297–1307, oct 2000.

[5] G. Pfurtscheller, T. Solis-Escalante, R. Ortner, P. Linortner, and G. R. Müller-Putz. Self-paced operation of an SSVEP-based orthosis with and without an imagery-based "brain switch": A feasibility study towards a hybrid BCI. *IEEE Trans. Neural Syst. Rehab.*, pages 409–414, aug. 2010.

[6] R. C. Panicker, S. Puthusserypady, and Y. Sun. An asynchronous P300 BCI with SSVEP-based control state detection. *Biomedical Engineering, IEEE Transactions on*, 58(6):1781–1788, june 2011.

[7] F. Lotte, Y. Renard, and A. Lécuyer. Self-paced brain-computer interaction with virtual worlds: A quantitative and qualitative study "out of the lab". In *4th international BCI Workshop*, 2008.

[8] Y. Renard, F. Lotte, G. Gibert, M. Congedo, E. Maby, V. Delannoy, O. Bertrand, and A. Lécuyer. OpenViBE: An open-source software platform to design, test and use brain-computer interfaces in real and virtual environments. *Presence*, 19:35–53, Apr. 2010.

Performance of a P300-Based BCI System Improved by a Bayesian Single-Trial ERP Estimation Technique

A. Goljahani[1], C. D'Avanzo[1], C. Genna[1], S. Silvoni[2], F. Piccione[2], G. Sparacino [1]

[1]Deparment of Information Engineering, University of Padova, Padova, Italy
[2]IRCCS, San Camillo Hospital, Venice, Italy

anahita.goljahani@dei.unipd.it

Abstract

Brain computer interface (BCI) systems based on electroencephalographic (EEG) signals are appealing given their non invasiveness, high temporal resolution, portability and low set-up cost. In particular, P300-based BCI does not require any previous long training of the subject. In this work we assess the improvement of classification performance obtained in a P300-based BCI system by "preprocessing" the signal by a Bayesian filtering procedure for single trial ERP estimation. The reference system is the BCI prototype designed at the IRCSS San Camillo Hospital (Venice, Italy), which embeds a preprocessing procedure based on independent component analysis (ICA). Results from two healthy subjects and four patients affected by amyotrophic lateral sclerosis (ALS) show that classification errors relative to the Bayesian approach for single-trial ERP estimation are at least halved with respect to the reference ICA method.

1 Introduction

In BCI systems, brain activity is treated as a signaling source that is directly utilized as a communication means without the activation of any motor output channels. Both electrophysiological [1] and hemodynamic signals [2] from the brain have been utilized as means to decode the subject intent; however, at present days BCIs for human research are mainly based on non invasive EEG recordings. In particular, three types of EEG signals are the mostly addressed: slow cortical potentials, sensorimotor rhythms and the P300 event related brain potential (ERP), see [1] for a review. Systems based on the first two signals require a training lasting some days or weeks [1], whereas P300 based systems usually need short single sessions for pilot/training [3] or calibration/machine learning purposes [4]. BCI systems that are based on P300 exploit the cognitive potential elicited when a rare task-relevant stimulus (target) is presented among non significant ones (non target). In this kind of BCI systems one major goal is that of separating the evoked potential from the superimposed background EEG. In order to achieve this, advanced signal processing and estimation techniques are applied to the signal after acquisition and before feature extraction steps, see [5] for the description of the elaboration blocks of a general BCI system. In this work, for the sake of reasoning, this intermediate step is referred to as "preprocessing".

The present study was carried out to investigate the convenience of utilizing a recent Bayesian approach for single-trial ERP estimation [6] as preprocessing procedure in a P300-based BCI system. To this aim, as a preliminary work, the performance of a BCI system employing this preprocessing technique was evaluated off-line on a small dataset of 6 subjects. Features extraction and classification procedures were implemented according to the BCI system designed at the IRCSS San Camillo Hospital (Venice, Italy) [7]. The latter, with its own preprocessing technique which is based on the independent component analysis (ICA), was taken as reference system. The performances were evaluated off-line in terms of classification errors obtained utilizing EEG data previously recorded during BCI sessions from two healthy subjects and four patients affected by amyotrophic lateral sclerosis (ALS). Results show that the Bayesian approach for single-trial ERP estimation determines an improvement of the BCI classification errors, which are at least halved with respect to the reference ICA method.

2 Methods

2.1 Data Base

Participants. Two healthy subjects (S1, S2) and four patients with ALS (P1, P2, P3 and P4) participated in a BCI experiment. ALS diagnosis was based on the El Escorial criteria and the patient was admitted to the S. Camillo Hospital for rehabilitation treatment.
Stimulation paradigm and data acquisition. Participants were asked to control the movement of a cursor from the center of the screen to one of four possible points in peripheral positions. Upward, rightward, downward and leftward arrows were randomly flashed to indicate possible movement directions. The participant had to pay attention to the flashing of the arrow pointing to the desired position. This flashing was considered as the onset of the target stimulus and all the others as the onset of non target stimuli. During the BCI sessions, under the hypothesis that the target stimulus elicits the P300 wave, the cursor was moved on the screen in the direction of the flashed arrow, every time the P300 was detected (see [7] for a comprehensive protocol description). EEG was recorded from electrodes placed at Fz, Cz, Pz and Oz according to the international 10/20 system. The electrooculogram (EOG) was also recorded from a pair of electrodes below and laterally to the left eye. Data were band-pass filtered at 0.15 and 30 Hz and digitized at 200 Hz sampling rate. Further details can be found in [7]. Epochs lasting 1.5 s ([-0.5 1] s with respect to the stimulus onset) were extracted from the recorded signal. Time-accuracy of the stimulus onset was assessed by means of an external optical sensor [8] (the delay between stimulus onset triggering and its actual appearance had a mean of 100 ms and a standard deviation of 18 ms). After discarding the epochs corrupted by artifacts, a total of N_t target and N_{nt} non target epochs were available for S1 ($N_t = 118$, $N_{nt} = 327$), S2 ($N_t = 196$, $N_{nt} = 543$), P1 ($N_t = 205$, $N_{nt} = 615$), P2 ($N_t = 137$, $N_{nt} = 387$), P3 ($N_t = 118$, $N_{nt} = 283$) and P4 ($N_t = 169$, $N_{nt} = 591$). These epochs, from now on called "raw" epochs, were utilized off-line to evaluate the classification performance.

2.2 Preprocessing Algorithms under Test

Independent component analysis (ICA). For target and non target epochs, the reference preprocessing procedure consists in an ICA decomposition of the raw epochs and the selection, by means of a fuzzy method, of the source reflecting the P300 wave [9]. The selected ICA component feeds the features extraction stage described in Section 2.3.
Bayesian two stages (B2S). The single trial estimation method proposed in [6] is a two stages procedure denoted as B2S. In both stages, a priori information on the statistical properties of the evoked potential and of the background EEG are utilized to extract the P300 wave at a single trial level from raw epochs. The method is applied to a set of N epochs. In the first stage, N filtered sweeps are obtained and are employed to compute the average ERP. This average is utilized as a further a priori information in the second stage in order to obtain the N single trial estimates. Mathematical details on the method can be found in [6]. The B2S approach was applied to target raw epochs from Pz electrode.

2.3 Features Extraction, Classification and Performance Assessment

For both procedures of Section 2.2, we denote as "preprocessed" the signals yielded by the preprocessing. For each of the preprocessed epochs, 78 parameters were determined by a feature extraction algorithm described in [9]. These features were used as input to a support vector machine (SVM) classifier, which categorized the epoch as target or non target [7]. The performance of the preprocessing approaches was evaluated in terms of classification errors of the two BCI systems explained in the introduction. In order to obtain stable results, 40 classifiers were utilized for each system as follows. For the ICA approach, each classifier was trained with a group of $N_t^{(tr)}$ target and $N_{nt}^{(tr)}$ non target preprocessed epochs, randomly selected among the available ones for the participant ($N_t^{(tr)} = 111.7$ and $N_{nt}^{(tr)} = 320$ on average). Then, another group of $N_t^{(test)}$ target and $N_{nt}^{(test)}$ non target preprocessed epochs ($N_t^{(test)} = 43.6$ and $N_{nt}^{(test)} = 127.2$ on average)

was utilized to obtain the classification errors on target, non target and total epochs, defined, respectively, as follows [9]: number of target epochs classified as non target divided by the number of target epochs; number of non target epochs classified as target divided by the number of non target epochs; number of total erroneously classified epochs divided by the number of total epochs. Finally, these errors were averaged over the 40 classifiers to obtain e_t, e_{nt} and e_{tot}, respectively. Similarly, for the B2S approach, each classifier was trained with $N_t^{(tr)}$ preprocessed target epochs and $N_{nt}^{(tr)}$ non target raw epochs, randomly selected from the available ones. Classification errors were obtained utilizing another group of $N_t^{(test)}$ preprocessed target and $N_{nt}^{(test)}$ non target raw epochs and e_t, e_{nt} and e_{tot} were computed as in the ICA case.

3 Results

B2S single trial estimation. As examples of B2S estimates for healthy subjects and patients, in Figure 1 we report two representative raw epochs (red dashed line) from S1 (graph (a)) and P4 (graph (b)) and the correspondent single trial estimates (black solid line).
Classification results. Table 1 shows results on classification errors on target (e_t), non target (e_{nt}) and total epochs (e_{tot}) after ICA (first row) and B2S (second row) preprocessing for all of the 6 subjects.

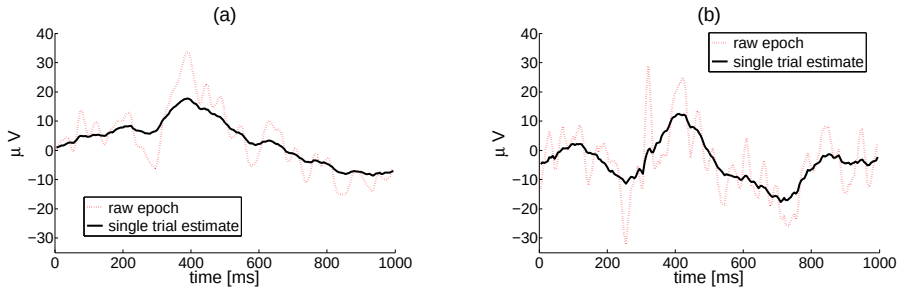


Figure 1: Representative raw epochs (red dashed line) and single trial estimates (black solid line) from S1 (graph (a)) and P4 (graph (b)).

	S1 (e_t,e_{nt},e_{tot})	S2 (e_t,e_{nt},e_{tot})	P1 (e_t,e_{nt},e_{tot})	P2 (e_t, e_{nt},e_{tot})	P3 (e_t,e_{nt},e_{tot})	P4 (e_t,e_{nt},e_{tot})
ICA	0.28, 0.07, 0.13	0.42, 0.13, 0.21	0.82, 0.13, 0.27	0.54, 0.12, 0.25	0.44, 0.2, 0.26	0.33, 0.06, 0.11
B2S	0.05, 0.01, 0.02	0.05, 0.03, 0.04	0.05, 0.01, 0.01	0.04, 0.01, 0.02	0.04, 0.02, 0.03	0.06, 0.03, 0.03

Table 1: Classification errors on target (e_t), non target (e_{nt}) and total epochs (e_{tot}). Columns correspond to subjects, rows to the ICA and B2S preprocessing approach, respectively.

4 Discussion

The effectiveness of the B2S approach in detecting single trial evoked potentials is demonstrated by the single trial estimates reported in graphs (a) and (b) of Figure 1. In fact, from Figure 1, the technique is visibly capable of extracting the P300 wave (black solid line) from the raw signal (red dashed line) even when the background EEG heavily distorts the underlying evoked potential, as in the case of the ALS patient (panel (b)). The signal denoising achieved with the B2S method is beneficial to classification performance. In fact, as reported in Table 1, all classification errors were at least halved with respect to the current ICA approach. In particular e_t, that is closely related to the P300 detection, is at least about five times smaller for B2S with respect to ICA

(e.g., for S1) and is best improved for P1, whose P300 classification error with ICA preprocessing is about sixteen times the one with B2S preprocessing.

5 Conclusion

This work was intended to be a preliminary off-line study to evaluate the convenience of utilizing a single trial estimation technique, the B2S, as preprocessing procedure in a BCI system. Results demonstrated a substantial improvement of classification errors in the case of B2S preprocessing with respect to the currently adopted ICA procedure. The satisfying outcomes encourage further investigation to extend the analysis to a wider dataset and to assess the usability of the B2S procedure in an on-line system.

6 Acknowledgments

This work was supported by the University of Padova grant N. 17/2011 to AG and by the Regione Veneto FSE grants cod. 2105/1/6/2215/2009 to AG and cod. 2105/201/4/1102/2010 to CDA.

References

[1] N. Birbaumer and L. G. Cohen. Brain-computer interfaces: communication and restoration of movement in paralysis. *The Journal of Physiology*, pages 621–636, 2007.

[2] R. Sitaram, A. Caria, and N. Birbaumer. Hemodynamic brain-computer interfaces for communication and rehabilitation. *Neural Networks*, 22:1320–1328, 2009.

[3] 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.

[4] B. Blankertz, G. Dornhege, M. Krauledat, K. R. Müller, and G. Curio. The non-invasive berlin brain-computer interface: Fast acquisition of effective performance in untrained subjects. *NeuroImage*, pages 539–550, 2007.

[5] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan. Brain-computer interfaces for communication and control. *Clinical Neurophysiology*, 113:767–791, 2002.

[6] C. D'Avanzo, S. Schiff, P. Amodio, and G. Sparacino. A bayesian method to estimate single-trial event-related potentials with application to the study of the P300 variability. *Journal of Neuroscience Methods*, 2011, In press.

[7] S. Silvoni, C. Volpato, M. Cavinato, M. Marchetti, K. Priftis, A. Merico, P. Tonin, K. Koutsikos, F. Beverina, and F. Piccione. P300-based brain-computer interface communication: evaluation and follow-up in amyotrophic lateral sclerosis. *Neural Networks*, 22:1320–1328, 2009.

[8] J. A. Wilson, J. Mellinger, G. Schalk, and J. Williams. A procedure for measuring latencies in brain-computer interfaces. *IEEE Trans Biomed Eng.*, 57:1785–1797, 2010.

[9] 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*, 117:531–537, 2006.

Comparison of Feature Extraction Methods for Brain-Computer Interfaces

P. Ofner[1], G. R. Müller-Putz[1], C. Neuper[1,2], C. Brunner[1]

[1]Institute for Knowledge Discovery, Graz University of Technology, Graz, Austria
[2]Department of Psychology, University of Graz, Graz, Austria

patrick.ofner@tugraz.at

Abstract

This paper compares classification accuracies of feature extraction methods (FEMs) as used in sensory motor rhythm (SMR) based Brain-Computer Interfaces (BCIs). Features were extracted offline from 9 subjects and classified with linear discriminant analysis. The following FEMs were compared: adaptive autoregressive parameters, band power, phase locking value, time domain parameters, and Hjorth parameters. FEM parameters were optimized individually with a genetic algorithm in advance. In summary, time domain parameters combined with a bipolar spatial filter yielded the best classification accuracies.

1 Introduction

In most sensory motor rhythm (SMR) based Brain-Computer Interfaces (BCIs), feature extraction methods (FEMs) typically extract features from preprocessed electroencephalographic (EEG) recordings. Afterwards, these features are classified. In general, different FEMs in combination with a specific classifier - here, we used linear discriminat analysis (LDA) as a binary classifier - lead to different classification accuracies. Usually, one wants to use the FEM yielding the highest classification accuracy. Therefore, this work compares the following popular FEMs with respect to the achieved classification accuracies: adaptive autoregressive (AAR) parameters [1], bilinear AAR (BAAR) parameters [2], multivariate AAR (MVAAR) parameters [2], band power (BP) [3], phase locking value (PLV) [4], time domain parameters (TDP) [5], and Hjorth parameters [6].

Most FEMs contain meta parameters that must be set before the method can be applied. It is crucial to tune these meta parameters carefully to tap the full potential of these methods. Therefore, all meta parameters were optimizied in a subject-specific way with a genetic algorithm (GA) [7].

Previous work compared AAR, BP, and fractal dimension features without optimizing meta parameters [8]. Another study compared a slightly different method set (with and without subject-specific optimization): AAR, BP, Hjorth, TDP, Barlow, Wackermann, and Brain-Rate [9].

2 Test Setup

EEG Data Prerecorded data from the BCI competition IV (data set 2A) [10] were used for an offline analysis. There 22 Ag/AgCl electrodes with inter-electrode distances of 3.5 cm were used. Two sessions from each of 9 participants were recorded. One session comprised 6 runs, each with 48 trials. Originally, four equally distributed classes - motor imagery (MI) of the left hand (class 1), right hand (class 2), both feet (class 3) or tongue (class 4) - were recorded in a cue-based paradigm without feedback. However, for the analysis in this work, only classes 1 and 2 were used. Cues were presented on a computer screen as follows: first, a fixation cross appeared together with a short signal tone indicating the start of a trial; at second 2, a cue appeared for 1.25 s (left arrow

for class 1, right arrow for class 2) prompting participants to perform the required MI task; at second 6, the fixation cross disappeared and a short break followed.

The EEG signals were originally band-pass filtered from 0.5 Hz to 100 Hz and recorded monopolarly at 250 Hz (left mastoid served as reference, right mastoid as ground). In addition, a 50 Hz notch filter was applied. To reduce data, signals were lowpass-filtered with 55 Hz and downsampled to 125 Hz. Laplace, common average reference (CAR), and bipolar derivations were calculated from the original monopolar data.

Comparison of Feature Extraction Methods First, FEMs were optimized in the *optimization step* with a genetic algorithm, and afterwards tested in the *evaluation step*. Data from session 1 of each subject was used in the optimization step and for training an LDA classifier in the evaluation step. Data from session 2 was used solely for testing the LDA classifier in the evaluation step.

All FEMs used data from C3 and C4 positions (10-20 system) with a specific spatial filter (monopolar, bipolar, Laplace, CAR). Thereby, three types of bipolar spatial filters were used: FC3/4-C3/4, C3/4-CP3/4, FC3/4-CP3/4. The type of bipolar spatial filter with the best fitness score (best classification accuracy) in the optimization step for a subject was used. In addition, the PLV FEM used four channels (two channels per hemisphere) in various arrangements, because inter-hemispheric coupling was expected to contain discriminative information [4]. The channel combination leading to the best fitness score in the optimization step was used for further analysis.

Optimization Step A GA optimized the meta parameters of a certain FEM. Therefore, the actual meta parameters were represented by an *individual*, and features were extracted according to these parameters from session 1. The classification accuracies over the trial were calculated with a 10×5 cross-validation procedure using an LDA classifier. Finally, the 0.9 quantile (which is more robust as e.g. the maximum) of these classification accuracies between the cue and end of trial was used as the *fitness score*. The GA maximized this fitness score.

Evaluation Step Features were extracted using the optimized meta parameters from the previous step. An LDA classifier was trained with features from session 1 of a subject and tested against features from session 2 of the same subject. The 0.9 quantile of the classification accuracy reached by the LDA classifier (between cue and end of trial) was used as the final classification accuracy reported for that FEM. Thus, for each method/spatial filter/subject combination, the classification accuracy was evaluated.

3 Results

Descriptive Statistics Figure 1 shows a box-and-whisker plot including mean values (dotted lines). Methods are sorted from left to right by their mean classification accuracies, and only spatial filters yielding the highest mean classification accuracy are shown. TDP with a bipolar spatial filter reaches the highest *mean* classification accuracy of 78 % (standard deviation 11 %, median 82 %). MVAAR with a bipolar spatial filter reaches the highest *median* classification accuracy of 83 % (standard deviation 13 %, mean 74 %). A bipolar spatial filter yields the highest mean (and median) classification accuracies for AAR, BP, Hjorth, MVAAR and TDP.

Figure 2 shows mean values and standard deviations of all FEMs and spatial filters. Bipolar and Laplacian filters yield higher mean classification accuracies than monopolar and CAR spatial filters - except for PLV, where the opposite is true.

Inferential Statistics A two-way repeated measures ANOVA was applied to test for significant effects of factors SPATIAL-FILTER ($F_{3,24} = 8.63$) and FEATURE-EXTRACTION-METHOD ($F_{6,48} = 8.20$) on the classification accuracy (dependent variable). If the sphericity assumption was violated, p-values have been corrected according to Huynh & Feldt. Both main effects and the interaction effect show significant p-values ($p < 0.05$). Tukey's Test has been used to test for significant differences in group mean values of the groups shown in Figure 1 (FEMs with their best spatial filter) and spatial filters when using TDP. PLV (CAR) differs significantly from TDP

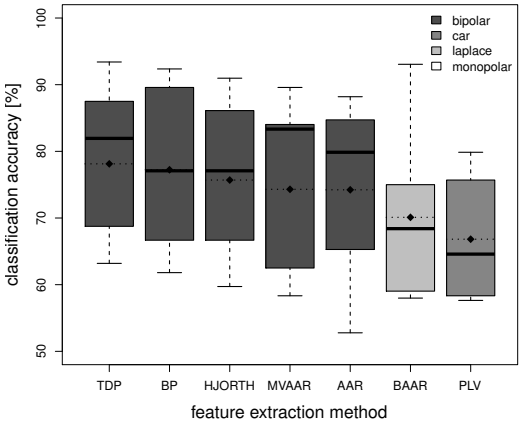


Figure 1: Box-and-whisker plot of classification accuracies when the spatial filter with the highest mean accuracy for each FEM is used. Additionally, the dotted line with the square marks mean classification accuracies.

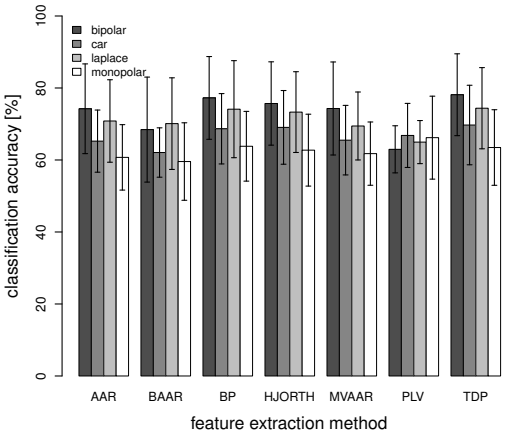


Figure 2: Mean values and standard deviations of classification accuracies for all FEMs and all spatial filters.

(bipolar) and BP (bipolar). A monopolar spatial filter is significantly worse than bipolar and Laplacian filters when using the method with the highest mean classification accuracy (TDP).

4 Discussion

No significant differences in the classification accuracies between AAR, BP, Hjorth, MVAAR, TDP and BAAR have been found when using the best (mean/median) spatial filter for a method. These results are corroborated by Vidaurre et. al [5], who also found no significant differences between AAR, BP, Hjorth, and TDP when using subject-specific optimization. However, PLV is significantly worse than BP/TDP. A reason could be that synchronization *between* neuronal assemblies occurs when they are processing information and therefore showing a decrease in power (event-related desynchronization ERD [11]) *within* each. This ERD makes PLV prone to noise.

Furthermore, a significant difference between monopolar and bipolar/Laplacian filters was found using the FEM with the highest mean classification accuracy (TDP). This is because bipolar/Laplacian filters, as opposed to a monopolar filter, eliminate noise common to the used electrodes. Figure 2 suggests that bipolar filters reach higher classification accuracies than Laplacian filters. Areas producing ERD patterns due to MI [11] are located anterior of C3/C4, which can be better covered by FC3/4-C3/4 bipolar filters than Laplacian filters focused directly on C3/C4. Thus, in this setup, bipolar filters yielded higher classification accuracies.

Results are only valid when using an LDA classifier, because in general, the performance of a classifier depends on the distribution of the input data, which depends on the FEM. However, LDA is widely used in the field of BCI research and thus, these results are useful for many BCI implementations.

To get meaningful results, it is absolutely necessary that the final testing is performed on *unseen* data to avoid underestimating the error. Therefore, session 2 of a subject was only used in the evaluation step.

Futhermore, one must keep in mind that a GA is a meta-heuristic optimization method and therefore likely finds a good solution, but not necessarily a global optimum.

5 Conclusion

No significant differences were found between TDP, BP, Hjorth, MVAAR, AAR, BAAR. However, TDP with a bipolar spatial filter yielded the highest mean classification accuracy, a high median classification accuracy, is computationally efficient, has less parameters to set (less need for subject-specific optimization), and is therefore favorable of all compared feature extraction methods. This conclusion is restricted to the usage of an LDA classifier, a continuous MI task of left hand versus right hand, and the usage of a small number of electrodes.

Acknowledgments

This work is supported by the European ICT Programme Project FP7-224631. 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] A. Isaksson, A. Wennberg, and L. H. Zetterberg. Computer analysis of EEG signals with parametric models. *Proceedings of the IEEE*, 69:451–461, 1981.

[2] C. Brunner, M. Billinger, and C. Neuper. A comparison of univariate, multivariate, bilinear autoregressive, and bandpower features for brain-computer interfaces. In *Fourth International BCI Meeting, Asilomar*, 2010.

[3] G. Pfurtscheller, D. Flotzinger, M. Pregenzer, J. R. Wolpaw, and D. J. McFarland. EEG-based brain computer interface (BCI) - search for optimal electrode positions and frequency components. *Medical Progress Through Technology*, 21:111–121, 1995.

[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(12):2501–2506, December 2006.

[5] 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, November 2009.

[6] B. Hjorth. EEG analysis based on time domain properties. *Electroencephalography and Clinical Neurophysiology*, 29(3):306–310, September 1970.

[7] J. Holland. *Adaptation in Natural and Artificial Systems.* University of Michigan Press (reprinted in 1992 by The MIT Press), 1975.

[8] R. Boostani, B. Graimann, M. H. Moradi, and G. Pfurtscheller. A comparison approach toward finding the best feature and classifier in cue-based BCI. *Medical and Biological Engineering and Computing*, 45:403–412, 2007.

[9] C. Vidaurre and A. Schlögl. Comparison of adaptive features with linear discriminant classifier for brain computer interfaces. In *30th Annual International IEEE EMBS Conference, Vancouver*, 2008.

[10] C. Brunner, R. Leeb, G. Müller-Putz, A. Schlögl, and G. Pfurtscheller. BCI competition 2008 - Graz data set A. http://www.bbci.de/competition/iv/desc_2a.pdf, 2008.

[11] G. Pfurtscheller and F. H. Lopes Da Silva. Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical Neurophysiology*, 110(11):1842–1857, 1999.

Wavelet Design by Means of Multi-Objective GAs for Motor Imagery EEG Analysis

J. Asensio[1], E. Galvan, R. Palaniappan[1], J. Q. Gan[1]

[1]School of Computer Science and Electronic Engineering, University of Essex, United Kingdom
jasens@essex.ac.uk;edgar.galvan@gmail.com;palani,jqgan@essex.ac.uk

Abstract

Wavelet-based analysis has been broadly used in the study of brain-computer interfaces (BCI), but in most cases these wavelet functions have not been designed taking into account the requirements of this field. In this study we propose a method to automatically generate wavelet-like functions by means of genetic algorithms. Results strongly indicate that it is possible to generate (evolve) wavelet functions that improve the classification accuracy compared to other well-known wavelets (e.g. Daubechies and Coiflets).

1 Introduction

Wavelets are wave-like oscillations used to extract relevant information from a given signal. They have useful features such as the time-frequency resolution, which make them suitable for electroencephalography (EEG) signals unlike the Fourier Transform (FT) that only gives information in the frequency domain [1] (although related methods like the short time FT provides time-frequency resolution). Several related approaches have already been used in the BCI research such as the continuous wavelet transform [2], the discrete wavelet transform (DWT) [3] and wavelet packages [4].

The main goal of this paper is to explore the possibility of automatically creating a wavelet that can be at least as good as other well-defined wavelets. Inspired by other works (e.g., [5–7]), we will explore the possibility to automatically evolve wavelets by means of Genetic Algorithms (GAs). It should be noted that the main difference between this work and the works mentioned previously is that we use GAs in order to evolve directly the filter coefficients for a one-way DWT.

This paper is organised as follows. In the next section we describe our approach. In Section 3, we present and discuss the results. In Section 4 we draw some conclusions.

2 Approach

2.1 Data Description

Each of the subjects, three in total, sat in a comfortable arm chair, one metre away from a 19" screen. EEG signals were recorded from five channels using bipolar electrode positions with respect to the international 10-20 system: FC3-PC3, FC1-PC1, Cz-Pz, FC2-PC2, FC4-PC4. The recording was made with a 16-channel EEG amplifier from g.Tech and sampled at 250 Hz. Each subject went through four sessions and 30 trials were recorded in each one. In each trial an arrow was shown from t=3 s to 8 s, with the arrow pointing left if the user was meant to imagine left hand movement, right for right hand and down for feet. Only two classes are taken into account in this study, left hand and feet, making a total of 80 trials for each subject.

2.2 Wavelet Processing

To assess the performance of the wavelet function obtained by means of GAs, first it is necessary to perform an equivalent experiment using a standard wavelet. For this purpose, we selected Daubuchies and Coiflets wavelets as they have been widely used in other studies and their accuracy has been well assessed [3, 4, 8]. For every cued instant in the dataset, a second of the signal from a specific point is transformed using the DWT. This procedure is repeated for each of the five channels, and all the outputs are joined and labeled with its class to build an input pattern. Then

a sliding window is applied, moving a 1/8 of a second ahead to take the next second of the signal to build the next pattern and so on, until the cued data is completely consumed.

A multi-frequency decomposition (MFD) is performed to obtain the most interesting coefficients from the wavelet transform. As the original data was acquired with a sampling frequency of 250 Hz the wavelet decomposition is done down to the 5^{th} level, by discarding D0 and D1 the coefficients for A5 (0–3.90 Hz), D5 (3.90–7.81 Hz), D4 (7.81–15.62 Hz), D3 (15.62–31.25 Hz) and D2 (31.25–62.5 Hz) are obtained. We selected this range as it covers from delta to gamma bands, letting the GA decide which elements are more important.

Dimensionality of the output from the MFDs is too high for classification purpose. Thus, Davis-Bouldin Index (DBI) [9] is used to measure the features' separability and then the best 25 are selected. The patterns obtained are classified using Fisher's linear discriminant (LDA) [10]. This model was used due to its simplicity and efficiency, it was also shown during BCI competitions 2003 and 2005 that LDA performs as well as Neural Networks or Support Vector Machines in terms of classification accuracy [11].

The classification process uses a five-fold cross-validation model. Also, the classification rates are calculated using trial by trial validation, i.e. a trial is classified as c if the majority of its samples are classified as c.

2.3 Evolving Wavelet Filters

Often, wavelets must fulfil some constraints such as the admissibility condition [1], so that the inverse transform can be calculated. In our case, there is no requirement to have a proper wavelet as we are not interested in applying the inverse DWT to reconstruct the original signal.

So, the aim is to generate a wave-like oscillation which increases the separability of the data when the wavelet transform is applied, but ignoring those restrictions that make the operation reversible. This, in consequence, makes the design of the fitness function used in our GA (described later in this section) easier than other scenarios (where the wavelet inverse transform is required).

The discrete wavelet transform can be described as a filtering and down-sampling waterfall process, the approximation and detail function for each level within the down sampling process is defined by a filter of length L and is given by $\phi(x) = \sum_n^L h_0(n)\phi(2x - n)$ and $\psi(x) = \sum_n^L h_1(n)\phi(2x - n)$, where h_0 and h_1 are the high-pass and low-pass filters, respectively.

2.3.1 Multi-Objective Genetic Algorithm

The aim of the proposed study, as stated previously, is to automatically evolve wavelets. Thus, it is necessary to evolve the coefficients of both filters h_0 and h_1. The filter length varies depending on the wavelet order. Also, these coefficients are defined by real numbers. The length of each chromosome in the population set is 20. Each chromosome is composed by two parts: the first half is designed to obtain the best value for h_0 and the second part aims at finding the best value for h_1. The original population is generated with random filter coefficient values in the range $[-1, 1]$.

The experiments were conducted using GA with tournament selection (size = 7), bit mutation (0.01) and two-point crossover (0.7). To obtain more reliable results, we performed 20 independent runs. Runs were stopped when the maximum number of generations (120) was reached.

As a first approach, we used a fairly simple fitness function based on the individual accuracy (raw fitness). A wavelet function is generated using the coefficients defined by each individual in the population; this function is then applied to every signal from the data set. Next, the fitness value is calculated as the outcome from the FDA. To assess a correct behaviour in the generated wavelets, the data set is divided into two different subsets (of 40 trials each): the first is used during the GA execution and the second is used to validate the performance of this newly generated function against the Daubechies and Coiflets wavelets.

We also tested a more robust approach by considering Multi-Objective (MO) GAs, in specific using a well proved algorithm called Non-Dominated Sorting GA [12]. Thus, the fitness function is formed by two elements: (a) the classification accuracy from FDA, as discussed earlier, and (b) the minimisation of the standard deviation in the classification accuracy among the different classes. For the latter objective, a weight of 0.6 has been applied in order to decrease its importance in the selection of the 'best so far' solution. In the following section, we present and discuss our findings.

Table 1: GA Mean result. SF stands for Single Fitness and MO for Multi-objective.

Subject	SF GA	SF Validation	MO GA	MO Validation
1	0.612 +/- 0.046	0.624 +/- 0.047	0.696 +/- 0.045	0.658 +/- 0.045
1 best	0.725 +/- 0.0447	0.680 +/- 0.057	0.762 +/- 0.0406	0.757 +/- 0.037
2	0.597 +/- 0.045	0.664 +/- 0.043	0.709 +/- 0.040	0.695 +/- 0.038
2 best	0.762 +/- 0.042	0.817 +/- 0.031	0.817 +/- 0.038	0.795 +/- 0.031
3	0.786 +/- 0.035	0.698 +/- 0.041	0.767 +/- 0.038	0.725 +/- 0.042
3 best	0.892 +/- 0.029	0.782 +/- 0.041	0.852 +/- 0.030	0.785 +/- 0.051

Table 2: Benchmark comparison.

Subject	Coif5	Db10	PSD	Generated Wavelet
1	0.688 +/- 0.116	0.690 +/- 0.130	0.730 +/- 0.120	**0.752 +/- 0.037**
2	0.667 +/- 0.162	0.633 +/- 0.110	**0.827 +/- 0.080**	0.690 +/- 0.130
3	0.665 +/- 0.140	0.690 +/- 0.150	0.697 +/- 0.147	**0.767 +/- 0.130**

3 Results and Discussion

The GA was run for 20 different rounds applying both single and multi-objective. The results are shown in Table 1. This table shows the mean of ten executions for the best individual in those rounds. The GA column values are the averaged accuracy of the best individuals against the set used for evolving the wavelets, whereas the validation column values show the averaged accuracy of the best individuals for the validation set. For each subject the results for the best wavelet are shown. In every case, the MO produced better results compared to the single fitness approach. This indicates that encouraging the GA to assure a balanced rate among class through classification accuracy leads to obtain a more robust wavelet against unseen trials. During the early experiments, an over-fitting problem arose and this was addressed by selecting only a random subset of every test fold during the evaluation. In most cases the population converged before the 120^{th} generation and was stable for at least 20 generations, therefore we can assume that the numbe of iterations was sufficient to allow the algorithm to evolve.

In Table 2 a comparison among the generated wavelets and other techniques is displayed. The columns *Db10* and *Coif5* shows the result for the process described in Section 2.2. The column *PSD* is the results of applying *Power Spectral Density* from 0 Hz to 62 Hz where the mean values for bands with width of 2 Hz each are computed, ending with 155 features per pattern. The DBI is applied to select the most 25 discriminating features. The column named *Generated Wavelet* shows the result of the best wavelet obtained by the GA and applied to the validation set, therefore the decision on which wavelet to use is not biased by the results against the validation set as shown in Table 1. Notice that the results shown in Table 2 are the average accuracies for the validation set from ten different runs where the data was randomly shuffled.

The results obtained show that the generated wavelet always perform better than the Daubechies and Coiflets. When compared to PSD the generated wavelet performed better for subjects 1 and 3 but worse in case of subject 2.

Frequency response study of the best generated filters in the Table 1 shows that all of them are stable as indicated by their phase responses. The Daubechies wavelet filters are designed in such a way that they are high-pass and low-pass filters. However, the behaviour found in the generated wavelets is different. All the filters obtained show a multi-band behaviour with the pass bands different for each h_0 and h_1. Due to the analytic complexity of the DWT, to state what each multi-band filtering implies is out of the scope for this study.

It is clear that the DBI feature selection strongly affects the evolutionary process but it allows us to draw out the most important frequency bands and channels involved in the classification. Studying the selected features, we can observe that for every subject the importance of each bipolar electrode varies. E.g. for the first subject, 5 % of the selected features come from FC4-PC4, whereas for the third subject, this ratio was up to 35 %.

If we focus our attention on the frequency bands, we find that for any user the wavelet decomposition level D2 corresponding to the 31.25 to 62.5 Hz frequency range occurs in 40 % of the features selected by DBI (although beta band features tend to rank higher in the DBI). This result

supports the outcome presented in [13] where it is stated that gamma band contains highly useful information for motor imagery classification. This behaviour is consistent among the generated functions but not in the Daubechies wavelet where every different subject presents a different distribution in the gamma range.

4 Conclusion

The results obtained in this paper show that the evolved wavelet performs better than Daubechies and Coiflets wavelets. This implies that using out-of-the-box wavelets might not be the best approach when dealing with EEG data. Thus, other evolved functions could improve the performance in wavelet based solutions, as shown in this study. It should be noted that one disadvantage of this approach is its efficiency, as it generates a different wavelet for each user. Future work will address this issue.

References

[1] P. S. Addison. *The Illustrated Wavelet Transform Handbook*. Institute of Physics Publ., 2002.

[2] V. Bostanov. BCI competition 2003-data sets ib and IIb: feature extraction from event-related brain potentials with the continuous wavelet transform and the t-value scalogram. *Biomedical Engineering, IEEE Transactions on*, 51(6):1057–1061, 2004.

[3] R. Scherer, S. P. Zanos, K. J. Miller, R. P. Rao, and J. G. Ojemann. Classification of contralateral and ipsilateral finger movements for electrocorticographic brain-computer interfaces. *Neurosurgical Focus*, 27(1):12–4967, 2009.

[4] J. Sherwood and R. Derakhshani. On classifiability of wavelet features for EEG-based brain-computer interfaces. In *Neural Networks, 2009. IJCNN 2009. International Joint Conference on*, pages 2895–2902. IEEE, 2009.

[5] M. M. Lankhorst, M. D. van der Laan, and W. A. Halang. Wavelet-based signal approximation with genetic algorithms. *Systems Analysis Modelling Simulation*, 43(11):1503–1528, 2003.

[6] E. Jones, P. Runkle, N. Dasgupta, L. Couchman, and L. Carin. Genetic algorithm wavelet design for signal classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 890–895, 2001.

[7] D. Farina, O. F. do Nascimento, M. F. Lucas, and C. Doncarli. Optimization of wavelets for classification of movement-related cortical potentials generated by variation of force-related parameters. *Journal of Neuroscience Methods*, 162(1-2):357–363, 2007.

[8] B. G. Xu and A. G. Song. Pattern recognition of motor imagery EEG using wavelet transform. *J. Biomedical Science and Engineering*, 1:64–67, 2008.

[9] J. C. Bezdek and N. R. Pal. Some new indexes of cluster validity. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 28(3):301–315, 1998.

[10] R. A. Fisher. The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, 7:179–188, 1936.

[11] B. Blankertz, K. R. Müller, D. J. Krusienski, G. Schalk, J. R. Wolpaw, A. Schlögl, G. Pfurtscheller, J. R. Millán, M. Schröder, and N. Birbaumer. The BCI competition III: validating alternative approaches to actual BCI problems. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 14(2):153–159, 2006.

[12] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2):182–197, 2002.

[13] J. Ginter Jr., K. J. Blinowska, M. Kaminski, P. J. Durka, G. Pfurtscheller, and C. Neuper. Propagation of EEG activity in the beta and gamma band during movement imagery in humans. *Methods Inf Med*, 44(1):106–113, 2005.

A Brain-Switch using Riemannian Geometry

A. Barachant[1], S. Bonnet[1], M. Congedo[2], C. Jutten[2]

[1]CEA Leti, DTBS, Minatec Campus, 38054 Grenoble, France
[2] GIPSA-lab, CNRS/UJF-INPG, Domaine Universitaire, 38402 Saint-Martin-d'Hères, France

alexandre.barachant@cea.fr

Abstract

This paper addresses the issue of asynchronous brain-switch. The detection of a specific brain pattern from the ongoing EEG activity is achieved by using the Riemannian geometry, which offers an interesting framework for EEG mental task classification, and is based on the fact that spatial covariance matrices obtained on short-time EEG segments contain all the desired information. Such a brain-switch is valuable as it is easy to set up and robust to artefacts. The performances are evaluated offline using EEG recordings collected on 6 subjects in our laboratory. The results show a good precision (Positive Predictive Value) of 92 % with a sensitivity (True Positive Rate) of 91 %.

1 Introduction

Asynchronous BCIs are of major interest in order to set-up practical BCI system in real-life conditions. Such BCI attempts to detect predefined brain pattern during operation and the user can send a command at any time. This is the most natural way but also the more difficult since one has to discriminate specific brain patterns with respect to the ongoing brain activity. A particular case of asynchronous BCI is called a brain-switch when only one brain state is to be detected in the ongoing brain activity [1]. A brain-switch could allow the user to indicate the system s/he is ready for an action [2] or simply sends a binary command very much like an electrical switch. The false rate should be kept as low as possible to let the communication system silent, when the user does not intend to communicate. Different brain-switches have been proposed based on different sources of BCI control, such as the post-imagery beta ERS [3].

Motor Imagery (MI) is a mental simulation or rehearsal of a movement without motor output, which results in the activation of dedicated cortical areas (somatosensory cortex) in given frequency bands. MI is often used in BCI applications since the brain patterns and areas involved are well known. Moreover, from the user point of view, it is also a natural way to interact. It has been observed that spatial covariance matrices are central to the detection process and that the respect of their topology using Riemannian geometry can improve the problem at hand. In this paper we propose to build a brain-switch based on direct manipulations of spatial covariance matrices using concepts of distance and geometric mean using a Riemannian metric. The classifier is first trained using a self-paced motor imagery BCI system and then evaluated offline on test recording which follows the same paradigm.

2 Methods

In the proposed approach, spatial covariance matrices are the features of interest for the brain-switch detection task and they are manipulated by doing computations within the manifold of symmetric positive-definite (SPD) matrices. The brain-switch is implemented on a sliding trial manner, where each trial is detected as a (specific) brain pattern and the result is aggregated within time to activate or not the brain-switch.

2.1 Riemannian Metric

Let $\mathbf{X} \in \mathbb{R}^{C \times N_t}$ be a short-time segment of EEG signal, which corresponds to a trial of motor imagery: C is the number of channels and N_t is the number of samples in a trial. We assume further that $\mathbf{X}$ has been band-pass filtered in a pre-processing step, and that the spatial covariance matrix of the trial is estimated by the Sample Covariance Matrix (SCM):

$$\mathbf{P} = \frac{1}{N_t - 1}\mathbf{X}\mathbf{X}^T \tag{1}$$

A key observation is that SCMs are symmetric positive-definite (SPD) matrices that belong to the manifold $P(C)$. Using Riemannian geometry, it is possible to define precisely the distance between any two SPD matrices $\mathbf{P}_1$ and $\mathbf{P}_2$ as:

$$\delta_R(\mathbf{P}_1, \mathbf{P}_2) = \left[\sum_{c=1}^{C} \log^2 \lambda_c\right]^{1/2} \tag{2}$$

with $\{\lambda_c, c = 1 \dots C\}$ the eigenvalues of $\mathbf{P}_1^{-1}\mathbf{P}_2$ [4]. This formula has some reminiscence with the common spatial pattern technique as demonstrated in [5].
If we are now given a set of M covariance matrices, it is possible to estimate the (Frechet) mean of this set using:

$$\mathfrak{G}\left(\mathbf{P}_1, \dots, \mathbf{P}_M\right) = \underset{\mathbf{P} \in P(C)}{\arg\min} \sum_{m=1}^{M} \delta_R^2\left(\mathbf{P}, \mathbf{P}_m\right) \tag{3}$$

Although no closed-form expression does exist, this (geometric) mean can be computed efficiently using an iterative algorithm [6].

2.2 Brain-Switch based on Covariance Matrices

The estimate of the covariance matrix of the band-pass filtered EEG signal in the specific condition $y = s$, (e.g., left-hand or right-hand movement imagination) can be readily obtained using the previous section as:

$$\mathfrak{G}_s = \mathfrak{G}(\mathbf{P}_i,\ i \,|y_i = s)$$

It can be expected that all these points will be clustered in a very delimited portion of the manifold $P(C)$. This allows us to define a region of interest (ROI) in the manifold based on the intra-class covariance matrix and a threshold distance. The ROI is parameterized by the equation: $\Omega = \{\mathbf{P} \mid \delta_R(\mathbf{P}, \mathfrak{G}_s) < \epsilon\}$ where $\epsilon = m + 3\,s$ with m and s denote respectively the median and standard deviation of the point-to-center distances $\{\delta_R(\mathbf{P}_i, \mathfrak{G}_s),\ i \,|y_i = s\}$.

This simple criterion is very efficient to remove large amplitude EEG artifacts from the analysis. However, it may not be not specific enough during a brain-switch operation mode. To improve the specificity of the detection, a second step consists in estimating the geometric mean $\mathfrak{G}_{\bar{s}} = \mathfrak{G}(\mathbf{P}_i,\ i \,|y_i = \bar{s}$ and $\delta_R(\mathbf{P}_i, \mathfrak{G}_s) < \epsilon)$ of the covariance matrices corresponding to the unspecific activity restricted to the computed ROI. By definition, the unspecific activity class is composed by all the activities excluding the specific one. The restriction to the ROI reduces the diversity of the unspecific activity class and thus allows the estimation of the corresponding mean covariance matrix. Again a very simple criterion is to decide whether a test trial belongs to the class of the closest mean in terms of Riemannian distance. The whole algorithm can be described as follows:

$$y = \begin{cases} s & \text{if } \delta_R(\mathbf{P}, \mathfrak{G}_s) < \epsilon \text{ and } \delta_R(\mathbf{P}, \mathfrak{G}_s) < \delta_R(\mathbf{P}, \mathfrak{G}_{\bar{s}}) \\ \bar{s} & \text{otherwise} \end{cases} \tag{4}$$

This classification output is then integrated along time to improve robustness. As a consequence, the brain-switch will be triggered ON when the specific activity is detected continuously for a given period of time T_s, typically 1 s. After being triggered, the brain-switch remains halted until the unspecific activity is detected continuously for another period of time $T_{\bar{s}}$, typically 1 s.

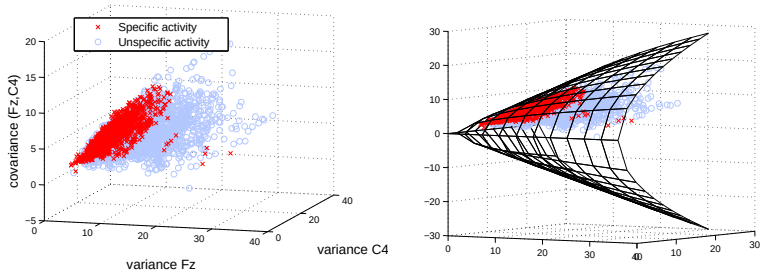


Figure 1: Covariance matrices positions for user E, channel Fz and $C4$. Training stage.

3 Results

The brain-switch precision and sensitivity is now evaluated using EEG recordings collected on 6 subjects in our laboratory. These datasets have been recorded before the development of the algorithm presented in this paper, with an adaptive implementation of spatial filtering by CSP and classification by LDA. For this reason, the results presented hereby are related to offline processing of these measurements. For each subject, 16 active electrodes (from g.tec) are disposed on the whole scalp according to the 10/20 system (Fpz F7 F3 Fz F4 F8 T7 C3 Cz C4 T8 P7 P3 Pz P4 P8). A sampling frequency of 512 Hz and a general band-pass filter between 8 and 30 Hz were used. The training stage consists in a self-paced motor imagery paradigm and lasts up to 15 min. The user is sited in front of a computer screen and a continuous visual feedback is provided. It consists in a blue vertical bar which evolves jointly with the detection output. The user strokes a key to indicate the system that s/he wants to begin a specific mental task. Each mental task is performed until the switch is triggered ON, with a maximal duration of 10 s. The algorithms are continuously adapted. The testing phase of 5 min is done on a self-paced operation mode similarly to the training stage, but without the adaptation of algorithms. It is important to notice that the unspecific activity corresponds to a concentrated resting state and does not reflect the whole diversity of the unspecific activity, which can be seen during long-term EEG recording.

Figure 1 shows the manifold of spatial covariance matrices in case of two channels. Each point represents an 2×2 SPD matrices estimated on a sliding windows of 1 second, either in specific condition (red points) or unspecific one (blue points). We note that the matrices belonging to the specific activity class lie in a compact area while the other ones are spread over a larger area. Moreover, we see that the points corresponding to the specific activity are close to the boundary of the manifold which is a cone. Such locations correspond to high correlations between the two chosen electrodes.

Figure 2 shows the class-dependent distribution of Riemannian distances from the center of the ROI during the learning stage for user E. In this example, 96.5 % of the trials corresponding to the specific activity are within the ROI while 20 % of the trials corresponding to the unspecific activity are outside the ROI. A visual inspection of the signal shows that most of rejected trials correspond to artifacts.

Figure 3 shows the classification process on the test data for user E. The threshold of detection given Equation 4 is represented by the dashed line. In this example, 85 % of unspecific activities and 72 % of specific activities have been correctly classified.

Figure 4 shows the temporal output of the brain-switch for 10 realisations of specific mental tasks of various length after the integration along time. During this period, only one event is a false positive. We can also see that the switch triggered ON at the beginning of most trials, shows the reactivity of the brain-switch.

Finally, the results for the 6 subjects are given in Table 1. The average performance is about 91 % for both PPV (Positive Predictive Value) and TPR (True Positive Rate) with 0.4 FP (False Positive) per minute, that is similar to the state of the art [3].

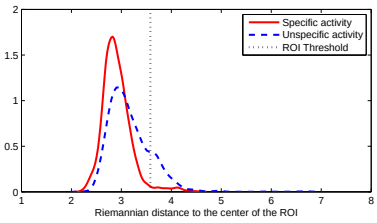


Figure 2: Distributions of the distance to the center of the ROI $\delta_R(\mathfrak{G}_s, \mathbf{P}_i)$ for the training data, user E.

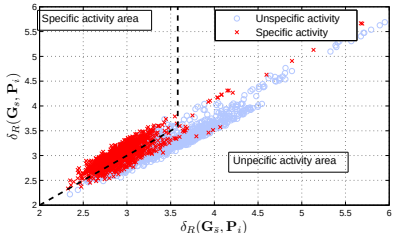


Figure 3: Classification of test data, user E.

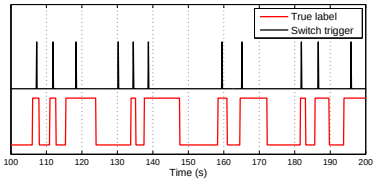


Figure 4: Temporal output of the brain-switch for test data, user E.

User	PPV	TPR	TP/min	FP/min
A	100	84	4.8	0
E	94	100	6.4	0.4
G	94	100	6.8	0.4
I	80	64	2.6	0.6
J	83	100	5.2	1
M	100	100	7.2	0
mean	**91.8**	**91.3**	**5.5**	**0.4**

Table 1: Results obtained offline for the 6 users.

4 Conclusion

The presented method is easy to set up in the EEG experiments. No parameters are needed to be set, and the large amplitude artifacts are natively handled. Offline results appear promising to go beyond the state-of-the-art. In addition, the training stage relies only on the estimation of class-related covariance matrices. The precision (PPV) could be improved by increasing the time T_s in return of losing sensitivity (TPR). Such an optimisation of timing should be done according to the application. This algorithm will be used in our next brain-switch experiments.

References

[1] S. G. Mason and G. E. Birch. A brain-controlled switch for asynchronous control applications. *IEEE Transactions on Biomedical Engineering*, 47(10):1297–1307, 2000.

[2] G. Pfurtscheller, B. Z. Allison, C. Brunner, G. Bauernfeind, T. Solis-Escalante, R. Scherer, C. Neuper T. O. Zander, G. R. Müller-Putz, and N. Birbaumer. The hybrid BCI. *Frontiers in Neuroscience*, 4, 2010.

[3] G. R. Müller-Putz, V. Kaiser, T. Solis-Escalante, and G. Pfurtscheller. Fast set-up asynchronous brain-switch based on detection of foot motor imagery in 1-channel EEG. *Medical and Biological Engineering and Computing*, 48(3):229–233, 2010.

[4] W. Förstner and B. Moonen. A metric for covariance matrices. *Technical report, Dept. of Geodesy and Geoinformatics, Stuttgart Universität*, 1999.

[5] A. Barachant, S. Bonnet, M. Congedo, and C. Jutten. Common spatial pattern revisited by Riemannian geometry. In *IEEE Workshop on Multimedia Signal Processing*, pages 472–476, 2010.

[6] A. Barachant, S. Bonnet, M. Congedo, and C. Jutten. Riemannian geometry applied to BCI classification. In *Conference on Latent Variable Analysis (LVA/ICA), St Malo*, 2010.

Restricted Boltzmann Machines as Useful Tool for Detecting Oscillatory EEG Components

D. Balderas[1], T. Zander[2], F. Bachl[2], J. Faller[1], C. Neuper[1], R. Scherer[1]

[1] Institute for Knowledge Discovery, Graz University of Technology, Austria
[2] Team PhyPA, Chair of Human-Machine Systems, Berlin Institute of Technology, Germany

d.balderassilva@tugraz.at, reinhold.scherer@tugraz.at

Abstract

Brain-Computer Interfaces (BCIs) are in need of different classification techniques, capable of having feature extraction and pattern classification. In this work we introduce Restricted Boltzmann Machines (RBMs) as a generative classification technique for BCI. It is demonstrated that RBMs are capable of inducing a powerful internal representation of spatio-temporal EEG patterns. Our results show that the emerging joint distribution of features and class labels yields expressive conditional distributions in both directions. We report initial results of the capabilities of RBMs used as classifiers.

1 Introduction

One of the challenges when designing a Brain-Computer Interface (BCI) is to correctly interpret user-intention from brain signals, requiring feature extraction and pattern classification. In particular, the separation of motor imagery patterns has been subject to a large variety of such techniques over the past years [1]. Although, many of these methods already proved to be effective, there is still need for techniques that explore the close interrelation of feature extraction and pattern classification, which are frequently designed as independent components.

A Restricted Boltzmann Machine (RBM) [2] is a graphical model that use hidden variables to derive a highly expressive marginal distribution over a given observation space. Despite their generative nature, they have been shown to be powerful in classification settings such as number recognition, text classification as well as for speech representation. The aim of this study is to evaluate the use of RBMs as an alternative classification method for BCIs that is able to automatically extract intrinsic information, from meaningful oscillatory features from the EEG.

2 Restricted Boltzmann Machines

RBMs can be understood to be a two-layer neural network of stochastic, usually binary units with symmetrically weighted connections, of *visible* units $\mathbf{v} = (v_1, \cdots, v_i)$ representing the data, and a layer of *hidden* units $\mathbf{h} = (h_1, \cdots, h_j)$, (referred to as neurons or feature extractors) to model dependencies over visible variables. In the RBM framework, the probability of the configuration $(\mathbf{v}, \mathbf{h})$ is defined by the Boltzmann distribution: $p(\mathbf{v}, \mathbf{h}; \Theta) \propto e^{-E(\mathbf{v},\mathbf{h};\Theta)}$ comprising the energy function $E(\mathbf{v}, \mathbf{h}; \Theta) = -\mathbf{h}^T\mathbf{W}\mathbf{v} - \mathbf{b}^T\mathbf{v} - \mathbf{a}^T\mathbf{h}$, with parameters $\mathbf{W}$ as the symmetric weights, $\mathbf{b}$ and $\mathbf{a}$ as the bias of the visible and hidden units respectively, and $\Theta = (\mathbf{W}, \mathbf{b}, \mathbf{a})$. The conditional distributions of the hidden variables given the visible variables and of the visible variables given the hidden are, $p(h_j = 1|\mathbf{v}) = \sigma_h(a_j + \sum_i v_i w_{ij})$ and $p(v_i = 1|\mathbf{h}) = \sigma_v(b_i + \sum_j h_j w_{ij})$ where σ called the activation functions of the respective unit. The latter can be adapted to model different kinds of units, e.g. binomial or Gaussian. However, as the hidden variables are not observed, the objective function is p(v), i.e. the joint distribution of the visible variables. Fortunately there

exists a computationally appealing algorithm called Contrastive Divergence (CD) [2] that allows to fit p(v) to a given set of observations, e.g. speech representation or in our case features extracted from EEG signals.

3 Methods

3.1 Signal Processing

This study used previously recorded EEG data [3]. The data was recorded from 10 subjects in a single training session, which consisted of 8 runs with 3 classes and 10 trials per class. The subjects had to perform 3 types of kinesthetic motor imagery tasks: left hand, right hand and foot motor imagery. One trial lasted for 8 s, the cue being present at second 3, followed by 5 s of imagery period.

Laplacian derivations were calculated over C3, C4 and CZ to obtain reference free data. The data was transformed into a time-frequency representation using the fast Fourier transform with a Hamming sliding window of 1 s and step size of 0.01 s (overlap is needed since RBM require a large number of data to be trained), using frequencies within the bandwidth 8 to 30 Hz, and centered between times 4.5 and 5.5 s, where significant information was expected [3]. The spectrum from the three channels were concatenated as single feature vectors and normalized by subtracting the vectors using the mean value over the training data set and scaled with the maximum standard deviation from all the trials. Normalization is essential to avoid saturation of the hidden units, which produces a small learning signal. The feature vectors of five out of eight runs were used as training data set, while the rest was used as the test data set. The dimensionality of the feature vector is: $space(\#channels) * frequency(\#bins) * time(\#windows)$. For the labels a Softmax representation was used with the form $\vec{y} = (1_{y=l})_{l=1}^{C}$ for C number of classes.

3.2 RBM Parametrization and Training

The training method used a double entry visible layer (class and features), to maximize the likelihood of the parameters Θ having generated a given pair of class and the corresponding features (see Figure 1). This method has previously been used for training the top layer of a deep belief network [4] or as a classifier [5]. The activation functions used for the visible layer were: Gaussian distributed units with unit variance for the real-value inputs and Softmax activation function for the labels. The hidden layer used the logistic function, with the number of units to be in the next order of magnitude (base 2) than the total number of visible units; more hidden units than visible units are needed since real-valued units normally contain more information than binary units.

The initial values taken for the training parameters (weights, biases and rates) were taken from [2] and adapted based on preliminary analysis. The RBMs was trained over 100 sweeps (or epochs) each comprising CD updates derived from 10 Gibbs sampling iterations (CD10), and taking the training parameters from the epoch with the highest classification accuracy. The training sets were composed of mini-batches of 100 randomly selected EEG trials. The weights were drawn from $N(0, 0.1^2)$ for the Gaussian-Binary connections and $N(0, 0.01^2)$ for the Softmax-Binary connections, with the different biases initialized at zero. The weights and biases were updated with a learning rate of 10^{-3} and a momentum of 0.5 with an increment gradually up to 0.7 during the learning process (higher increments made the learning unstable). A cost of $2 \cdot 10^{-4}$ is imposed since this facilitates the learning process of CD by increasing the mixing rate of the Markov chain.

4 Results

For validation we compared the classification results of our RBM implementation with those of a standard method, which used a linear discriminant analysis (LDA) classifier based on 6 logarithmic

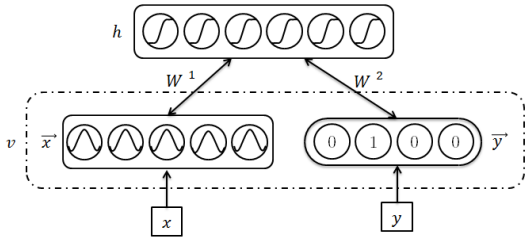


Figure 1: An RBM modeling the joint distribution of feature vector and the target classes. It shows a double entry visible layer, with x as the concatenated short Fourier transform of the three different channels, y as the classification labels; $\vec{x}$ as the normalized Gaussian representation of the features and $\vec{y}$ as the Softmax representation of the labels; h as the binary hidden units. $W1$ and $W2$ represent the bidirectional weights between the features and the hidden units, and the class label and the hidden units respectively.

		al1	al2	al3	al4	al5	al6	al7	al8	al9	al10
R vs. F	RBM	82.57	57.10	57.28	71.13	78.77	66.94	52.96	64.18	62.93	62.93
	LDA	83.47	56.89	62.77	62.48	83.12	70.10	48.13	66.81	58.37	58.28
L vs. F	RBM	76.46	54.77	52.69	76.60	78.47	66.66	55.27	68.87	69.63	66.97
	LDA	78.18	52.96	53.99	76.23	79.26	61.90	58.76	57.34	65.64	57.55
L vs. R	RBM	83.41	53.22	54.42	78.04	68.65	51.30	56.96	54.85	59.94	59.94
	LDA	80.51	49.64	64.01	79.46	60.07	55.84	52.86	48.08	51.64	51.09
L vs. R vs. F	RBM	71.16	37.14	39.37	61.04	63.88	44.20	40.87	46.32	45.27	45.33

Table 1: Accuracy results over the different classes over the test data set. The classes are shown as Right hand (R), Left hand (L) and Foot (F) motor imagery.

band-power features (2 frequency bands [10,13] and [16,24] from each of 3 Laplacian derivations). Table 1 shows the accuracy for two classes for the 10 different subjects for both RBM and LDA. It also show that RBMs can be easily extended to a higher number of classes, in our case three classes for the 10 different subjects. Figure 2 shows for a single subject how the RBM reconstructs a signal for a specific class label from its internal representation, and it is compared to the mean over the testing data set.

5 Discussion

The aim of this work was to provide a proof of concept, showing that RBMs, as generative models, have a potential as classifiers for BCIs, that can be used along current techniques and be useful for future research. Although RBMs are not yet optimized for BCIs systems they have shown potential giving accuracies comparable to the ones obtain from LDA, with mean accuracies of 64.79 % for RBM and 62.84 % for LDA.

It was shown that RBMs create an inner representation over signal vectors and class label, being able to generate a class label given a signal or generate a representation of the signal given a class label. This generative capability of RBMs is better illustrated when a class label is inserted and the signal is generated, showing similarities between the original signal and the reconstructed signal.

Noteworthy there was no need of a specific band selection to generate a representation of the oscillatory components hidden in an EEG, simplifying the amount of work on data. Also, RBMs can be easily extended to distinguish a higher number of classes, showing accuracies in our case of over 70 % for 3 classes (sig. above chance level), thing that is not easily implemented in other

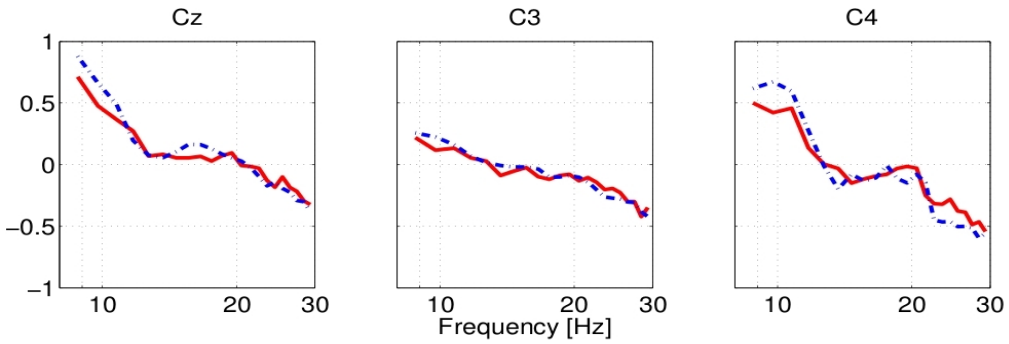


Figure 2: Example of signal reconstruction from RBMs over one subject. The dashed line is the average for a single subject over all the trials for right motor imagery, the full line is the RBM reconstructed signal of the class representing the right hand motor imagery.

methods.

Future work

We aim to use RBMs for on-line or adaptive classification, tracking the changes of the parameters and adapting their connections. Therefore a better understanding and testing of RBMs on BCIs is required. Testing different configurations like tuning of the learning rates, trying different types of activation units or increasing the number of the layers. RBMs are also interesting as feature extractors or as Boosting classifiers for other methods, providing a new classification technique that could enhance the capabilities of current BCIs classification.

Initial testing has shown that limiting the frequency bands or different normalization changes the accuracy. This shows that there is potential to improvement in the signal representation, by changing the normalization, limiting the frequency or time ranges. This has to be done carefully trying not to lose the internal representation of the signal.

Acknowledgments

This study was supported by Wings for Life- Spinal Cord Research Foundation (project 06/002) and ICT Collaborative Project BrainAble (247447).

References

[1] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi. A review of classification algorithms for EEG-based brain–computer interfaces. *Journal of Neural Engineering*, 4:R1–R13, 2007.

[2] G. E. Hinton. A Practical Guide to Training Restricted Boltzmann Machines. http://www.cs.toronto.edu/~hinton/absps/guideTR.pdf, August 2010.

[3] G. R. Müller-Putz, R. Scherer, G. Pfurtscheller, and C. Neuper. Temporal coding of brain patterns for direct limb control in humans. *Frontiers in Neuroscience*, 4, 2010.

[4] G. E. Hinton, S. Osindero, and Y. W. Teh. A fast learning algorithm for deep belief nets. *Neural Computation*, 18:1527–1554, 2006.

[5] H. Larochelle and Y. Bengio. Classification using discriminative Restricted Boltzmann Machines. In *In ICML '08: Proceedings of the 25th international conference on Machine learning. ACM*, 2008.

A New BCI Classification Method Based on EEG Sparse Representation

Y. Shin[1], S. Lee[1], M. Ahn[1], S. C. Jun[1], H. N. Lee[1]

[1]Gwangju Institute of Science and Technology,
Gwangju, Republic of Korea

{shinyh,seungchan,frerap,scjun,heungno}@gist.ac.kr

Abstract

We propose a new sparse representation based classification (SRC) method for the motor imagery based Brain Computer Interface (BCI) system. For the success of SRC method, construction of a good dictionary matrix is critical. We provide a method based on common spatial pattern (CSP). The performance of the proposed method has been evaluated over the motor imagery data sets collected from five subjects. We observe that the proposed method gives higher classification accuracy results than the linear discriminant analysis (LDA) method does, one of the well-known classification methods for BCI systems.

1 Introduction

BCI system provides a new communication and control channel between people who have severe motor disabilities and an external device without any muscle movements. Many researchers have developed BCI systems using EEG signals. EEG signals can be easily acquired from the scalp; but they are very noisy and show nonstationary characteristics. Thus, powerful preprocessing methods are used for EEG signals such as the principal component analysis (PCA), the independent component analysis (ICA) and the common spatial pattern (CSP). Widely used classification methods in the EEG based BCI field have been adopted from the pattern recognition community, including the linear discriminant analysis (LDA) and the support vector machine (SVM).

In this paper, we aim to develop a new sparse representation based classification (SRC) method for the motor imagery based BCI system. The central problem of the sparse representation theory aims to search for the most compact representation of a given signal in terms of linear combination of atoms in an overcomplete dictionary [1]. When compact description of seemingly complex signals becomes possible, it can be used for a number of applications including noise reduction, dimensionality reduction, and pattern recognition. The sparse representation is achieved by the ell-1 minimization which is a suboptimal but closely attains the sparsest solution within a polynomial time if enough number of observation samples have been obtained. This ell-1 minimization approach can be utilized as a sparse representation tool in this paper for the classification purpose. For the success of this SRC method, construction of a good dictionary matrix is critical. We provide a method based on the CSP.

2 Methods

2.1 Experiment

In this study, we use data sets of BCI Competition III (Data set IVa) [2] which were recorded from five subjects. Subjects have taken the same procedure of a BCI experiment in which there are three classes, left and right hand, right foot of motor imagery movements. However, only

data corresponding to right hand 'R' and right foot 'F' were provided for analysis. The data recording was made using BrainAmp amplifiers and a 128 channel Ag/AgCl electrode cap from ECI. 118 EEG channels were measured at positions of the extended international 10/20-system. Signals were band-pass filtered between 0.05 and 200 Hz and then digitized at 1000 Hz. For off line analysis signals were downsampled to 100 Hz.

2.2 Preprocessing

We take a data segmentation for following analysis. We use 1000∼2000 ms of signal samples (100 samples) after the Cue has been presented. Next, to eliminate the noise that is not related with sensorimotor rhythms (SMRs), we use a band-pass filter with 8∼15 Hz cut off frequency.

To reduce the dimension of feature vector and make distinguishable features, we use the CSP method. CSP is a powerful signal processing technique that has been successfully applied for EEG-based BCIs [3].

Let $\mathbf{X} \in \mathbb{R}^{C\times T}$ be a segment of EEG signals where C is the number of EEG channels. In this study, C is 118, and T is the number of sampled time points collected in all the trials. We use 100 samples (one second). We have two classes of EEG training trials $\mathbf{X}_R \in \mathbb{R}^{C\times T}$ and $\mathbf{X}_F \in \mathbb{R}^{C\times T}$ each corresponding to the Right hand and Foot movement. Using the CSP method, we can obtain the CSP filters $\mathbf{W} \in \mathbb{R}^{C\times C}$. We call each column vector $\mathbf{w}_i \in \mathbb{R}^C (i = 1, 2, ..., C)$ of $\mathbf{W}$ a spatial filter. Among them, we use nCSP filters from the front and another set from the back. Then, we can make this as the CSP filtering matrix $\overline{\mathbf{W}} \in \mathbb{R}^{C\times 2n}$, i.e., $\overline{\mathbf{W}} := [\mathbf{w}_1, ..., \mathbf{w}_n, \mathbf{w}_{C-n+1}, \mathbf{w}_C]$. Given the two classes of EEG training signals, we can define the CSP filtered signals, i.e.,

$$\overline{\mathbf{X}}_R \in \mathbb{R}^{2n\times T} := \overline{\mathbf{W}}^{\mathrm{T}}\mathbf{X}_R, \quad \overline{\mathbf{X}}_F \in \mathbb{R}^{2n\times T} := \overline{\mathbf{W}}^{\mathrm{T}}\mathbf{X}_F \tag{1}$$

Next, we compute FFT of each set of $T = 100$ samples. Because the sampling rate is 100 samples/sec, the frequency resolution is 1 Hz. We then compute the frequency power of the Mu band (8∼15 Hz). Thus, there are $N_f = 8$ columns in each matrix. They are:

$$\overline{\mathbf{X}}_R(f) \in \mathbb{R}^{2n\times N_f}, \quad \overline{\mathbf{X}}_F(f) \in \mathbb{R}^{2n\times N_f} \tag{2}$$

2.3 Linear Sparse Representation Model

Let N_t be the number of total training signals for each class $i = R, F$. We define the dictionary matrix $\mathbf{A}_i = [\mathbf{a}_{i,1}, \mathbf{a}_{i,2}, ..., \mathbf{a}_{i,N_t}]$ for $i = R, F$ where each column vector $\mathbf{a} \in \mathbb{R}^{m\times 1}$ having dimension $m = 2n\times N_f$ is obtained by concatenating the 2n rows of $\overline{\mathbf{X}}_R(f)$ and taking the transpose. Let's call this vectorization. The same procedure is repeated for the right foot part, $\overline{\mathbf{X}}_F(f)$. By combining the two matrices, we form the complete dictionary, $\mathbf{A} := [\mathbf{A}_R; \mathbf{A}_F]$. Thus, the dimension of $\mathbf{A}$ is $m \times 2N_t$. We apply the same procedure done to obtain the columns of the dictionary to the test signal. That is, the test signal is transformed to a vector $\mathbf{b} \in \mathbb{R}^{m\times 1}$ through the CSP filtering, FFT, and vectorization. Thus, the dimesion of $\mathbf{b}$ is the same as the dimension of the columns of the dictionary $\mathbf{A}$. Then, this test signal $\mathbf{b}$ can be sparsely represented as a linear combination of some columns of $\mathbf{A}$:

$$\mathbf{b} = \sum_{i=R,F} x_{i,1}\mathbf{a}_{i,1} + x_{i,2}\mathbf{a}_{i,2} + \cdots + x_{i,n_t}\mathbf{a}_{i,N_t} \tag{3}$$

where $x_{i,j} \in \mathbb{R}, j = 1, 2, ..., N_t$ are scalar coefficients. Then, we can represent this as a matrix algebraic form:

$$\mathbf{b} = \mathbf{A}\mathbf{x} \tag{4}$$

where $\mathbf{x} = [x_{R,1}, x_{R,2}, ..., x_{R,N_t} x_{F,1}, x_{F,2}, ..., x_{F,N_t}]^T \in \mathbb{R}^{2\cdot N_t}$.

# of CSP filter	al	ay	aw	aa	av
2×1	96.64	89.64	80.35	85.71	60.71
2×2	96.42	90.71	86.78	85.35	66.42
2×3	97.85	93.57	86.07	87.14	68.57
2×4	97.85	95.71	90.00	92.50	67.50
2×5	98.92	95.71	92.14	92.14	71.07
2×6	97.50	95.71	89.64	90.71	77.50

Table 1: Classification accuracy of SRC with different number of CSP filters.

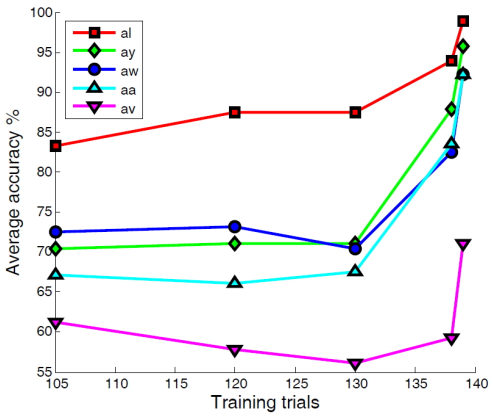


Figure 1: Average accuracy of SRC with different number of training signals.

2.4 Sparse Representation by ell-1 Minimization

We have the number of total training signals $2N_t$ which is larger than the number of frequency power ($m = 2n \times N_f$). The more trials, the larger the dictionary, and thus the better the sparse representation result will be. The column size is larger than the row size of the dictionary $\mathbf{A}$. Thus, the linear Equation 4 is under-determined ($m < 2N_t$). Recent studies in the Compressed Sensing theory have shown that the ell-1 norm minimization, given below, can solve this under-determined system well in polynomial time [4]: $\min \|\mathbf{x}\|_1$ subject to $\mathbf{b} = \mathbf{A}\mathbf{x}$. In this paper, we use the basis pursuit method, one of the standard linear programming methods.

2.5 Sparse Representation based Classification

After solving the ell-1 minimization problem, the nonzero elements of $\mathbf{x}$ must be corresponding to the column of class i. Because the EEG signals are very noisy, the nonzero elements may appear in the indices corresponding to the column of another class. To make use of the sparse representation result, in a classification problem, we introduce the characteristic function δ [5]. For each class i, we define its characteristic function $\delta_i : \mathbb{R}^{2N_t} \rightarrow \mathbb{R}^{2N_t}$ which selects the coefficients associated with class i. For $\mathbf{x} \in \mathbb{R}^{2N_t}$, $\delta_i(\mathbf{x}) \in \mathbb{R}^{2N_t}$ is a new vector which is obtained by nulling all the elements of $\mathbf{x}$ that are associated with the other class. Then we can obtain the residuals $r_i(\mathbf{b}) := \|\mathbf{b} - \mathbf{A}\delta_i(\mathbf{x})\|_2$ for R and F. Then, the classification rule is given by: $\text{class}(\mathbf{b}) = \arg\min_i r_i(\mathbf{b})$. Thus, we determine the class i that has the minimum residuals.

3 Results and Discussion

We have analyzed five data sets, which have the same 140 trials for each class. We use the statistical k-fold cross-validation method to evaluate the average performance of the classifiers. Table 1 shows the classification accuracy of SRC with different number of CSP filters. Here, we

Classification Method	al	ay	aw	aa	av	Mean	Std.
LDA	92.50	92.50	91.07	91.07	63.92	86.21	12.48
SRC	98.92	95.71	92.14	92.14	71.07	90.00	10.95

$* \ P = 0.0428 < 0.05$ (Using paired T-test)

Table 2: Classification accuracy of SRC and LDA.

use 139 training signals for each class (leave-one-out). From these results, we choose the most important 10 (2×5) CSP filters among 118.

Figure 1 shows the classification accuracy of SRC method when we use the 10 CSP filters. The SRC method shows good performance, especially when the number of training signals is large enough. We compare the classification accuracies of the LDA and the SRC method. Table 2 shows the comparison results when we use the leave-one-out method. In the LDA classification, we also use the CSP filtering, the FFT to extract features. For all subjects, we use the same 8∼15 Hz frequency power, number of CSP filters (10) and time sample (1000∼2000 ms). The results indicate that the accuracy of the SRC is better than that of the LDA method. It is statistically improved with the level of p which is less than 0.05, however, more datasets will be helpful for solid conclusion and it is currently under the investigation.

4 Conclusion

We have applied the idea of sparse representation as a new classification method to the motor imagery based BCI. The sparse representation method needs a well-designed dictionary matrix constructed from the training data. We use the CSP filtering and the FFT to produce the columns of the dictionary matrix. We have evaluated the proposed method over the BCI competition data sets. We observe that the performance of the proposed method is very satisfactory when the number of training signals is large enough. We have compared with the conventional approach such as LDA method, which is well known for the BCI system. For all subjects, our proposed method shows better classification accuracy than LDA method.

5 Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (Do-Yak Research Program, No. 2010-0017944) and NRF grant (No. 2010-0006135).

References

[1] K. Huang and S. Aviyente. Sparse representation for signal classification. *Neural Information Processing Systems*, 2006.

[2] B. Blankertz. Berlin brain-computer interface. *http://www.bbci.de/*.

[3] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K. R. Müller. Optimizing spatial filters for robust EEG single-trial analysis. *IEEE Signal Proc. Magazine*, 25(1):41–56, 2008.

[4] E. Candes, J. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate measurements. *Comm. Pure and Applied Math.*, 59(8):1207–1223, 2006.

[5] J. Wright, A. Y. Yang, A. Ganesh, S. Shankar Sastry, and Y. Ma. Robust face recognition via sparse representation. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 31(2):210–227, 2009.

A Transient-VEP Based Spelling System by Using ICA and Adaptive Morphological Filter

K. Inoue[1], T. Yamaguchi[1], T. Mizoguchi[1], M. Fujio[1], M. Maeda[1]

[1]Department of Systems Design and Informatics, Kyushu Institute of Technology,
680-4 Kawazu, Iizuka, Fukuoka, 820-8502, Japan

inoue@ces.kyutech.ac.jp

Abstract

Recently, a lot of researches have studied the brain computer interface (BCI) system, and some systems have already been put to practical use. For instance, the system based on P300 were actively used. However the P300 system requires user's strong concentration. Therefore it is necessary to introduce another information from the EEG signal. In this study, we tried to construct transient-VEP spelling system which uses response by visual stimulus. The influence from another source of a light was removed by independent component analysis, and VEP component was extracted with a morphological filter. The classification accuracy among four classes were about 85 % on the average by using the correlation coefficient between the template and the EEG response for a period of 22.5 sec.

1 Introduction

The electroencephalogram (EEG) signals can be used to move a cursor on a computer screen. Such an EEG-based brain computer interface (BCI) can provide a new communication channel to replace an impaired motor function. In this system, EEG includes the information (e.g. ERD, ERS and P300) related to a cognitive task or a motor imagery, which is used to estimate an behavioral intention.

Many recognition methods of EEG signal have been proposed. Pfurtscheller et al extracted ERD/ERS from EEG over motor imagery tasks.They used it for the pattern recognition for single-trial wave form using a neural network [1]. We also have studied the fail safe system for BCI using error potential P300 [2]. Guger et al developed a spelling device using P300 feature [3]. This kind of spelling system was carefully-studied in recent years. Generaly, user can't use the system long time, since the system requires user's strong concentration. BCI systems based on visual evoked potential (VEP) were also studied. The system based on steady-state VEP is known to a good distinction performance [4]. This system has danger of causing a photosensitive epilepsy, and gives an unpleasant feeling to some users. Bin et al. [5] developed the BCI system based on code modulation VEP. This system is efficient as the spelling system. But, all pixels on the screen always blinks. Therefore, it is not suitable to BCI system in which only selected pixels on the screen blinks.

In this study, we tried to construct EEG typewriter based on visual evoked potential (VEP). The transient-VEP was chosen for stress-free spelling system and independent component analysis (ICA) was introduced to extract source components of VEP. In addition, VEP components were extracted using mathematical morphology that has nonlinear characteristics [6]. These methods were designed by using information of correlation between the extracted signal and template signal. Effectiveness of our proposed method was confirmed through experimental studies. As the results, we found out that BCI system using transient-VEP was available.

2 Experimental Paradigm

The VEP spelling system is constructed by using EEG response influenced from visual stimulations of several patterns. Four kinds of characters are displayed in the monitor (Figure 1), and subject is instructed to gaze a character in this experiment. Each character flashes in a different patterns

with cycle of 1.5 sec. respectively. A cycle is composed of stimulation and non-stimulation pattern of $N_b = 15$ bits. The character lights during 50 ms. at the stimulation. This experiment was continued 2640 cycles (120 cycles $\times$ 22 set) a day. The monitor (ProLite E2607WS, iiyama) is from the distance of about 100 cm in front of a subject. Five subjects (S1, S2, S3, S4, S5) participated in this experiment for three days. EEG signals were measured from 17 channels (Figure 2) with 512 Hz sampling frequency (g.USBamp, g.tec).

3 Signal Analysis and Pattern Recognition Method

If there is another optical source besides the gaze point, the source influences VEP related to the gaze character (Figure 3). These influences appear in ipsilateral side as stimulation in occipital region. Here, ICA is used for preprocessing of all channels (17 channels). Additionally, VEP components are extracted by using morphological filter (M-filter) from decomposed signal. Finally, the gaze character is distinguished by using correlation coefficient between template and these features.

3.1 Mathematical Morphology

In this section, mathematical morphology is described. As basic operations, we employ Minkowski addition $\oplus$ and Minkowski subtraction $\ominus$, those are respectively defined as follows.

$$[f \oplus g](t) = \max_{u \in G, t-u \in F} \{f(t-u) + g(u)\}, \quad [f \ominus g](t) = \min_{u \in G} \{f(t-u) - g(u)\} \tag{1}$$

where, $f(t)$ is an original signal and $g(t)$ is the structuring function which characterizes the filter. Furthermore, F and G denote the domains of $f(t)$ and $g(t)$, respectively. Conventionally, it is assumed that every signal takes the value $-\infty$ out of its domain.

Morphological filters are composed of these combinations. These filters behave nonlinear low, high, band pass and band stop filter (Figure 4).

$$\text{low pass} : \psi_g(t) = \frac{1}{2}\left((f_g)^g(t) + (f^g)_g(t)\right) \tag{2}$$

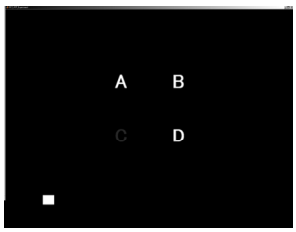
$$\text{high pass} : \omega_g(t) = f(t) - \psi_g(t) \tag{3}$$

$$\text{band pass} : p_{g_m,g_n}(t) = \psi_{g_m}(t) - \psi_{g_n}(t) \tag{4}$$

$$\text{band stop} : c_{g_m,g_n}(t) = f(t) - \psi_{g_m}(t) + \psi_{g_n}(t) \tag{5}$$

where, g_m and g_n are structuring functions of different shape and $f_g = [(f \ominus g^s) \oplus g]$, $f^g = [(f \oplus g^s) \ominus g]$ and $g^s(t) := g(-t)$.

The characteristics of filter is decided by the shape of structuring function. In this study, the structuring function SuperEllipse type is defined with a variable parameter k, width of window



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	1	1	0	1	1	0	0	1	0	1	0	0	0	0	1
B	0	1	1	0	1	0	1	1	1	1	0	0	0	1	0
C	1	0	1	1	0	0	1	0	1	0	0	0	0	1	1
D	1	1	0	1	0	1	1	1	1	0	0	0	1	0	0

1: stimulus, 0: non-stimulus

Figure 1: Stimulus Patterns.

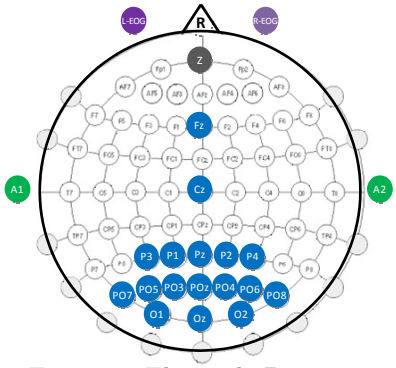


Figure 2: Electrode Positions.

$2 \times n$ and curvature parameter p as follows.

$$\text{SuperEllipse}: \quad g(t) = k\left(1 - \left|\frac{x}{n}\right|^p\right)^{\frac{1}{p}} \qquad (-n \le t \le n) \tag{6}$$

3.2 Pattern Recognition Method

ICA is preprocessed for 17 channels. The signal corresponding to the gazing point is chosen by using correlation coefficient between EEG response and template ($\text{cor}(\mathbf{x},\ \mathbf{x}_{\text{TEMP}})$). $\mathbf{x}_{\text{TEMP}}$ is composed by using average signal at single source of light.

$$i^* = \arg\max_i\{|r(i)|\}, \quad r(i) = \sum_{j=1}^{4}\sum_{k=1}^{N_b(=15)} e_j(k)\cdot \text{cor}\left(\mathbf{x}(i,k), \mathbf{x}_{\text{TEMP}}\right), \quad (i = 1, 2, \cdots, 5) \tag{7}$$

where, e_j denotes a stimulation pattern of class j (stimulus: 1, non-stimulus: -1). The roles of selected signals are fixed by using $\text{sign}(r(i))$.

The structuring function of morphological filter is constructed for extracting VEP component from EEG signal. In this study, we used morphological band stop filter whose characteristic is decided by two structuring functions (6 parameters). These parameters are selected by using genetic algorithm by the evaluation function as follows.

$$\theta^* = \arg\max_\theta\{u(\theta)\}, \quad u(\theta) = \sum_{j=1}^{4}\sum_{k=1}^{15} e_j(k)\cdot \text{cor}\left(c_{g_1,g_2}, \mathbf{x}_{\text{TEMP}}\right) \tag{8}$$

where, $\theta = [k_1, n_1, p_1, k_2, n_2, p_2]^T$.

The features for pattern recognition are selected to correlation coefficient between template wave and EEG response of 1 bit period (100 ms). Bayes classifier is adopted as the pattern recognition method with assumption that feature value $\mathbf{s}$ is random variable with normal distribution. The discrimination function is as follows.

$$i^* = \arg\max_i \Pr(\omega_i|\mathbf{s}) = \arg\max_i\left\{\frac{1}{2}\ln|\Sigma_i| - \Pr(\omega_i) + \frac{1}{2}(\mathbf{s}-\mathbf{m}_i)^T{\Sigma_i}^{-1}(\mathbf{s}-\mathbf{m}_i)\right\} \tag{9}$$

$$\mathbf{s} = [s_1, s_2, s_3, s_4]^T, \quad s_j = \sum_{k=1}^{15} e_j(k)\cdot \text{cor}\left(\mathbf{y}(k), \mathbf{x}_{\text{TEMP}}\right)$$

4 Results

Figure 5 shows the result of processing of ICA and M-filter. Using ICA, VEP component was extracted, although it is hard to extract VEP component from the observed original signal. Figure 6 shows the distributions of 2 features in 4 features averaged during 7.5 sec (5 times 15 bits).

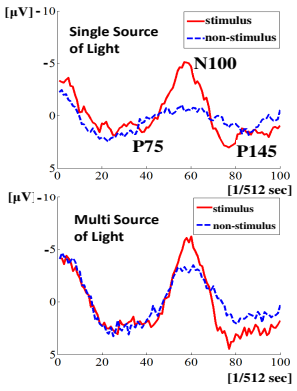


Figure 3: Averaged Signal.

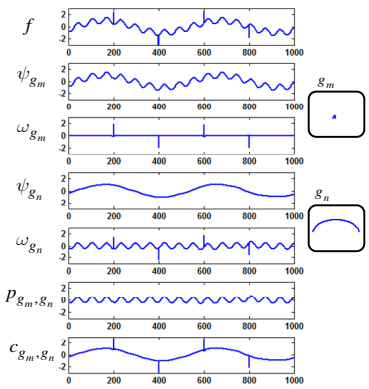


Figure 4: M-Filter.

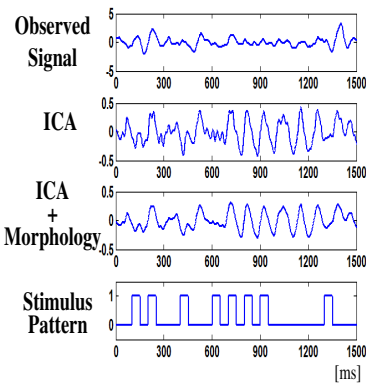


Figure 5: Processing Result.

The pattern recognition seems to become easily by using ICA and M-filter. Figure 7 shows the relations between the accuracy and the number of additions. It is confirmed that the accuracy improved as the number of additions is increased, although too many additions is not practicable. The averaged classification accuracy among 4 classes of 5 subjects was about 49 % and 74 % in case of EEG signals of one series (1.5 sec) and 5 series (7.5 sec), respectively. The classification accuracy of subject S4 was about 87 % when it used the EEG signals of 15 series. The number of addition that the averaged accuracies over all subjects achieves 80 % was 11 on ICA and was 9 on ICA+M-filter. The M-filter was useful for the accuracy improvement. On the other hand, after the stimulations appeared during 4 successive periods, the response like VEP appeared in the following periods without stimulation for all subjects. These phenomenon is similar to the response of steady-state VEP. There was a possibility of changing subject's internal state. Therefore, it is necessary to improve the experimental system or methods in order to avoid such phenomenon.

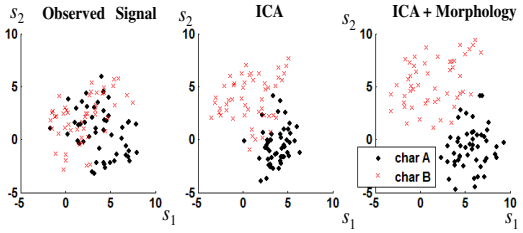


Figure 6: Distribution of Feature (Subject: S4, number of addition: 5 (7.5 sec)).

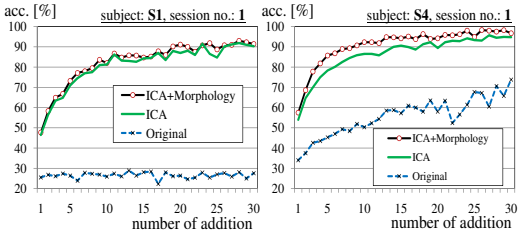


Figure 7: Pattern Recognition Results with number of addition.

5 Conclusion

In this paper, we tried to construct EEG typewriter system based on transient-VEP by applying ICA and morphological filter. By using information of correlation between template signal and processed signal for designing these methods, it was confirmed that VEP features was able to be extracted more clearly and the gaze point detection system using VEP was constructable. A BCI system with VEP will be able to be achieved by advancing this research further.

References

[1] G. Pfurtscheller, C. Brunner, A. Schlögl, 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.

[2] K. Inoue, K. Kumamaru, and G. Pfurtscheller. Robot operation based on pattern recognition of EEG signals. *Proc. of the 3rd International Brain-Computer Interface Workshop and Training Course 2006*, pages 116–117, 2006.

[3] C. Guger, S. Daban, E. Sellers, C. Holzner, G. Krausz, R. Carabalona, F. Gramatica, and G. Edlinger. How many people are able to control a P300-based brain-computer interface (BCI)? *Neuroscience Letters*, 462:94–98, 2009.

[4] P. Martinez, H. Bakardjian, and A. Cichocki. Fully online multicommand brain-computer interface with visual neurofeedback using SSVEP paradigm. *Computer Intelligence and Neuroscience*, 2007.

[5] G. Bin, X. Gao, Y. Wand, Y. Li, B. Hong, and S. Gao. A high-speed BCI based on code modulation VEP. *Journal of Neural Engineering*, 8(2), 2011.

[6] J. Serra. *Image Analysis and Mathematical Morphology*. U.K. Academic, London, 1994.

Largest Lyapunov Exponent Extraction for EEG-Based Brain-Computer Interface

P. Belluomo[1], M. Bucolo[1], L. Fortuna[1]

[1]Dipartimento di Ingegneria Elettrica, Elettronica e Informatica,University of Catania,Italy

Abstract

Signal processing and classification methods are essential tools in the improvement of Brain Computer Interface (BCI) technology. In this paper a new approach based on nonlinear time series analysis to extract EEG signals features is proposed. In particular a fast algorithm that computes the largest Lyapunov exponent was used. This signal processing approach was tested off line considering three sessions of imaginary motor tasks. The results obtained reveal the capability and the potentiality of this method in respect to the classical approach.

1 Introduction

Different signal processing algorithms have been used in BCI [1] in order to enhance their performance. The signal processing schemes include pre-processing, features extraction, feature selection/dimensionality reduction, features classification and post-processing blocks, as Bashashati indicates in his review [2]. With regard to the features extraction about one-third of BCI designs have used power-spectral features because of their ease of application, computational speed and direct interpretation of the results. The recent progresses in the theory of nonlinear dynamics and complex systems mathematics provide new methods for the study of brain signals [3]. Nonlinear time series analysis focuses on methodologies that distinguish chaotic signals from noise, and how properties of chaos can be used to model and classify dynamical systems [4]. Since the first nonlinear EEG studies [3] several steps forward have been done, and the nonlinear analysis of EEG has been widely used for diagnosis of neurovegetative pathologies. For example in Alzheimer's Disease a lower value of correlation dimension, maximum Lyapunov exponent and sample entropy, than in healthy control subjects have been found. In addition, Daly analyzed the performance of the Wackermann parameters in the classification of single-trial ERP responses [5] for BCI. The main issue related to the use of nonlinear time series analysis for EEG signals is the high computational time often involved in these procedures, so not easily adaptable to the BCI requirements. In this paper a nonlinear features extraction method based on the Divergence Algorithm(DivA) that has shown very good performance both in the parameter evaluation and computation time is considered [6]. The proposed method has been already applied to MEG signals in order to characterize the occurrence of spatiotemporal nonlinear patterns [4]. The largest Lyapunov exponent (λ) was evaluated during a left and right hand imagery movement task. The DivA and the experiment design are presented in Section 2 and 3 while the results of the offline analysis are discussed in Section 4.

2 Divergence Algorithm

Signals generated by a nonlinear systems, such as the brain, present irregular and continuous waveforms in time and in frequency domains [3]. In this context the largest Lyapunov exponent (λ) is representative of the sensitivity to initial condition (stretching phase) of a given dynamic, being positive for chaotic behaviors. The relationship between λ and the divergence of j^{th} pairs

Dataset 1 (14 electrodes)	Dataset 2 (16 electrodes)	Dataset 3 (16 electrodes)
1-Training session (60 sec)	1-Training session (60 sec)	1-Training session (60 sec)
2-Training session (60 sec)	2-Training session (60 sec)	2-Training session (60 sec)
3-Training session (60 sec)	3-Training session (60 sec)	3-Training session (60 sec)

Table 1: Experimental data sets over a three-day time period.

of system trajectories starting from nearest initial conditions can be represented by the following relation:

$$d_j(i) \approx C_j e^{\lambda(i\Delta t)} \tag{1}$$

where C_j is the initial separation between trajectories, i is the discrete time step and Δt the sampling period of the time series. Equation 1 is valid for a finite number of time-steps after that the diverging curve saturates since the effects of *folding process* take place keeping all the trajectories bounded. λ is estimated through the finite size Lyapunov exponent method [7]:

$$\lambda(\delta, \Delta) = \left\langle \frac{1}{T(\delta, \Delta)} \right\rangle ln\left(\frac{\delta}{\Delta}\right) \tag{2}$$

where δ is the initial error and Δ our tolerance, while T is the predictability time.

For our experiment DivA, an alternative methodology for the extraction of the asymptotic distance and of the λ was used [6]. This implementation results to be computationally less onerous than the conventional ones, since it is not based on the time-delay embedding concept and also no intermediate computational steps are needed to obtain the final result being particularly suitable for realtime analysis. A set of starting points is found: X = $x_0^{(i)}$ where i = 0,...n. Each point $x_0^{(i)}$ identifies a trajectory made up of all the time series samples in the range: [$x_0^{(i)}$, $x_0^{(i)}$ + *lenghtTrj*-1], where *lenghtTrj* is the length of the trajectories chosen in a way that, when the distance between them is computed. The d_j is computed averaging all the trajectory differences and the largest Lyapunov exponent can be extracted as the initial slope of this curve. The slope is simply computed considering the straight line joining the first occurrence of the lowest value with the first value settle around the asymptotic value of d_j.

3 EEG Tasks

For our experiment the user, without any muscular involvement, modifies his/her neuronal activity in the primary sensory-motor areas performing a motor imagery task. During the training session, the computer screen is either blank, or displaying an arrow pointing left or right. The two different stimuli appear for several times in a random sequence. Depending on the direction of the arrow, the subject is instructed to imagine a movement of the left hand or of the right hand. If the screen was blank the user is instructed to have a rest. The same task was performed three times in three different days by an healthy right-handed subject (female, aged 28 years) who wore a EEG cap with integrated electrodes. The EEG potentials were recorded using two channels configuration: for the first session the signals was acquired at 14 locations (FC5, FC1, FC2, FC6, C3, Cz, C4, CP5, CP1, CP2, CP6, T7, Pz and T8) and for the second and third sessions at 16 locations (replacing T7 and T8 with P3, P4, F3 and F4) sites in the the standard 10-20 System and digitized at 2000 Hz. This change has been decided to have augmentative information after a first view of the results obtained for the first trial.

4 Discussion

The signals acquired during all the training sessions were filtered in six different bands: delta (1–4 Hz), theta (4–8 Hz), alpha (8–12 Hz), beta I (12–16 Hz), beta II (16–20 Hz), and gamma (20–49 Hz). In each band and for all the channels the power and λ were computed. Then, the

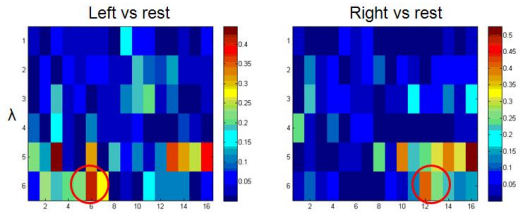


Figure 1: Space-frequency map of the correlation coefficient for λ, the fist row reports the LA versus rest and the second row RA versus rest.

correlation coefficients [8] associated with the two conditions, Left (LA) and Right (RA) arms imaginary movement versus rest phase (RP) were evaluated.
As it is expected from previous well-known results, according to the International 10-20 System, the evolution of the power in gamma band, reveals that imagination of hands or arms movement cause a radial current flow in the sensory-motor area close to the positions C3 or C4, respectively associated with the right and the left arms [9]. Figure 1 plots the space-frequency map of the correlation coefficient (r) for λ, it is possible to distinguish in the gamma-band a higher level of sensitivity during the LA imaginary activity in the electrode $C4(6)$ and during the RA in $C3(13)$. The same conclusion can be drawn considering Figure 2 (on the left) where for each dataset the r value for both power (blue line) and λ (red line) in gamma band versus the channels have been taken into account. In addition Figure 2 (on the right) shows for both parameters the head maps of the r-value averaging the dataset 2 and 3. The first row of pictures is referred to the LA activity and the second to the RA.
These results could lead to investigate on the potentiality of this new parameter to drive a BCI system. In this direction first of all a series of statistical comparisons were performed, to confirm the attitude of this feature to characterize LA-RA activity:

-between λ in $C4$ during LA and RA ($n1 = n2 = 45$, $t = 4.1$ and $p = 0.0071\,\%$);
-between λ in $C3$ during LA and RA ($n1 = n2 = 45$, $t = -2.9$ and $p = 0.42\,\%$);
-between λ in $C4$ and in all the other channels during the LA ($n1 = 45$, $n2 = 645$, $t = 3.8$ and $p = 0.012\,\%$);
-between λ in $C3$ and in all the other channels during the RA ($n1 = 45$, $n2 = 645$, $t = 1.7$ and $p = 8\,\%$);

where t is the t-test value, n1 and n2 are the samples size, and p is the probability of observing a value as extreme or more extreme than the obtained t-value. The rejection of the null hypothesis at the 5 % significance level has been confirmed in the first three cases, meanwhile in the last comparison is required the 8 %. In this contest considering the size of the sample n1 and n2 a power of about 60 % can be assumed.

From these results it is possible to confirm the previous analysis and further more can be enhanced the attitudes of the subject investigated to an higher level of brain wave control during the LA than the RA. For both tests involving channel $C4$ bigger t values were obtained.

To point out on the parameter sensitivity the indicator D = max_{left} - max_{right} has been evaluated, where max_{left} and max_{right} are the maximum r value of the signal recorded from the electrodes that cover respectively the left and right motor cortex area. Figure 2 shows that the value of D for λ is higher than for the power or at least comparable. This results highlight the potentiality of λ to be used as more robust feature for BCI. In future encouraged from these preliminary results an extended case study will be designed taking into account more subjects, and higher precision methods for classification will be applied. Once the valence of this approach will be proved, to move toward the on-line implementation another important aspect that has to be faced is the algorithmic complexity. Up to now the proposed algorithm for the evaluation of the *lambda* has not been fully-optimized but the low computational time of about 30 sec involved in the process one complete dataset (on a Quad 1.58 GHz PC) is a promising beginning.

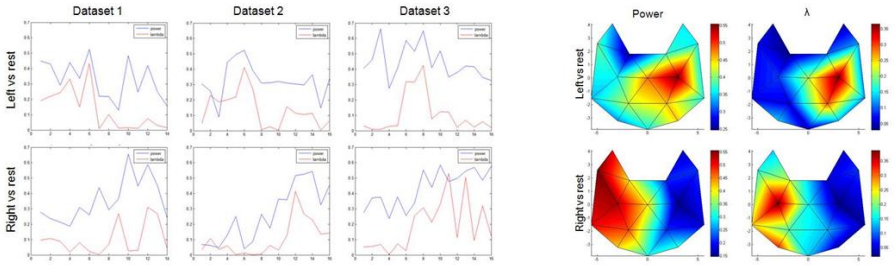


Figure 2: On the left: the correlation coefficient for power (blue line) and λ (red line) in the gamma band for each experimental datasets. Channels from 1 to 7 are related to the right hemisphere and channels from 10 to 16 to the left hemisphere. On the right: head map distribution of the r value for the second and third dataset.

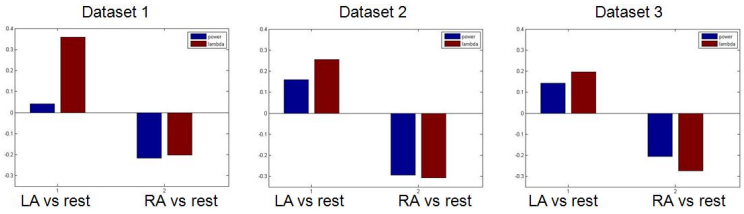


Figure 3: D value computed for the power (blue) and λ (red) during LA and RA for all the dataset.

References

[1] J. R. Wolpaw, N. Birbaumer, W. J. Heetderks, D. J. McFarland, P. H. Peckham, G. Schalk, E. Donchin, L. A. Quatrano, C. J. Robinson, and T. M. Vaughan. Brain-computer interface technology: a review of the first international meeting. *IEEE Transactions on Rehabilitation Engineering*, 8:164–173, 2000.

[2] A. Bashashati, M. Fatourechi, R. K. Ward, and G. E. Birch. A survey of signal processing algorithms in brain-computer interfaces based on electrical brain signals. *Journal of Neural Engineering*, 4:32–57, 2007.

[3] H. Kantz and T. Schreiber. Nonlinear time series analysis. *Cambridge*, 2004.

[4] F. Sapuppo, E. Umana, M. Frasca, M. La Rosa, D. Shannahoff-Khalsa, L. Fortuna, and M. Bucolo. Complex spatio-temporal feature in MEG data. *Mathematical Biosciences and Engineering*, 3(4):697–716, 2006.

[5] I. Daly, N. Williams, S. J. Nasuto, K. Warwick, and D. Saddy. Single trial BCI operation via wackermann parameters. *IEEE Intern. Workshop, MLSP2010*.

[6] M. Bucolo, F. Di Grazia, F. Sapuppo, and M. C. Virzí. A new approach for nonlinear time series characterization, "DivA". *The 16th Medit. Conf. on Control and Automation (MED2008)Ajaccio*, 2008.

[7] E. Aurell, G. Boffetta, A. Crisanti, G. Paladin, and A. Vulpiani. Predictability in the large: an extension of the concept of Lyapunov exponent. *Journal of Physics A*, 30:1–26, 1997.

[8] T. H. Wonnacott and R. Wonnacott. Introductory Statistics. *John Wiley and Sons, New York*, 3 edition, 1977.

[9] R. Neshige, H. Lüders, and H. Shibasaki. Recording of movement-related potentials from scalp and cortex in man. *Brain*, 111:719–736, 1988.

Towards an EEG-Based BCI Controlled by Expectation

S. Kanoh[1,2], K. Miyamoto[2], T. Yoshinobu[2,3]

[1]Department of Electronics and Intelligent Systems, Faculty of Engineering,
Tohoku Institute of Technology, Sendai, Japan
[2]Graduate School of Engineering, Tohoku University, Sendai, Japan
[3]Graduate School of Biomedical Engineering, Tohoku University, Sendai, Japan

kanoh@tohtech.ac.jp

Abstract

An EEG-based BCI (brain-computer interface) system which controls the movement of computer cursor only by user's mental expectation was proposed and investigated. The cursor moved at a fixed speed to one out of four directions (up, down, left and right), and the direction was changed to another after a random time interval. Subjects were requested to gaze at the cursor and expect that the cursor moves to the direction instructed by the experimenter. The event-related potentials elicited by the sudden change of the direction of cursor movement were extracted, and were classified by LDA (linear discriminant analysis) to determine the direction to which the subject expected the cursor to move. By offline analysis, it was found that the elicited event-related potentials were modulated by the expectation of the subject on the direction of cursor movement, and this phenomenon could be applied to the BCI controlled only by the intrinsic brain activation induced by user's expectation.

1 Introduction

In the current EEG-based BCI systems, users are requested to execute the specified and explicit tasks to activate the brain [1]. In the BCI system based on motor imagery, for example, users are requested to imagine the movement of their extremities. In the BCI system to detect evoked potentials or event-related potentials (ERPs), users choose one from simultaneously presented sensory objects (e.g. blinking visual targets [2] or auditory sounds [3]), and perceive or pay attention to the selected object. In such BCI systems based on evoked potentials or ERPs, however, the brain activation to detect is evoked by the presented sensory stimuli or is induced by their irregularity (e.g. oddball sequence), and no intrinsic process was present.

In this study, the possibility of an EEG-based BCI (brain-computer interface) system to control the movement of the computer cursor only by user's mental expectation was investigated. Subjects were requested to gaze at the moving cursor displayed on the LCD computer screen, and to expect the cursor to move to a preferred direction. By offline analysis, it was shown that the ERPs elicited by a sudden change of the direction of cursor movement were modulated by user's expectation or preference, and it was suggested that such an intrinsic neuronal process could be used for a BCI system based on implicit mental processes.

2 Methods

2.1 Experiments

The experiments in this study were approved by the Ethics Committee on Clinical Investigation, Graduate School of Engineering, Tohoku University, and were performed in accordance with the policy of the Declaration of Helsinki.

Four male volunteers (22 $\sim$ 24 years old) took part in the experiments as subjects. The EEG signal was recorded with Ag-AgCl electrodes at 32 locations (AFz, Fz, F1, F2, F3, F4, FCz, FC1, FC2, FC3, FC4, FC5, FC6, Cz, C1, C2, C3, C4, C5, C6, CPz, CP1, CP2, CP3, CP4, CP5, CP6,

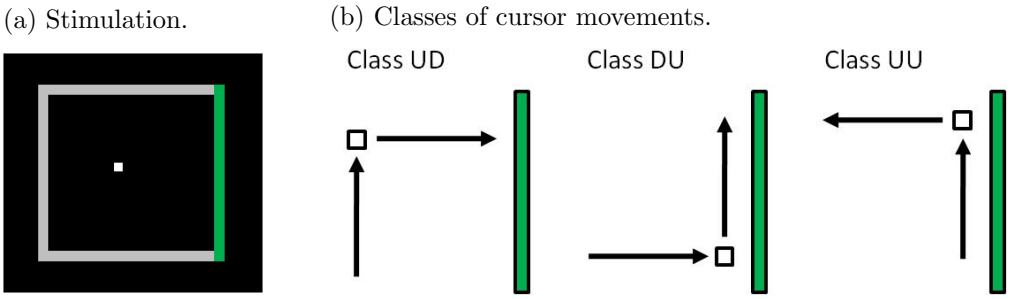


Figure 1: Visual stimulus and classes of cursor movements.

Pz, P1, P2, P3 and P4) based on the International 10-20 electrode system and referred to right earlobe with a left earlobe ground. The signals were amplified (gain × 500), bandpass-filtered (0.15 ∼ 100 Hz) and acquired at the sampling frequency of 1000 Hz.

The visual stimuli were presented by a personal computer with 17-inch LCD screen (1024 × 768 pixels). The schematic diagram of the visual presentation is shown in Figure 1 (a). The white computer cursor (0.5 cm×0.5 cm) moved within a square field (15 cm×15 cm) whose edges were colored gray. The background color was black. The cursor moved to up, down, left or right directions at a fixed speed (2 cm/s). The direction was randomly changed from one to another after a random time interval of 1 ∼ 4 s.

The cursor started to move from the center of the square field. Before starting each trial, one of the edges of the square was randomly chosen as the goal of the cursor, and the color of this edge was changed to green. Subjects were requested to gaze at the cursor with expecting it to reach the green edge. Each trial was terminated when the cursor reached one of the edges of the square. For each subject, 65 trials of experiments were conducted.

For accuracy, the actual timing of the change of the direction of cursor movement was detected by a photo-detector mounted on the LCD computer screen and it was recorded simultaneously with the EEG signal.

2.2 Data Analysis

The acquired data was analyzed offline. EEG responses to the changes of the direction of cursor movement were extracted by using a time window of −100 ∼ 600 ms relative to the change of the direction of cursor movement. The data was baseline-corrected using the averaged potential at −100 ∼ 0 ms. To remove the data with artifacts, the epoch data whose amplitude exceeded ± 50 µV was discarded and excluded from further analysis.

Prior to the pattern classification, the epoch data was lowpass-filtered (60 Hz), and down sampled to 125 Hz. The epoch data at time 0 ∼ 600 ms (76 points × 32 ch. = 2432 dimensions) was used as a feature vector for pattern classification.

The Fisher's LDA (linear discriminant analysis) [4] was applied for pattern classification. The purpose of the pattern classification was to classify the epoch data into the following three classes according to the directions before and after the change of cursor movement: undesired to desired direction (UD), desired to undesired direction (DU), and undesired to undesired direction (UU). Examples of these three classes are shown in Figure 1 (b).

3 Results

The ERP responses of all the subjects to the sudden changes of the direction of cursor movement are shown in Figure 2. These responses were averaged for each class (UD, DU and UU), subject and electrode. It was found that the ERPs were observed in all classes. These ERPs were elicited by the sudden change of the direction of cursor movement at the latency of around 100 ∼ 400 ms. In three out of four subjects (Subjects B, C and D), the ERP responses of each class had two

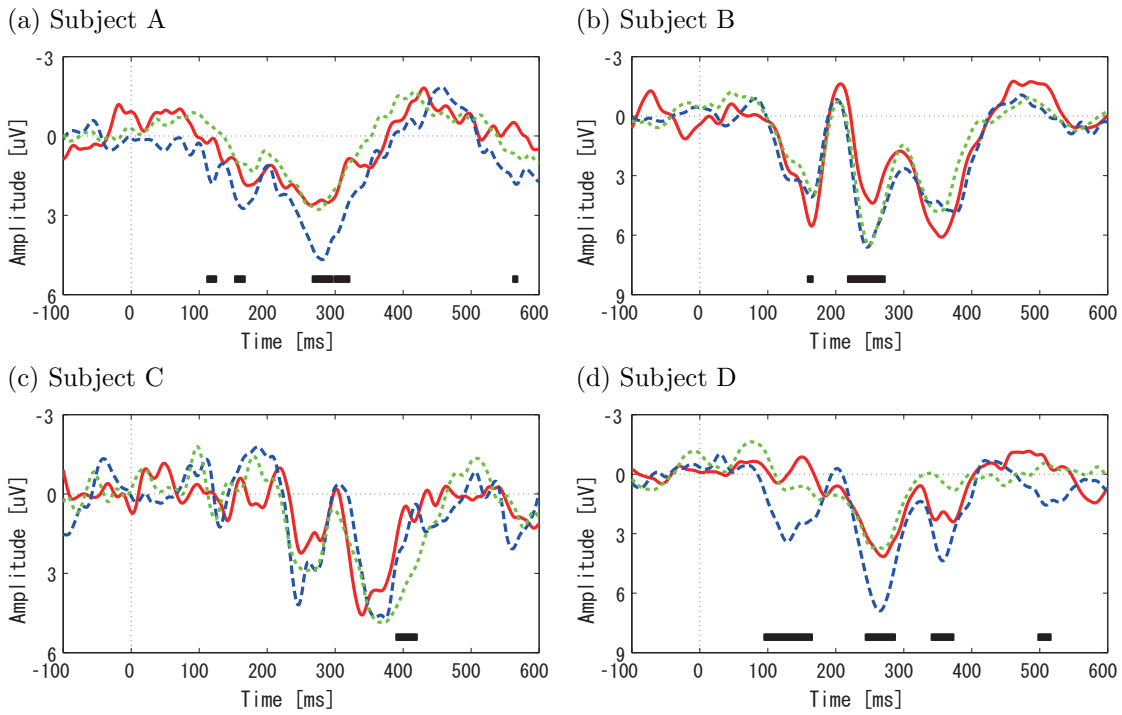


Figure 2: Averaged ERP responses to the change of the direction of cursor movement. Averaged responses of class UD (solid line), DU (dashed line) and UU (dotted line) at the electrode location Cz are shown for each subject. Black bars at the bottom of the figures denote the time period with significant differences of the amplitudes between one class (DU (a, d), UD (b) and UU (c)) and other two classes ($p < 0.05$).

positive ERP components at the latency of around 250 and 350 ms. In Subject A, only one positive ERP component which was similar to P300 was observed. In Subjects A, B and D, additional one ERP component at the latency of around $100 \sim 200$ ms was observed.

The amplitudes of elicited ERP components were different between three classes. A Welch's t-test revealed that the amplitudes of elicited ERP components had significant differences between one and other two classes ($p < 0.05$). It was suggested that these components were modulated by user's expectation.

The pattern classifiers were applied to the epoch data measured in these experiments so that the epoch data of one class are distinguished from those of other two classes (e.g. class UD vs. classes DU and UU). Note that the desired direction can be determined by the classifier if there is a difference of the ERP response between one and other two classes.

The true positive rate (TPR) and false positive rate (FPR) of the pattern classification are shown in Table 1. The scores were evaluated by the 10-fold cross validation. TPR values were higher than FPR values, and the classification of class UU and other two classes (UD and DU) had the highest TPR and FPR. But by the present preliminary evaluations, the overall scores were not high enough for BCI applications.

4 Discussion

These results suggested that the amplitudes of the ERP components with latencies of around $100 \sim 400$ ms were modulated by subject's expectation on the direction of cursor movement. As the paradigm in this study is related to the error processing between desired and actual directions of the cursor movement, these ERP components are thought to be the error-related potential [5]. The error-related potentials can be used to suppress undesired operations of the BCI system, in

Score	TPR [%]			FPR [%]		
Class to detect	UD	DU	UU	UD	DU	UU
Subject A	33.53	41.60	47.34	33.54	27.31	47.15
Subject B	36.05	33.88	51.96	24.63	31.96	50.15
Subject C	34.78	45.38	46.45	31.29	27.68	45.85
Subject D	36.92	41.43	50.89	28.37	37.86	48.68

Table 1: Classification accuracies.

which the error-related potentials induced by the mismatch between the output of BCI and the actual classification results are detected to ignore the incorrect classification results [6].

In this paradigm, the P300 components should also be elicited by the sudden change of the direction of cursor movement, regardless of the class depicted in Figure 1 (b), in addition to the error-related potentials induced by the expectation of the subjects. Thus, the pattern classifier was required to detect the difference of the ERPs in two or more intrinsic mental processes. That means the pattern classification in the present study is much harder than that to detect the presence or absence of the ERP components e.g. in the case of the P300 speller [2].

The detailed investigations on the preprocessing and pattern classification of the data, and the neurophysiological basis of the target brain activities, including the spatial distribution of the ERPs during the task, were left for the further study.

5 Conclusion

An EEG-based BCI system for controlling the movement of the computer cursor only by user's mental expectation was proposed and investigated. By the offline analysis, it was found that the event-related potentials elicited by the sudden changes of the cursor movement were modulated by the subject's expectation or preference on the direction of cursor movement. This phenomenon could be applied to the BCI controlled only by the intrinsic brain activation induced by user's expectation.

References

[1] J. R. Wolpaw, N. Birbaumer, W. J. Heetderks, D. J. McFarland, P. H. Peckham, G. Schalk, E. Donchin, L. A. Quatrano, C. J. Robinson, and T. M. Vaughan. Brain-computer interface technology: a review of the first international meeting. *IEEE Transactions on Rehabilitation Engineering*, 8, 2:164–173, 2000.

[2] E. Donchin, K. M. Spencer, and R. Wijesinghe. The mental prosthesis: assessing the speed of a P300-based brain-computer interface. *IEEE Transactions on Rehabilitation Engineering*, 8, 2:174–179, 2000.

[3] S. Kanoh, K. Miyamoto, and T. Yoshinobu. A brain-computer interface (BCI) system based on auditory stream segregation. *Journal of Biomechanical Science and Engineering*, 5, 1:32–40, 2010.

[4] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification*. Wiley Interscience Publication, 2nd edition, 2000.

[5] M. Falkenstein, J. Hoormann, S. Christ, and J. Hohnsbein. ERP components on reaction errors and their functional significance: a tutorial. *Biological Psychology*, 51, 2-3:87–107, 2000.

[6] P. W. Ferrez and J. del R. Millán. Error-related EEG potentials generated during simulated brain-computer interaction. *IEEE Transactions on Biomedical Engineering*, 55, 3:923–929, 2008.

Theoretical Framework and Simulation of an Adaptive BCI Based on Movement-Related and Error Potentials

X. Artusi[1], I. Khan Niazi[2], M. F. Lucas[1], D. Farina[2]

[1]IRCCyN Nantes - Centrale Nantes, France
[2]BCCN Göttingen - Georg-August University, Germany

xavier.artusi@irccyn.ec-nantes.fr

Abstract

Error potentials generated in response to an error made by the translation algorithm can be used to improve the performance of a BCI, as a feedback extracted from the user. The present study addresses the inclusion of error potentials in a BCI system based on the decoding of movement-related cortical potentials (MRCPs). First, we theoretically quantified the improvement in accuracy of the BCI system when using error potentials for correcting the output decision, in the general case of multiclass BCI. The derived theoretical expressions can be used during the design phase of any BCI system. Second, we studied in simulation the performance of the closed-loop system in order to evaluate its ability to adapt to the changes in the mental states of the user. By setting the parameters of the simulator to experimentally determined values, we showed that updating the learning set with the examples estimated as correct based on the decoding of error potentials leads to convergence to the optimal solution.

1 Introduction

New paradigms for brain computer interfacing (BCI), such as based on imagination of task characteristics, requires long training periods, have limited accuracy, and lack adaptation to the changes in the users' conditions. Error potentials generated in response to an error made by the translation algorithm can be used to improve the performance of a BCI, as a feedback extracted from the user [1–3]. Our work addresses the inclusion of error potentials in a BCI system based on the decoding of movement-related cortical potentials (MRCPs) associated to the speed of a task. The overall system needs an online learning classifier, a detector of error potentials and strategies for taking a decision and updating the learning set. The purpose of the classification algorithm is the decoding of the user intentions. In our experimental protocol, it corresponds to the discrimination of kinetic parameters from single-trial MRCPs. The detector of error performs an evaluation (as *correct* or *wrong*) of the user reaction to the displayed decisions. Its output is used to correct the decision of the system and to update the learning set with pertinent examples, in order to allow adaptivity of the task classification. Classification and error detection methods are detailed in [4]. In the present study, we first theoretically quantify the improvement in accuracy of the corrected BCI system. Second, we study, in simulation, the performance of the closed-loop system in order to evaluate its ability to adapt to changes in the mental states of the user, and select the strategy for using the detection of errors.

2 Methods

2.1 Theoretical Improvement of the BCI System

The detector of error provides information on the accuracy of the response of the classifier (noted SVM) from EEG recorded after the display of this response. If the response is detected as *wrong*, it is not possible with more than two classes ($n_c > 2$) to deduce the true class even with a perfect detector. So a natural decision rule (noted 1) to integrate the detector consists in taking a decision

only if the SVM response is estimated as *correct*. When the response is estimated as *wrong*, the subject repeats the imagination task. We propose to use this decision rule also in the biclass case and to compare it to the decision rule 2 used in [3] that consists, when the error detector estimates the response of the SVM as wrong, in taking the opposite decision.

In order to calculate the probability of error of the corrected system, the problem is formalized using four random variables: Ω and $\hat{\Omega}$ with values in $\{\omega_1, \ldots, \omega_{n_c}\}$ represent respectively the movement intention and the intention decoded by the SVM, $\omega, \hat{\omega}$ being the corresponding realizations; E and $\hat{E}$ with values in $\{correct, wrong\}$ represent the accuracy of the SVM decision and its estimation by the detector, $e, \hat{e}$ being the corresponding realizations; $e = correct$ if $\hat{\omega} = \omega$ else $e = wrong$. We denote $P_\Omega(\omega)$ the *a priori* probability of the class ω, $P_{\hat{\Omega}|\Omega}(\hat{\omega}, \omega) = P_E(e)$ the conditional probability of the intention decoded by the SVM knowing the true intention of the user and $P_{\hat{E}|\hat{\Omega}\Omega}(\hat{e}, \hat{\omega}, \omega) = P_{\hat{E}|E}(\hat{e}, e)$ the conditional probability of the detector.

Following decision rule 1, the probability of error of decision corresponds to the probability $P_{\hat{E}|E}(correct, wrong)$ that, when the detector indicates *correct*, the class $\hat{\omega}$ given by the SVM is different from the true intention ω of the user. We denote it P_{Er_1} :

$$P_{Er_1} = \sum_{i,j \neq i} P_{\Omega\hat{\Omega}|\hat{E}}(\omega_i, \hat{\omega_j}, correct) = P_{\hat{E}}^{-1}(correct) \sum_{i,j \neq i} P_{\hat{E}|\hat{\Omega}\Omega}(correct, \hat{\omega_j}, \omega_i).P_{\hat{\Omega}|\Omega}(\hat{\omega_j}, \omega_i).P_\Omega(\omega_i)$$

with $P_{\hat{E}}(\hat{e}) = \sum_e P_{\hat{E}|E}(\hat{e}, e).P_E(e)$. The performance of the corrected system is also characterized by the probability that the subject has to imagine again the task; it corresponds to the probability that the detector estimates the SVM decision as *wrong*: $P_r = P_{\hat{E}}(wrong)$. The probability for the user to repeat m times the same task is $(P_r)^m$. For decision rule 2, the probability of error is :

$$P_{Er_2} = P_E(wrong).P_{\hat{E}|E}(correct, wrong) + P_E(correct).P_{\hat{E}|E}(wrong, correct)$$

with no repetitions ($P_{r_2} = 0$).

The bit transfer rate [5], taking into account the above probabilities, is calculated for an n_c classes problem as: $BpT = [\log_2(n_c) + (1 - P_{Er}) \log_2(1 - P_{Er}) + P_{Er} \log_2(P_{Er}/(n_c - 1))](1 - P_r)$.

These theoretical formulations will be used in Section 3 in a biclass case with values obtained by experimental recordings.

2.2 Updating the Learning Set and Simulation of the Overall System

In order to allow the adaptation of the BCI to the user changes, it is necessary to define a strategy to modify the learning set with pertinent examples. As the adaptation is the goal, we use a moving window: when a new example is integrated to the learning set, the oldest example is removed. To integrate a new element, it is necessary to have a good estimation of its class. This is true only for examples that are considered as *correct* by the error detector. Two strategies are tested to update the learning set:

- taking all examples that are considered as *correct* by the detector;
- among the examples that are considered as *correct*, taking only examples near the current decision surface, because they are considered as carrying more information. This corresponds to the technique of sampling by uncertainty used in active learning [6]. This proximity is measured by the value of the decision function f (the trial x is taken if $|f(x)| \leqslant 1$).

This investigation can be done only in simulation because it needs to estimate at each time the performance of the evolving system (that cannot be performed on experimental data from one trial) and to compare it to a known reference. We have simulated a system corresponding to the SVM and the error detector. Signals are not simulated but only feature vectors in a representation space of two dimensions. We consider in $\mathbb{R}^2$ n_c classes $\{\omega_1, ..., \omega_i, ..., \omega_{n_c}\}$. The elements of each class ω_i are distributed by a gaussian density $\mathcal{N}(\mu_i, \Sigma_i)$ that can evolve. The opinions $\hat{e}$ of the detector are the realizations of a random binary variable $\hat{E}$ of conditional density $P_{\hat{E}|E}$. As initialization, the learning set is composed of n_{init} labeled examples. Each time a new example is generated, the classification is performed by the online multiclass SVM. Then, the new elements may be integrated with the label estimated by the SVM depending on the error detector

Subjects	1	2	3	4	5	6	Av.
ED : Good detection of *correct* trials	92 %	91 %	93 %	87 %	83 %	80 %	88 %
Good detection of *wrong* trials	73 %	43 %	60 %	60 %	37 %	63 %	56 %
SVM : Movement classification error rate	28 %	24 %	32 %	30 %	22 %	22 %	26 %
Initial Bit transfer rate	0.14	0.20	0.10	0.12	0.24	0.24	0.17
Error probability of the corrected BCI, P_{Er_1}	10.0 %	16.5 %	16.8 %	16.5 %	17.6 %	11.5 %	15 %
Probability of repetition, P_r	26.2 %	17.2 %	23.4 %	27.1 %	21.4 %	29.5 %	24 %
Final bit transfer rate, BpT_1	0.39	0.29	0.26	0.26	0.26	0.34	0.30
Error probability of the corrected BCI, P_{Er_2}	13.3 %	20.5 %	17.6 %	21.1 %	27.1 %	23.7 %	21 %
Final bit transfer rate, BpT_2	0.43	0.27	0.33	0.26	0.16	0.21	0.28

Table 1: The performances of the error detector (ED) and of the SVM are obtained from experimental data. The performances of the corrected BCI are calculated for both decision rules using theoretical formulations of Section 2.1.

and the strategy. The size of the learning set increases until n_w (size of the moving window), and then remains constant by removing the oldest elements. At each time, the probability of error of task classification is estimated empirically from a test set generated using the current densities $\mathcal{N}(\mu_i, \Sigma_i)$, in four cases: the online SVM without error detector, the ideal Bayesian classifier knowing the theoretical densities of the classes at each time and without error detector, the online SVM corrected and updated using the error detector, the ideal Bayesian classifier corrected by the error detector. Bayesian classifiers are used as optimal references for the SVM.

In order to use realistic values for the sources of error, we analyze the performance of the task classification algorithm and error potential recognition on experimental data in the case of a biclass BCI, on six healthy subjects. Each subject was asked to perform two motor imagination tasks, slow and fast right arm flexion (60 trials of each class). After performance of each task the subjects were presented with a pseudo feedback (*slow/fast*) that randomly corresponded to a correct or wrong classification (75 % correct and 25 % wrong).

3 Results and Conclusion

In Table 1, the experimental results of the task classification and of the detection of error potentials are used as estimation of the conditional probabilities of the SVM ($P_{\hat{\Omega}|\Omega}$) and of the detector ($P_{\hat{E}|E}$) to calculate the theoretical global error probability and the bit transfer rate of the corrected system. On average over all subjects, with the decision rule 1 the error rate was 15 % (to be compared to the 26 % in the case of SVM alone) with a repetition rate of 24 %. With the decision rule 2 the error rate was 21 %. The bit transfer rate was improved by 76 % in the first case and by 64 % with the second one. Furthermore with the decision rule 1, the bit transfer rate is improved for all the subjects that is not the case for the decision rule 2. According to these results we used the decision rule 1 to correct the SVM decision in the next simulations.

The results of the experimental analysis concerning the performance of the detector were used for simulations performed in the general case of multi-class classification. The evolution of the 4 error rates described in Section 2.2 are reported in Figure 1 for four simulations with 3 classes. Classes are moving in the feature space and the distances between the centers are decreasing or increasing, simulating changes in the brain signals due to user adaptations. The conditional probabilities taken for the detector are the ones estimated on average from the experimental data. $n_{init} = 10$, $n_w = 40$. These simulations show that the looped system can adapt and allow us to choose the best strategy for updating the BCI system. In the simulated conditions of Figure 1a and Figure 1b where distances between the centers are increasing, both strategies simulated converge to the ideal Bayesian classifier, whereas the SVM alone diverges, as expected. The second simulation, where the classes are translated and distances between the centers are decreasing, shows that the first strategy (taking all examples estimated as *correct*) allows correct adaptation and convergence of the classifier whereas the second strategy (uncertainty sampling) does not converge.

In this study, we show that including a detection of errors, although far from ideal, increases

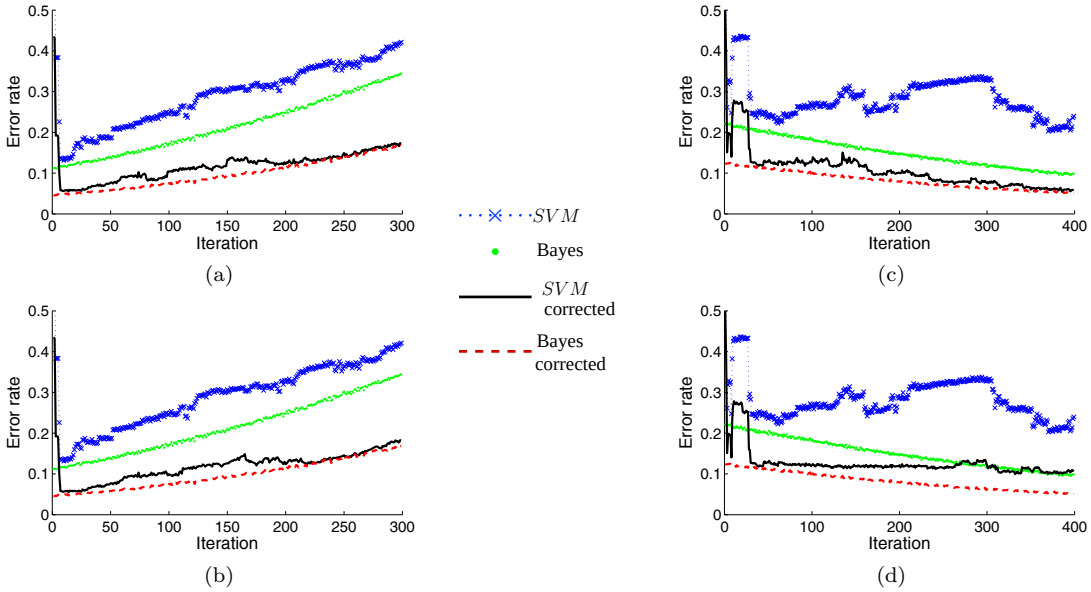


Figure 1: Simulation of the adaptive system with distance between the centers of the classes decreasing (a, b) and increasing (c, d). The strategies used for updating the learning set consist in: (a, c) taking all examples estimated as *correct* by the detector, (b, d) applying a technique of uncertainty sampling among these examples, taking only examples near the decision surface.

substantially the accuracy of a BCI system with the decision rule 1. The repetition of the imaginary movement on average once every 4 commands seems to be reasonable for clinical applications. Then in the representative examples shown, we also show that the strategy of using all the well classified signals as learning set, and forgetting the oldest labeled signals, allows adaptation to the user changes and converges in performance to the optimal solution. Simulations reflect the real online application, algorithms used works in real-time, especially the incremental/decremental SVM decoding the movement intention.

References

[1] R. Chavarriaga and J. d. R. Millán. Learning from EEG error-related potentials in noninvasive brain-computer interfaces. *IEEE Trans. on Rehabilitation Engineering*, 18(4):381–388, 2010.

[2] B. Dal Seno, M. Matteucci, and L. Mainardi. Online detection of P300 and error potentials in a BCI speller. *Computational Intelligence and Neuroscience*, pages 1–5, 2010.

[3] P. W. Ferrez and J. d. R. Millán. You are wrong! automatic detection of interaction errors from brain waves. *International Joint Conference On Artificial Intelligence*, 19:1413–1418, 2005.

[4] X. Artusi, I. Khan Niazi, M. F. Lucas, and D. Farina. Accuracy of a BCI-based on movement-related and error potentials. Technical report, IRCCyN, 2011.

[5] B. Blankertz, C. Schäfer, G. Dornhege, and G. Curio. Single trial detection of EEG error potentials: A tool for increasing BCI transmission rates. *Artificial Neural Networks-ICANN 2002*, pages 138–138, 2002.

[6] G. Schohn and D. Cohn. Less is more: Active learning with support vector machines. In *Proc. of the 70th Int. Conf. on Machine Learning*, pages 839–846. Morgan Kaufmann, 2000.

Adaptive Classification Improves Control Performance in ERP-Based BCIs

S. Dähne[1], J. Höhne[1], M. Tangermann[1]

[1]BBCI Lab, Machine Learning Dept., Berlin Institute of Technology, Berlin, Germany

Abstract

This contribution investigates the effects of applying an unsupervised adaptation mechanism to linear classifiers for Brain-Computer Interfaces (BCI). Specifically, we track changes in the first two moments of the unlabeled data distribution. Changes are adaptively compensated by recalculating the classifier based on short, consecutive data segments. The approach is validated on three auditory oddball data sets containing a total of $N = 37$ subjects, of which 6 were used for model selection and the remaining 31 for validation. We find a significant performance increase (up to 14 %) for the adaptive scheme compared to a fixed classifier. The increase is largest for subjects with low performance.

1 Introduction

Brain-Computer Interface (BCI) systems observe and analyze neural signals recorded for example via EEG. A BCI aims to translate these signals into control commands for a computer program or neural prosthesis. In order to reach this goal, BCIs must not only be accurate in their class-discrimination, they must also be robust with respect to the inherently low signal to noise ratio. Furthermore, the robustness to changes in the data distributions of ongoing or task-specific signals, often referred to as non-stationarities, is a desirable characteristic of a BCI.

In principle two approaches can be identified to reach the latter goal: finding signal representations that are maximally invariant with respect to such non-stationarities [1] or trying to detect and to compensate them [2]. Adaptive methods, which fall into to class of the compensating approach have been successfully applied in the context of BCIs based on motor-imagery [3].

In this work we pursue a similar approach in the context of BCIs based on auditory event-related potentials (ERP). Specifically, we track changes in moments of the feature distributions over time and carefully adapt the classifier to alleviate the effects of non-stationarities.

2 Methods

2.1 Data

The EEG data sets used for offline-analysis stem from three separate auditory oddball ERP BCI studies, referred to as PASS2D, AMUSE, and WARP, respectively. All data sets were recorded using a 63 channel layout and 1000 Hz sampling rate. The data was band-pass filtered with a pass band of 0.4 to 40 Hz and down-sampled to 100 Hz prior to analysis. Epochs were extracted between -150 ms and 800 ms relative to the stimulus onset. Details specific to each of the studies are as follows [PASS2D, AMUSE, WARP]: number of stimulus epochs in calibration phase = [3402, 4320, 972], number of stimulus epochs in the online phase = [11987, 8100, 10692], target to non-target ratio [1:8, 1:5, 1:8], number of participants = [10, 21, 6]. Further details about the PASS2D and AMUSE paradigms are provided by [4] and [5].

2.2 Classification

2.2.1 Feature Extraction

After the calibration phase, temporal features were extracted from the raw epochs and utilized to train a linear classifier. In this context, a temporal feature is defined as the mean EEG value of a post-stimulus interval. The intervals are the same for all channels and were chosen by hand by the experimenter (PASS2D and AMUSE) or by a heuristic (WARP). Depending on the the subject, a number of 2, 3, or 4 intervals are chosen. This yields a feature vector of dimensionalityN_{fv}, which is the product of the number of intervals and the number of recording channels.

2.2.2 Fixed Classifier

The data from the calibration measurements was used to train a Linear Discriminant Analysis (LDA) classifier. This classifier was not adapted during the (offline) classification of the online data and thereby represents the baseline for comparison. LDA was used, as this method has proven to yield robust and accurate results in many studies [6] and allows for inspection and analysis of the decision mechanism (i.e. inspection of the feature-weights). Thus, the classifier output for a data sample $\mathbf{x}$ is given by

$$\begin{aligned} \mathrm{LDA}_{fix} &= \mathbf{w}^T\mathbf{x} + b, & (1)\\ \text{with } \mathbf{w} &= \mathbf{C}^{-1}(\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1) & (2)\\ b &= -0.5\mathbf{w}^T(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2). & (3) \end{aligned}$$

In this formalism, $\mathbf{C}^{-1}$ is the inverse of the covariance matrix of the pooled data and $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ denote the mean of the class *target* and *non-target*, respectively. Since N_{fv} can become quite large (e.g. $N_{fv} = 252$ for 4 intervals and 63 EEG channels) relative to the number of training samples, we applied a regularization method (shrinkage [6]) to estimate $\mathbf{C}$.

The binary decision about whether $\mathbf{x}$ belongs to either of the two classes is based on the sign of LDA_{fix}. This classifier is termed *fixed* because it is computed on the calibration data and remained fix thereafter.

2.2.3 Adaptive Classifier

Note that the computation of w (Equation 2) consists of two factors, namely $\mathbf{C}^{-1}$ and the difference of the class means $(\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1)$. Only the second factor requires class labels which is only available during calibration. The first factor, on the other hand, can be computed in an unsupervised manner during calibration as well as during the online phase [2].

Thus, after being initialized on the calibration data, the classifier is made adaptive by updating the pooled mean and pooled covariance. These two moments are re-estimated in small consecutive segments of the (online) data. The segment size we used corresponded to the number of epochs that within a single trial, i.e. the selection of an element in a spelling application. For PASS2D and WARP data sets a single trial consisted of 135 epochs (AMUSE data: 90 epochs). The segment-wise mean and covariance estimates are used to update the weight vector $\mathbf{w}_n$, which is now a function of the segment index n, according to

$$\mathbf{w}_{n+1} = \mathbf{C}_{n+1}^{-1}(\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1), \quad (4)$$

with $\mathbf{C}_{n+1} = (1-\alpha)\mathbf{C}_n + \alpha\mathbf{C}_{\text{segment}}$ The bias b_n is updated accordingly.

The update ratio α is a crucial parameter of the adaptive scheme. It is not obvious how to determine the parameter value a priori, because it should be adjusted to the time-scale of the ongoing changes in the feature distribution. We chose a purely data driven approach to tackle this problem: First, we computed the α value that optimizes the classification results on the WARP data set. This was done by running the adaptive classifier with a range of α values for all subjects of the data set and comparing the resulting classification performance to that of the

static classifier. Second, we took the value that yielded the best performance (averaged over the 6 WARP data set subjects) and used it for the adaptive classification of the two remaining data sets.

3 Results

The classification results of the remaining $N = 31$ subjects is shown in Figure 1 **a**. In this scatter plot each dot corresponds to one subject. Its x-coordinate represents the classification result of LDA_{fix} and its y-coordinate to the classification result obtained with LDA_{adapt}. For points above the diagonal the adaptive classification scheme performs better than the fixed scheme. This was the case for 17 subjects, while for 11 subjects there was no difference in performance and for 3 subjects the performance decreased by using the adaptive classifier. The difference in classification performance was significant with $p < 0.001$ in favor of the adaptive scheme.

Figure 1 **b** to **d** shows details about the ongoing dynamics during the online run of the subject for which the benefit of adaptive LDA classification was largest. In **b** we show the evolution of the adapted bias of the first 63 feature dimensions (i.e. the early interval of each channel, cf. Section 2.2.1). Scalp maps show the spatial distribution for 4 selected time segments (indicated by the index n). In subplot **d** of the same figure the dynamics of the weight vector $\mathbf{w}(\text{n})$ is visualized, again with scalp maps of these time segments. Finally, in subplot **c** we show the evolution of the confidence of the classifier, computed as the average difference between the classifier output for *targets* and classifier outputs for the *non-target* class that is closest to the *target* output. The larger this difference is, the greater is the class discrimination. Negative values indicate misclassifications. Note the deviation of the two lines towards the end of the BCI session (segments 80 to 128). In this interval, the fixed classifier made 21 mistakes, while the adaptive classifier was wrong in only four times. Furthermore, the dynamics of the adapted means reveal drifts that are strongest for the frontal electrodes. This was observed in addition in a visualization of the evolution of the variances (not shown). A very interesting effect appears after segment number 42. A drastic change in feature means and variance must have taken place, which is eminent in addition by the classifier confidence for these segments. Such changes might have been caused by a sudden move of the subject or other physical influences.

4 Discussion

In this work we have investigated the effect of using a simple and unsupervised adaptation scheme which allows to compensate for slow changes in the first two moments of the pooled feature distribution. The obtained improvement in classification rates indicate the usefulness of the applied approach.

However, care has to be taken when choosing the update parameter. Too fast adaptation rates are prone to overcompensate sudden influences (e.g. trials contaminated by artifacts) and may thereby deteriorate the resulting classification performance notably. We advocate in favor of a conservative choice for the update rates because further simulations have shown that (too) slow update rates are less detrimental (if at all) compared to (too) fast updates (data not shown). The pragmatic approach we chose for finding a suitable value of the update parameter still leaves room for improvement. We conjecture that it would be beneficial to find a subject specific value, rather than taking the same for the entire data set.

An additional benefit of the adaptive method is that it can be applied *post-hoc* after an experiment in order to investigate the time-scale of potential non-stationarities in the signal. Finding the optimal adaption time scale a posteriori may reveal insights about further improvements of the experimental design (such as block size, when to include breaks, etc.).

In conclusion, we find that it is beneficial to apply adaptive classification methods with carefully chosen update parameters in the context of both, auditory and visual ERP BCIs.

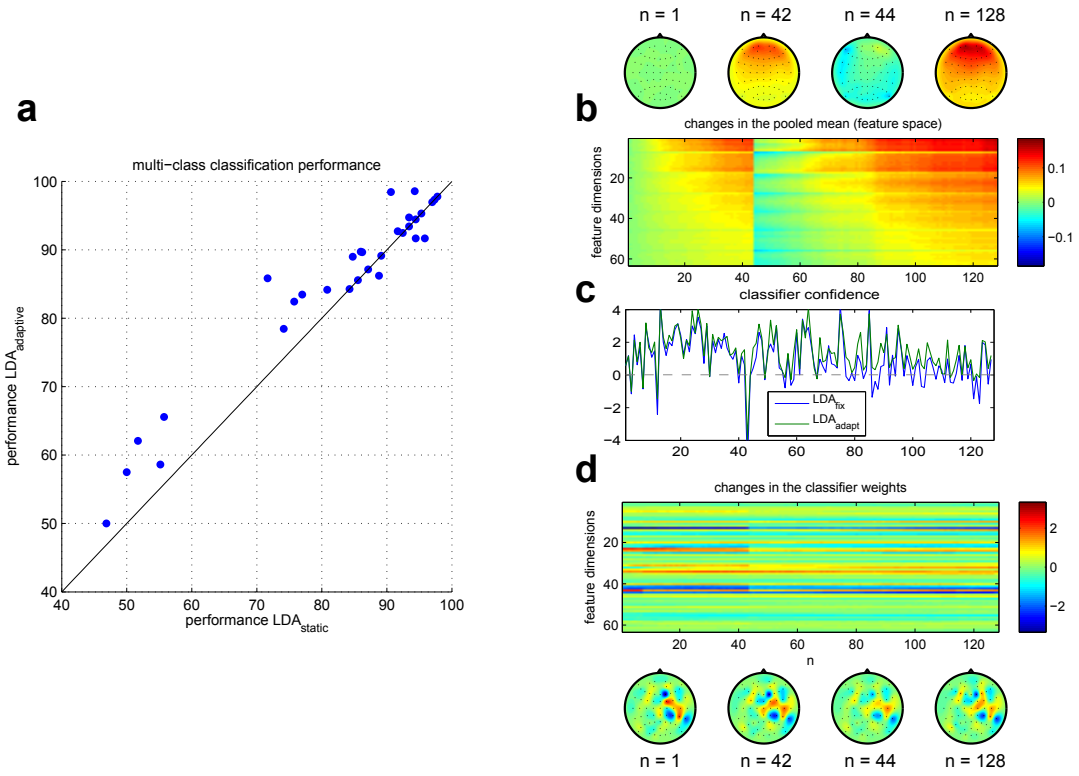


Figure 1: Results of adaptive classification. **a**: Performance of LDA_{fix} versus LDA_{adapt}. **b** – **d**: Dynamics of subject VPnz of the PASS2D data set. See text for details.

Acknowledgments

This work is supported by the European ICT Programme Project FP7-224631 and by GRK 1589/1.

References

[1] P. von Bünau, F. C. Meinecke, F. Király, and K. R. Müller. Finding stationary subspaces in multivariate time series. *Physical Review Letters*, 103: 214101, 2009.

[2] C. Vidaurre, C. Sannelli, K. R. Müller, and B. Blankertz. Co-adaptive calibration to improve BCI efficiency. *J Neural Eng*, 8(2): 025009 (8pp), 2011.

[3] C. Vidaurre, M. Kawanabe, P. von Bünau, B. Blankertz, and K. R. Müller. Toward unsupervised adaptation of LDA for brain-computer interfaces. *IEEE Trans Biomed Eng*, 58(3): 587 –597, 2011.

[4] J. Höhne, M. Schreuder, B. Blankertz, and M. Tangermann. Two-dimensional auditory P300 speller with predictive text system. *Conf Proc IEEE Eng Med Biol Soc*, 1: 4185–4188, 2010.

[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] B. Blankertz, S. Lemm, M. S. Treder, S. Haufe, and K. R. Müller. Single-trial analysis and classification of ERP components - a tutorial. *NeuroImage*, 2010, in press.

[7] J. Schäfer and K. Strimmer. A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. *Stat Appl Genet Mol Biol*, 4: Article32, 2005.

How do You Like Your P300 Speller: Adaptive, Accurate and Simple?

P. J. Kindermans[1], D. Verstraeten[1],
P. Buteneers[1], B. Schrauwen[1]

[1]Electronics and Information Systems department, Reservoir Lab, Ghent Univeristy

PieterJan.Kindermans@UGent.be

Abstract

We show how a P300 BCI can be improved by using an adaptive classifier based on a class-re-weighted version of Ridge Regression. The unsupervised adaptation leads to an improvement in both the spelling accuracy and the individual P300 detection. This can in turn be used to reduce the number of training trials while maintaining the same level of performance.

1 Introduction

Brain Computer Interfaces (BCI) are an innovative way to control machines and computers. The P300 spelling paradigm is one of the most widespread and stable systems, but it still suffers from some problems. Some of these problems are the tiresome training procedure and the inter session variability. In this paper we propose a solution to these two issues: an adaptive classifier.

Related work on adaptive P300 can be found in [1], where a general subject unspecific classifier is used and a new subject specific classifier is created during the use of the P300 system. Another adaptive system with 2 classifiers is proposed in [2]. Finally, a Support Vector Machine (SVM) based system can be found in [3].

2 Methods

P300 We performed the experiments on dataset 2 from BCI Competition III [4]. We chose this dataset because it is publicly available and challenging. The dataset itself contains 64 channel EEG from 2 different test subjects with 85 training and 100 test characters. The number of epochs per character is 15. Each row and each column is intensified once during an epoch. The P300 grid is 6 by 6.

Preprocessing Preprocessing is an essential part of a BCI setup but we tried to keep this to a minimum. The EEG is preprocessed character by character, beginning with a Common Average Reference filter followed by a subject specific bandpass filter. The lower cutoff frequency was 0.5 Hz, the upper was 8 Hz for subject A and 7 Hz for subject B. (optimized by cross validation). We then proceeded by normalizing each EEG channel to zero mean and unit variance and subsampling from 240 Hz to 60 Hz. In the final step we retained 10 samples per channel, centered at the expected timestep for the P300.

Basic Classifier The resulting feature vectors are classified by a class-re-weighted version of Ridge Regression [5]. This is an extension of Ridge Regression that is able to handle unbalanced data sets. We selected this classifier because it achieves good performance and it is computationally cheap to update on-line.

Subject	epochs	SUP	ONL	BOUND	ESVM [6]	CNN-1 [7]	MCNN-1 [7]
A	5	67.0	67.0	68.0	**72.0**	61.0	61.0
	10	88.0	89.0	**91.0**	83.0	86.0	82.0
	15	96.0	96.0	95.0	**97.0**	**97.0**	**97.0**
B	5	78.0	**84.0**	**84.0**	75.0	79.0	77.0
	10	**93.0**	**93.0**	**93.0**	91.0	91.0	92.0
	15	**96.0**	**96.0**	**96.0**	**96.0**	92.0	94.0

Table 1: Spelling accuracies when all training characters are used. Evaluation is done using the first 5, 10 or 15 epochs. The evaluation method is the one from BCI Competition III.

The first step in training the classifier is reweighing each sample by a factor $\frac{N}{N_c}$, where N is the total number of samples and N_c is the number of samples belonging to class c. For P300 this is 6 for positive samples and $\frac{6}{5}$ for negative samples. The classifier can be computed as follows:

$$\mathbf{w} = \left(X^T R X + \lambda I\right)^{-1} X^T R \mathbf{y}, \tag{1}$$

where λ is the regularization parameter which is optimized by cross validation, X is a matrix containing all the samples (one per row), R is a diagonal matrix with R_{nn} the weight for sample n and $\mathbf{y}$ a column vector containing the desired output (-1 or 1). The classifier output for a specific sample equals $\mathbf{xw}$, where $\mathbf{x}$ is a row vector.

When we predict a character, we first compute classifier outputs for all intensifications independently. The score for each character is the sum of classifier outputs over all intensifications highlighting that character. The character with the highest score gets predicted.

Adaptive Classifier The basic classifier is adapted in an unsupervised manner by simply adding on-line labeled test samples to the training set and re-training the classifier. To do this, we first predict the desired character with the old classifier, then we label each intensification containing the predicted character as a positive sample and the other as negative samples. For every six samples, five will be negative and one positive. This means that four negative samples will always be given the correct label. Only the positive and one negative sample might receive the wrong label. When half of the characters are predicted correctly then 83 % samples are labeled correctly. We update the $X^T R X$ and $X^T R \mathbf{y}$ matrices from Equation (1) as follows:

$$(X^T R X)_{new} = (X^T R X)_{old} + (X^T R X)_{online}, \tag{2}$$

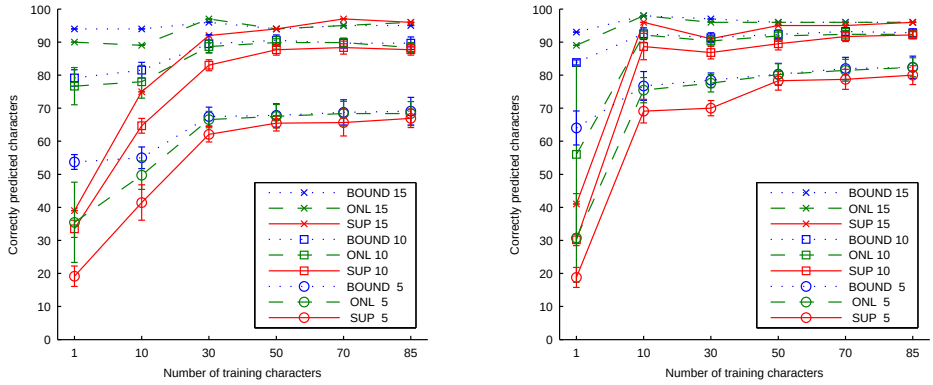
$$(X^T R \mathbf{y})_{new} = (X^T R \mathbf{y})_{old} + (X^T R \mathbf{y})_{online}. \tag{3}$$

The classifier weight vector is then retrained using Equation (1). The regularization parameter is not updated but the one from the original cross validation is kept.

3 Results

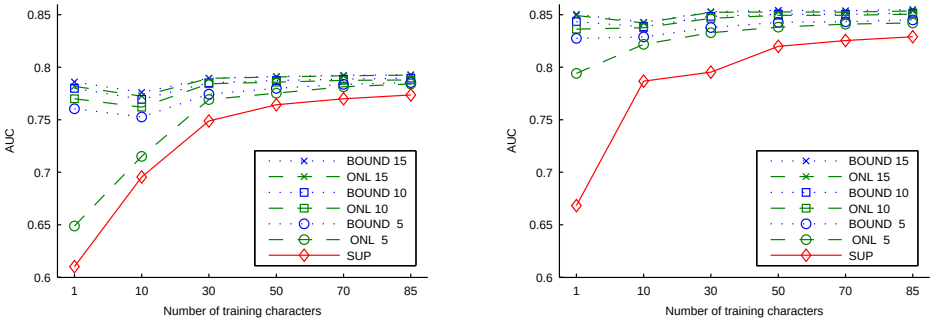
We evaluated three different classifiers. The first one is the basic classifier which is only trained in a supervised manner (SUP). The second classifier starts out as the basic classifier but is adapted after each character prediction (ONL). The last classifier uses the same principle as ONL but uses the true labels in the update, therefore it is an upper bound on the performance (BOUND).

The first experiment mimics the procedure in BCI Competition III. We trained the classifier for each subject using the entire training session for that subject. The first 5, 10 or 15 epochs are used to predict the characters. The results are shown in Table 1, together with the results from [6], where an ensemble of Support Vector Machines (ESVM) was used - this method won BCI Competition III - and those from [7], where Convolutional Neural Networks (CNN-1 and MCNN-1) were used.



(a) Subject A. (b) Subject B.

Figure 1: Character prediction results as a function of the number of characters used during supervised training.



(a) Subject A. (b) Subject B.

Figure 2: Classifier performance measured as Area Under Curve (AUC) as a function of the number of characters used during supervised training.

To test the influence of the size of the initial training set, we performed a second experiment. Figure 1 shows the character prediction accuracy as a function of the number of labeled training characters from the reduced training set. The reduced training set contains only the first k characters from the entire training set. We averaged the results on 5 and 10 epochs over all possible start epochs (e.g. we used the first five epochs starting at epochs 1,2,...,11).

The last experiment evaluates the classification performance on the individual intensifications. This is done by computing the Area Under Curve (AUC) for a Receiver Operator Characteristic plot. In each run, we used 90 % of the test set to update the classifier and the remaining 10 % to compute the AUC for the classifier. Results are averaged out over 10 runs, such that each part of the test set was used once to compute the AUC.

4 Discussion

The results from the first experiment (Table 1) indicate that SUP is a powerful classifier for P300 and can outperform much more complex and computationally demanding systems. The ESVM approach classifies only 1 extra character correctly on 15 epochs. If we reduce the number of epochs to 10, we see that our classifiers outperform the other approaches. The ESVM approach has better performance on subject A when only 5 epochs are used, but subject B gives the opposite

result and the difference in accuracy is higher in the latter case.

The classification results from Figure 1 show that the adaptivity improves the spelling accuracy. The effect is most noticeable when the number of training samples and epochs is low. A remarkable result is the fact that on dataset B the highest accuracies are achieved when only 10 characters are used for training and 15 epochs during testing (98 % correct) although the base classifier performs best when all the training data is used. We assume that this is caused by the higher influence of the self-labeled samples when less training data was available. When the basic classifier has low accuracy, around 20 % for 5 epochs after training on 1 character, the performance still increases through adaptation. As the classification accuracy increases, the difference in performance between ONL and BOUND becomes negligible.

In Figure 2 we observe the same behaviour for the AUC results as for the character prediction from the previous experiment. The SUP method benefits from an increased number of training samples. It is also clear that the adaptation ONL can drastically improve the performance even when the starting classifier performs badly (an AUC less than 0.65). Finally we see that there is no significant difference in performance between ONL and BOUND when the classifier used to label the character has high accuracy.

5 Conclusion

We proposed a simple, computationally cheap and accurate P300 classifier based on a re-weighted version of Ridge Regression for classification. Preprocessing is kept minimal by restricting it to Common Average Reference, a subject specific frequency filter and normalization of the input. This classifier is then adapted on-line to increase performance. This enables us to achieve high spelling accuracies with a limited number of epochs and training samples. Future work should focus on reducing the influence of bad training examples and inter subject transferability.

References

[1] S. Lu, C. Guan, and H. Zhang. Unsupervised brain computer interface based on intersubject information and online adaptation. *IEEE Trans. on Neural Systems and Rehabilitation Engineering*, 17(2):135 –145, 2009.

[2] R. C. Panicker, S. Puthusserypady, and S. Ying. Adaptation in P300 brain-computer interfaces: A two-classifier cotraining approach. *IEEE Trans. on Biomedical Engineering*, 57(12):2927 – 2935, 2010.

[3] 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 Recognition Letters*, 29(9):1285 – 1294, 2008.

[4] B. Blankertz, K. R. Müller, D. J. Krusienski, G. Schalk, J. R. Wolpaw, A. Schlögl, G. Pfurtscheller, J. del R. Millán, M. Schröder, and N. Birbaumer. The BCI competition III: validating alternative approaches to actual BCI problems. *IEEE Trans. on Neural Systems and Rehabilitation Engineering*, 14(2):153 –159, 2006.

[5] K. A. Toh. Deterministic Neural Classification. *Neural Computation*, 20(6):1565 – 1595, 2008.

[6] A. Rakotomamonjy and V. Guigue. BCI Competition III: Dataset II- Ensemble of SVMs for BCI P300 Speller. *IEEE Trans. on Biomedical Engineering*, 55(3):1147 –1154, 2008.

[7] H. Cecotti and A. Gräser. Convolutional neural networks for P300 detection with application to brain-computer interfaces. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 99, 2010.

A Model of BCI-Control

A. Kübler[1,2], B. Blankertz[3], K. R. Müller[3,4], C. Neuper[5,6]

[1]Department of Psychology I, University of Würzburg, Germany
[2]Institute of Medical Psychology and Behavioural Neurobiology, Univ. of Tübingen, Germany
[3]Machine Learning Laboratory, Berlin Institute of Technology, Germany
[4]Institute for Pure and Applied Mathematics, UCLA, Los Angeles, USA
[5]Institute for Knowledge Discovery, Laboratory of BCIs, Graz University of Technology, Austria
[6]Department of Psychology, Karl-Franzens University Graz, Austria

andrea.kuebler@uni-wuerzburg.de

Abstract

Control of brain-computer interfaces (BCI) is achieved either by voluntary regulation of EEG activity, e.g. sensorimotor rhythms (SMR), or by eliciting well defined responses of the brain to stimulation, e.g. the P300 event-related potential in an oddball paradigm. Many subjects, patients with neurological diseases as well as healthy volunteers, achieve high a level of accuracy. However, in a considerable amount of participants, no stimulus or task related EEG pattern can be detected by the BCI or the bitrate is not high enough to allow for meaningful and satisfactory communication despite statistically significant accuracy [1]. Thus, recently, few studies were dedicated to elucidate factors that contribute to BCI performance and to define predictors of BCI control. Available results indicate four aspects: (1) individual characteristics of the BCI user, (2) characteristics of the BCI, (3) type of feedback and instruction, and (4) the BCI-controlled application. An integration of these aspects leads to a neuro-bio-psychological, data analytical, and ergonomical model of BCI-control.

1 Introduction

It is well known that BCIs are not suitable for all users. Depending on the BCI approach, in up to one third of the participants the BCI is unable to detect classifiable task related EEG patterns. Consequently, these participants cannot be provided with a BCI-controlled application. The causes for this "BCI-inefficiency" have not yet been satisfactorily described. Few studies exist that explicitly investigated the predictive value of internal (user related) and external (BCI related) factors on BCI performance [2, 3]. In this paper, we present an effort to integrate the existing knowledge about factors that influence BCI performance into a model of BCI-control. From now more than 40 years of research on BCI including early studies on neurofeedback we suggest four different aspects that contribute to BCI-control: (1) individual characteristics of the BCI user which include physiological, neurological and psychological factors; (2) characteristics of the BCI which comprises hardware and software components, (3) type of feedback and instruction including feedback modality, presentation within each modality and the instruction that is provided prior to the training; and (4) the BCI-controlled application which can range from simple two-choice to multiple-choice paradigms for communication, neuroprosthetic control, and clinical applications.

2 Methods

We restricted our literature search in Pubmed on non-invasive BCI studies in humans and the authors' databases. A Pubmed search with the terms "BCI AND performance AND (influence OR predictor)" yielded 12 hits only; "BCI AND performance" resulted in 352 publications from

1996 to 2011. Focus was then laid on papers that included regression or correlation analysis between performance (accuracy/bitrate) and independent variables or that manipulated independent variables to investigate the effect on performance. Online BCI-control was preferred.

3 Results

Our proposed model of BCI control is composed of four categories that all influence performance with a BCI (Figure 1). Feedback, instruction and application in turn influence the psychological and physiological state of the user.

3.1 Individual Characterstics

Performance in the SMR-based BCI was predicted to 25 % by the μ-peak during rest with eyes open [2]. High correlations of performance were found with activation in prefrontal areas known to be involved in allocation of attentional resources and working memory and with premotor and supplementary motor areas [4]. Early performance predicted later performance in BCI based on slow cortical potentials (SCP) and SMR [3, 5]. Self-regulatory capacity as measured with heart rate variability correlated moderately with performance in the P300 BCI [6]. Generally speaking, individuals with neurological disease perform worse than healthy subjects probably due to brain damage, specifically in patients with neurodegenerative disease [3,7]. Motivational (fear of incompetence, challenge) and emotional states moderately influence BCI performance in patients and in healthy volunteers [8,9]. The ability to concentrate, visuo-motor coordination and internal locus of control of reinforcement were also found to moderately influence performance in P300 and SMR-based BCI [8–11].

3.2 Characteristics of the BCI

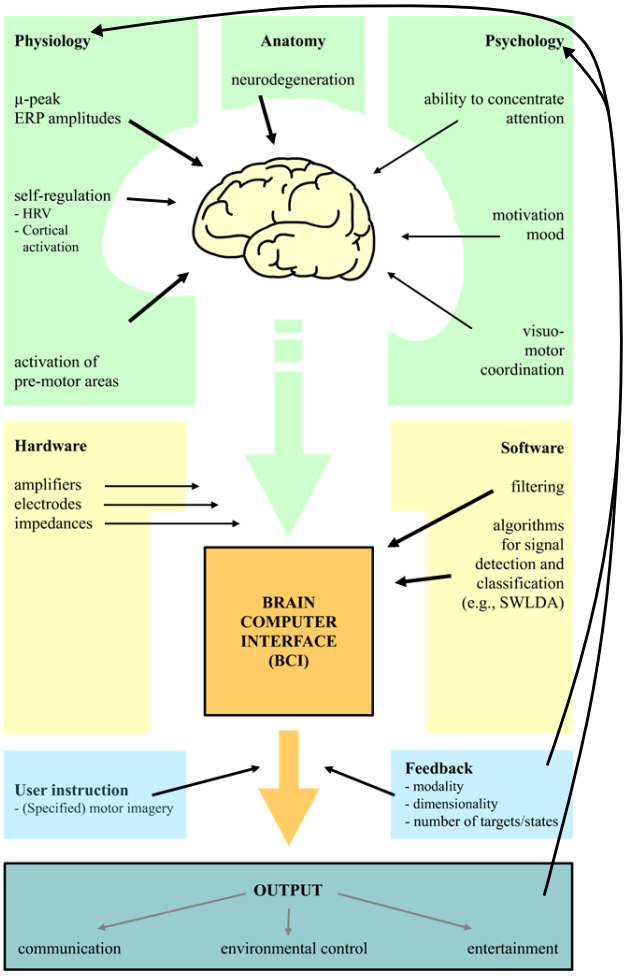
Classification performance depends on various factors such as the experimental paradigm, signal quality, the number of electrodes used, as well as feature extraction and classification algorithms. Systematic studies with a meaningful number of participants are rare, in particular with online BCI-control. For SMR-based BCIs methods including the Common Spatial Pattern (CSP) technique are often preferred and scored well in the BCI Competitions [12]. In ERP-based BCIs, Linear Discriminant Analysis (LDA) applied to spatio-temporal features was found to perform well when regularization (by stepwise-LDA [13] or Shrinkage-LDA [14]) is employed. BCIs based on visual evoked potentials (SSVEP, motion-VEP) typically perform well with very few electrodes (e.g, [15]), while P300-BCIs require about 6 electrodes for good performance [16], and for SMR-based BCIs a number of 17 electrodes has been suggested [17]. For practical applications and home use, also one-channel BCI systems, allied with careful selection of EEG parameters and training, proved efficient [18].

3.3 Feedback and Instruction

For the activation of appropriate areas in the brain related to motor imagery the instruction was found to play a major role; kinesthetic motor imagery proved more effective than visual, goal-directed motor imagery lead to more focused hand area activation [19, 20] and so-called quasi-movements were better detectable than ordinary motor imagery [21]. In general, motor imagery imposed less workload than other mental tasks [22]. Number of targets or items for selection presented by the BCI, type of stimulation (intensity, inter-stimulus-intervals) and modality (visual/auditory/tactile) affected performance; generally, participants perform better when stimuli, feedback and results are presented visually [23–25]. Feedback has a strong impact on task-related brain activity [26]. The presentation of feedback may also influence performance, e.g. virtual reality environment seems to positively affect outcome [27].

3.4 Application

The more complex the application the more difficult becomes BCI-control, because more information has to be processed [28, 29]. For example, internet browsing or painting lead to a drop in performance as compared to simple spelling with the P300 [30, 31].



4 Conclusion

BCI performance is influenced by many internal and external characteristics of the user and the BCI and thus, multiple toeholds for improvement exist. Further knowledge of these characteristics will foster the introduction of BCI operability 'standards'. For a full understanding of differences in performance all aspects have to be considered in a full online, closed-loop set up.

Figure 1: A model of BCI-control comprised of 4 aspects: individual characteristics, BCI characteristics, feedback and instruction, BCI-controlled application. Colours serve for distinction of categories only. Boldness of arrows indicate possible strength of influence on BCI control.

References

[1] A. Kübler, and N. Birbaumer. Brain-computer interfaces and communication in paralysis: extinction of goal directed thinking in completely paralysed patients? *Clin Neurophysiol*, 119: 2658–2666, 2008

[2] B. Blankertz, C. Sannelli, S. Halder, E. M. 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] A. Kübler, N. Neumann, B. Wilhelm, T. Hinterberger, and N. Birbaumer. Predictability of brain-computer communication. *Int J Psychophysiol*, 18(2-3): 121–129, 2004

[4] S. Halder, D. Agorastos, R. Veit, E. M. Hammer, S. Lee, B. Varkuti, M. Bogdan, W. Rosenstiel, N. Birbaumer, and A. Kübler. Neural mechanisms of brain-computer interface control. *Neuroimage*, 55: 1779–1790, 2011.

[5] E. Weber, A. Köberl, S. Frank, and M. Doppelmayr. Predicting successful learning of SMR neurofeedback in healthy participants: methodological considerations. *Appl Psychophysiol Biofeedback*, 36: 37–45, 2011.

[6] C. Vogele et al.. Effects of cardiac autonomic balance on performance in P300 brain-computer interface (BCI). *Clin Neurophysiol*, submitted.

[7] 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. *Clin Neurophysiol*, 117(3): 531–537, 2006

[8] S. C. Kleih, F. Nijboer, S. Halder, and A. Kübler. Motivation modulates the P300 amplitude during brain-computer interface use. *Clin Neurophysiol*, 121: 1023–1031, 2010.

[9] F. Nijboer, N. Birbaumer, and A. Kübler. The influence of psychological state and motivation on brain-comoputer interface performance in patients with amyotrophic lateral sclerosis - a longitudinal study. *Front Neuroscience*, 4, 2010.

[10] W. Burde, and B. Blankertz. Is the locus of control of reinforcement a predictor of brain-computer interface performance? *Proceedings of the 3rd International Brain-Computer Interface Workshop and Training Course 2006*, 76–77, Verlag der Technischen Universität Graz, 2006.

[11] C. E. Lakey, D. R. Berry, and E. W. Sellers. Manipulating attention via mindfulness induction improves P300-based brain-computer interface performance. *J Neural Eng*, 8: 025019, 2011.

[12] B. Blankertz, K. R. Müller, D. Krusienski, G. Schalk, J. R. Wolpaw, A. Schlögl, G. Pruftscheller, J. del R. Millàn, M. Schröder, and N. Birbaumer. The BCI competition III: validating alternativ approachs to actual BCI problems. *IEEE Trans Neural Syst Rehabil Eng*, 14(2): 153–159, 2006.

[13] 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.

[14] B. Blankertz, S. Lemm, M. S. Treder, S. Haufe, and K. R. Müller. Single-trial analysis and classification of ERP components - a tutorial. *Neuroimage*, 2011, in press.

[15] T. Liu, L. Goldberg, S. Gao, and B. Hong. An online brain-computer interface using non-flashing visual evoken potentials. *J Neural Eng*, 7(3): 036003, 2010.

[16] D. Krusienski, E. W. Sellers, D. J. McFarland, T. M. Vaughan, and J. R. Walpow. Toward enhanced P300 speller performance. *J Neurosci Methods*, 167: 15–21, 2008.

[17] C. Sannelli, T. Dickhaus, S. Halder, E. M. Hammer, K. R. Müller, and B. Blankertz. On optimal channel configurations for SMR-based brain-computer interfaces. *Brain Topogr*, 23(2): 186–193, 2010.

[18] G. R. Müller-Putz, V. Kaiser, T. Solis-Escalante, and G. Pfurtscheller. Fast set-up asynchronous brain-switch based on detection of foot motor imagery in 1-channel EEG. *Med Biol Eng Compout*, 48: 229–233, 2010.

[19] C. Neuper, R. Scherer, M. Reiner, and G. Pfurtscheller. Imagery of motor actions: differential effects of kinesthetic and visual-motor mode of imagery in single-trial EEG. *Brain Res Cogn Brain Res*, 25(3): 668–677, 2005.

[20] F. Pichiorri, F. De Vico Fallani, F. Cincotti, F. Babiloni, M. Molinari, S. C. Keih, C. Neuper, A. Kübler, and D. Mattia. Sensorimotor rhythm-based brain-computer interface training: the impact on motor cortical responsiveness. *J Neural Eng*, 8: 025020, 2011.

[21] V. V. Nikulin, F. U. Hohlefeld, A. M. Jacobs, and G. Curio. Quasi-movements: a novel motor-cognitive phenomenon. *Neuropsychologia*, 46(2): 727–742, 2008.

[22] E. V. C. Friedrich, R. Scherer, K. Sonnleitner, and C. Neuper. Impact of auditory distraction on user performance in a brain-computer interface driven by different mental tasks. *Clin Neurophysiol*, 2011, in press.

[23] 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: 617–625, 2009.

[24] E. W. Sellers, D. Krusienski, D. J. McFarland, T. Vaughan, and J. Walpow. A P300 event-related potential brain-computer interface (BCI): the effects of matrix size and inter stimulus interval on performance. *Biol Psychol*, 73: 242–252, 2006.

[25] A. M. Brouwer, and J. B. F. van Erp. A tactile P300 brain-computer interface. *Front Neuroscience*, 4(19): 036003, 2010.

[26] C. Neuper, R. Scherer, S. Wriessnegger, and G. Pfurtscheller. Motor imagery and action observation: modulation of sensorimotor brain rhythms during mental control of a brain-computer interface. *Clin Neurophysiol*, 120: 239–247, 2009.

[27] J. Gruzelier, A. Inoue, R. Smart, A. Steed, and T. Steffert. Acting performance and flow state enhanced with sensory-motor rhythm neurofeedback comparing ecologically valid immersive VR and training screen scenarios. *Neurosci Lett*, 480: 112–116, 2010.

[28] A. Kübler, N. Neumann, J. Kaiser, B. Kotchoubey, T. Hinterberger, and N. P. Birbaumer. Brain-computer communication: self-regualtion of slow cortical potential for verbal communication. *Arch Phys Med Rehabil*, 82(11): 1533–1539, 2001.

[29] G. Pfurtscheller, R. Leeb, C. Keinrath, D. Friedmann, C. Neuper, C. Guger, and M. Slater. Walking from thought. *Brain Res*, 1071: 145–152, 2006.

[30] J. I. Münssinger, S. Halder, S. C. Kleih, A. Furdea, V. Raco, A. Hösle, and A. Kübler. Brain painting: first evaluation of a new brain-computer interface application with ALS patients and healthy volunteers. *Front Neuroscience*, 4(0), 2010.

[31] C. Zickler et al. BCI and assistive technology: BCI as input channel in a standard assistive technology ICT-software. *J Clin Neurosci*, submitted.

SMR EEG-BCI Aptitude in Healthy Subjects Varies with the Integrity of Corpus Callosum White Matter

B. Varkuti[1,*], S. Halder[1,2,*], M. Bogdan[2,3], A. Kübler[4], W. Rosenstiel[2], R. Sitaram[1], N. Birbaumer[1,5]

[1]Institute of Medical Psychology and Behavioral Neurobiology, University of Tübingen, Gartenstr. 29, 72074 Tübingen, Germany
[2] Wilhelm-Schickard Institute for Computer Science, University of Tübingen, Sand 13, 72076 Tübingen, Germany
[3] Computer Engineering, University of Leipzig, Johannisgasse 26, 04103 Leipzig, Germany
[4] Department of Psychology I, University of Würzburg, Marcusstr. 9-11, 97070 Würzburg, Germany
[5] Ospedale San Camillo, Istituto di Ricovero e Cura a Carattere Scientifico, Laboratorio di Neuroscience comportamentale, Via Alberoni 70, 30126 Venezia, Italy
[*] **These authors contributed equally to the paper**

balint.varkuti@medizin.uni-tuebingen.de

Abstract

Healthy subjects differ in terms of their ability to control Electroencephalography (EEG) based Brain-Computer Interfaces (BCIs), but the reasons for this remain unclear. Deviating from the classical perspective of analyzing psychological intermediary variables we investigated whether the structural trait of inter-hemispheric white matter integrity differs between subjects classified as low or high aptitude users. For this purpose 17 healthy subjects participated in a Diffusion Tensor Imaging (DTI) measurement as well as in a sensori-motor rhythm (SMR) EEG-BCI session. We extracted the individual Corpus Callosum (CC) Fractional Anisotropy (FA) values from each subject and compared them across the two user groups. The difference is significant (two-sample t-test, $p = 0.0187$) and the values differ in the expected direction (CC mean FA low aptitude users 0.53 vs. high aptitude users 0.58). We conclude that individual inter-hemispheric connectivity traits might influence one's ability to master SMR control.

1 Introduction

The successful use of some electroencephalography (EEG) based Brian-Computer Interfaces (BCIs) depends critically on an individuals ability to voluntarily alter one's own brain activity in a directed manner and with relatively precise timing. Psychological variables such as attention, mood, training and the right mental strategy or technique [1], equipment related variables such as sufficient Signal-to-Noise-Ratio (SNR) during the EEG recording and adequate processing as well as functional variables such as individual neural plasticity and normality of the EEG spectral properties [2] and the modulability of BCI-critical features are relevant to successful BCI control.
Recent studies have indicated a critical role of central white matter for the spectral properties of an individuals EEG [3], underlining the significance of certain fiber bundles for appropriate inter-regional conduction velocity in intra-cerebral neural signaling. While the fiber bundles in the healthy brain form always some standard connections (the connectome core so to speak), there are connections that vary between subjects and which are an individual trademark of our brain - just like a fingerprint. While most of these connections are intra-hemispheric and located within diverse association fiber bundles (e.g. Superior Longitudinal Fascicle, Capsula Externa), roughly one third of all macroscopic inter-regional anatomical connections in the brain are inter-hemispheric and traverse through a few localized commissural structures such as the Corpus Callosum (CC) [4], which could be termed the bottlenecks of inter-hemispheric connectivity.

It is possible to control a BCI by performing motor imagery of different body parts such as the hands or feet which causes Event-Related-Desynchronization/Event-Related-Synchronization of the sensori-motor rhythm (SMR) also referred to as μ-rhythm [5, 6].
SMR modulation through means of motor imagery is known to be a complex task [7], involving multiple cortical and subcortical structures and the timely integration of their respective neural computations into a coherent and continuous mental construct accessible to consciousness and voluntary control. The overall integration and the time-critical signaling processes within this bi-hemispheric task-dependent neural network must therefore strongly rely on commissural connections. Such anatomical connections are structures which are known to be strongly affected in patients with Amyotrophic Lateral Sclerosis (ALS), who are the group which we mainly target with EEG-BCI applications [8].
We tested the hypothesis that the structural substrate of inter-hemispheric anatomical connectivity has an influence on individual BCI aptitude, by splitting a group of SMR EEG-BCI users into a high and a low aptitude group and comparing the structural properties of their main commissural connection (namely the Fractional Anisotropy of the CC).

2 Methods

2.1 Participants

Twenty healthy participants (seven female and 13 male, mean age 24.5 years SD 3.7, range 19–36) took part in the study which was approved by the Internal Review Board of the Medical Faculty, University of Tübingen. All participants had no prior experience with SMR BCIs, had no history of neurological diseases and were German native speakers. The data from three subjects had to be excluded due to image and scanner artifacts. A detailed overview of EEG-BCI performance, performed imagery tasks, methodic details of the split into low and high aptitude performers and demographic data of all participants can be found in [7].

2.2 EEG-BCI Setup and Definition of Low and High Aptitude Users

The EEG recording was performed using four 32 channel Brainamp DC amplifiers manufactured by Brainproducts, Munich, Germany. A 128 channel cap manufactured by Easy Cap, Munich, Germany was used. For details on EEG acquisition please see [9]. Each of the participants performed a single EEG-BCI session. This included measurements in which the participant had to perform motor execution, observation and imagery. Three imagery measurements which totaled 75 trials of three classes (left hand, right hand, foot) were used to calibrate the feedback parameters. The calibration trials had a duration of 8 seconds, in 4 of which the participant performed the task. After calibration three feedback measurements were performed in which the participant had to use the two classes which showed the highest discriminability in the calibration data to control a cursor [9]. In total 300 feedback trials were performed (150 per class). The accuracy of the participants achieved in these 300 trials was used to categorize them into high and low aptitude SMR EEG-BCI users by a median split.

2.3 MRI Recording

The MRI measurements were performed in a Siemens Magneton Trio Tim 3T whole body scanner using a standard 12 channel head coil. Each subject participated in a Diffusion Tensor Imaging (DTI) measurement ($1.8 \times 1.8 \times 6.5$ mm voxels, 5 mm gap, $TR = 3$ sec, $TE = 93$ ms, $FoV = 1150 * 1150$, Flip Angle = $90\,^{\circ}$, 20 transversal slices, 128×128 voxels per slice, 20 diffusion directions, b-value = $1000\,\mathrm{s/mm^2}$).

Subject	Gender	Age	Aptitude group	CC mean FA
1	M	23	high	0.6071
2	F	22	high	0.5815
3	M	22	low	0.5450
4	F	23	low	0.5183
5	F	22	low	0.6241
6	M	23	high	0.5567
7	F	30	high	0.5476
8	F	25	low	0.5379
9	F	26	low	0.5168
10	M	25	low	0.5458
11	M	22	low	0.4735
12	F	20	low	0.5557
13	M	22	high	0.5926
14	M	28	low	0.5250
15	M	23	high	0.5970
16	M	24	high	0.6321
17	M	19	high	0.5698

Table 1: Sample characteristics, EEG-BCI performance and CC FA.

2.4 Analysis of DTI Data

Voxel-wise Fractional Anisotropy (FA) values indicate the orientedness of the diffusion tensor - the more ellipsoid shaped a tensor is the more the surrounding tissue is restricting isotropic diffusion. The normalization parameters to Montreal Neurological Institute standard space were estimated for the FA-image aligned B0-image of the DTI sequence using SPM8. The normalization parameters were then inverted to warp the standard label image of the ICBM-DTI-81 Atlas [10] into each participants original DTI space. For region 4 of the ICBM-DTI-81 Atlas (Body of CC) the median and mean of FA values for all voxels with an FA value above 0.25 was extracted and saved in the participant/CC FA value table.

3 Results

CC mean FA does not significantly differ for Gender ($p = 0.64$) nor does it correlate with Age ($r = -0.24$, $p = 0.35$) but significantly differs between high and low aptitude users (two-sample t-test, $p = 0.0187$). The result remains significant if CC median FA is used instead of mean FA ($p = 0.0136$). High aptitude users have a higher CC mean FA value (0.58 vs 0.53), indicating higher fiber density, larger fiber diameter, higher myelination, stronger directional similarity of fibers or overall higher structural integrity (or a combination of all) for their CC region.

4 Discussion

A covariation of structural brain traits with BCI aptitude can be interpreted in various ways. Structural trait differences between high and low aptitude users could indicate the existence of a latent variable, that influences both BCI aptitude as well as the structural integrity of the CC. This variable could be a psychological or neurobiological trait that we have not yet been able to measure.
On the other hand CC structural integrity could directly influence the EEG-BCI use on the physical level of signal generation through restricting timely inter-hemispheric signaling and integration.
Further the fiber integrity in the CC region could influence the processes related to neural plasticity during the learning of BCI control, by affecting the timing of conditioning processes and thereby altering the outcome of learning after equal training periods. The higher structural integrity of

the CC could further indicate a higher number of inter-hemispheric anatomical connections in the high aptitude group that might correspond to superior inter-hemispheric communication and integration in general.
It is possible that these findings are related to the diminished BCI aptitude of late-stage ALS patients, as multiple central white matter structures (such as the CC) are affected in ALS.

5 Conclusion

While our first finding might indicate an influence of commissural connectivity on one's ability to use a SMR EEG-BCI, further data is necessary to evaluate the overall impact of variations in inter-hemispheric connectivity on learning processes. We plan to extend our approach to incorporate further measures of individual structural brain traits into a more comprehensive model and to explore the role of inter-hemispheric connectivity in EEG-BCI use.

References

[1] F. Nijboer, N. Birbaumer, and A. Kübler. 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, 2010.

[2] K. P. Thomas, C. Guan, L. C. Tong, and A. P. Vinod. *A Study on the impact of spectral variability in brain-computer interface*. IEEE, May 2010.

[3] P. A. Valdés-Hernández, A. Ojeda-González, E. Martínez-Montes, A. Lage-Castellanos, T. Virués-Alba, L. Valdés-Urrutia, and P. A. Valdes-Sosa. White matter architecture rather than cortical surface area correlates with the EEG alpha rhythm. *Neuroimage*, 49(3):2328–2339, 2010.

[4] G. Gong, Y. He, L. Concha, C. Lebel, D. W. Gross, A. C. Evans, and C. Beaulieu. Mapping anatomical connectivity patterns of human cerebral cortex using in vivo diffusion tensor imaging tractography. *Cerebral Cortex*, 19(3):524, 2009.

[5] C. Neuper, R. Scherer, M. Reiner, and G. Pfurtscheller. Imagery of motor actions: differential effects of kinesthetic and visual-motor mode of imagery in single-trial EEG. *Brain Res Cogn Brain Res*, 25(3):668–677, 2005.

[6] A. Kübler, F. Nijboer, J. Mellinger, T. M. Vaughan, H. Pawelzik, G. Schalk, D. J. McFarland, N. Birbaumer, and J. R. Wolpaw. Patients with ALS can use sensorimotor rhythms to operate a brain-computer interface. *Neurology*, 64(10):1775–1777, 2005.

[7] S. Halder, D. Agorastos, R. Veit, E. M. Hammer, S. Lee, B. Varkuti, M. Bogdan, W. Rosenstiel, N. Birbaumer, and A. Kübler. Neural mechanisms of brain-computer interface control. *NeuroImage*, January 2011.

[8] N. Filippini, G. Douaud, C. E. Mackay, S. Knight, K. Talbot, and M. R. Turner. Corpus callosum involvement is a consistent feature of amyotrophic lateral sclerosis. *Neurology*, 75(18):1645–52, November 2010.

[9] B. Blankertz, C. Sannelli, S. Halder, E. M. Hammer, A. Kübler, K. R. Müller, G. Curio, and T. Dickhaus. Neurophysiological predictor of SMR-based BCI performance. *NeuroImage*, 51(4):1309–1303, March 2010.

[10] S. Mori, K. Oishi, H. Jiang, L. Jiang, X. Li, K. Akhter, K. Hua, A. V. Faria, A. Mahmood, R. Woods, A. W. Toga, G. B. Pike, P. R. Neto, A. Evans, J. Zhang, H. Huang, M. I. Miller, P. van Zijl, and J. Mazziotta. Stereotaxic white matter atlas based on diffusion tensor imaging in an ICBM template. *NeuroImage*, 40(2):570–82, April 2008.

Motivation Influences Performance in SMR-BCI

S. C. Kleih[1], A. Riccio[2], D. Mattia[2], V. Kaiser[3], E. V. C. Friedrich[3,4], R. Scherer[3], G. Müller-Putz[3], C. Neuper[3,4], A. Kübler[1,5]

[1]Department of Psychology, University of Würzburg, Germany
[2]Laboratory of Neuroelectrical Imaging and Brain Computer Interface, Fondazione Santa Lucia, Rome, Italy
[3]Institute for Knowledge Discovery, Graz University of Technology, Austria
[4]Department of Psychology, University of Graz, Austria
[5]Institute of Medical Psychology and Behavioural Neurobiology, University of Tübingen, Germany

sonja.kleih@uni-wuerzburg.de

Abstract

In this study we investigated the effect of motivation on performance when using a Brain-Computer Interface (BCI) based on sensorimotor rhythms (SMR). After pooling the data acquired with four different SMR-BCI protocols in one sample of $N = 41$ participants, we found a positive correlation between the motivational components "challenge" and "incompetence fear" and accuracy in percent correct responses. As motivation seems to have a positive effect on SMR-BCI performance, we recommend to enhance motivation if possible and to monitor motivation in BCI settings.

1 Introduction

Brain-Computer Interfaces (BCIs) based on sensorimotor rhythms (SMR) allow for muscle independent device control [1]. However, there are considerable inter-individual differences in BCI performance [2] which might, to some extent, be attributed to the influence of psychological variables such as motivation or emotion or both. Only recently, research started monitoring and manipulating motivation to investigate its relation to BCI performance [3–6]. First results indicate that motivational factors positively influence BCI performance not only in healthy participants [3,6] but also in patients with amyotrophic lateral sclerosis [5]. As all studies which investigated the effect of motivation on BCI performance included small sample sizes only (highest number of participants $N = 16$ in [3]), it might be argued that intra-individual performance differences are emphasized disproportionally high. This renders the evaluation of a bigger sample necessary to judge the effect of psychological variables on BCI performance. In this study, we therefore pooled the data of several SMR-BCI studies to further establish the relation between motivation and performance. We hypothesized that higher motivation would be associated with higher BCI performance.

2 Methods

2.1 Participants and Procedure

Data from $N = 41$ healthy participants were analyzed. The mean age of participants was 25.05 (range 19–33). EEG data was acquired at the authors' institutes using different study protocols. All participants were measured with Ag/AgCl electrodes. A subgroup of $N = 14$ participants underwent 2 screening sessions and had to perform different mental tasks (word association, mental subtraction, motor imagery and spatial navigation) in 9 feedback sessions. In this protocol 29

electrodes were used for data acquisition. The 4-class classification was based on the method of common spatial pattern and realized by means of Fisher´s linear discriminant analysis with majority voting (for further details, please see [7]). Another subgroup consisted of 15 participants who were asked to imagine right hand movement (RH) or movement of both feet (FE) indicated by an arrow either pointing to the right or down on a computer screen. Based on this data, 2–3 bandpower features were selected to train an LDA-classifier. This classifier was used to distinguish between the two different classes (RH vs. FE) by generating an output of positive or negative class labels with a continuous distance value that was proportional to the probability of the actual classification belonging to one class. The classifier provided feedback for the training sessions. Feedback had the form of a fluid cloud which could change color (RH/positive class label = green, FE/negative class label = blue), intensity (continuous distance value) and direction (RH = right, FE = down; integrated classifier output). In this protocol 3 bipolar electrodes were used for the measurements. The third subgroup consisted of 8 participants who had to steer a cursor either to the top or bottom target area on a computer screen by imagining hand or foot movement. In this protocol 61 electrodes were used and the participants completed between 3 and 5 training sessions. The fourth group of participants ($N = 5$) had to steer a cursor into a target area presented on the bottom or top area of the computer screen. In this protocol 32 electrodes were used. All participants were provided with online feedback. Feedback was either provided by a usual moving cursor, a fluid cursor or a bar graph. As we were interested in the effects of motivation on BCI performance independent from the training protocol and we could not find significant differences between subgroups with regards to performance, we pooled all data into a sample of $N = 41$ participants and investigated the first session only, so that all participants experienced the same amount of training. Motivation was assessed prior to the BCI session.

2.2 Questionnaires

To measure motivation, we used the Questionnaire for Current Motivation (QCM) which was adapted for BCI by Nijboer and colleagues, (2008, [4]) and a visual analogue scale (VAS). The QCM-BCI comprises 18 items divided into four subscales incompetence fear, mastery confidence, interest and challenge which have to be rated on a seven point Likert-scale. The consistency coefficients for the 4 subscales are between alpha = .66 and alpha = .90 (Cronbach). The VAS ranged from 0 to 10 on a 10 cm long horizontal line (0 = extremely unmotivated and 10 = extremely motivated). Participants had to indicate with a vertical line their current motivation.

2.3 Data Analysis

As there was a range of training sessions from 1 up to 9 sessions depending on the study protocol, we correlated motivation and performance of the first session only. As accuracy in percent correct responses (performance) and motivation ratings were not normally distributed, Spearman's rank correlation was calculated.

3 Results

Average performance in session 1 was 62.73 percent correct responses online (SD±26.71). We found significant positive correlations between incompetence fear and accuracy ($\rho = .37$, $p < .05$, see Figure 1 A) and between challenge and accuracy ($\rho = .50, p < .01$, see Figure 1 B).

4 Discussion and Conclusion

In this study we demonstrated a relation between motivation and SMR BCI performance in a bigger sample of participants. Correlations were positive and moderate for incompetence fear and moderate to high for challenge. These results are in line with those of Nijboer and colleagues (2008, [4]) who interpreted their correlations such that the fear to fail might boost performance

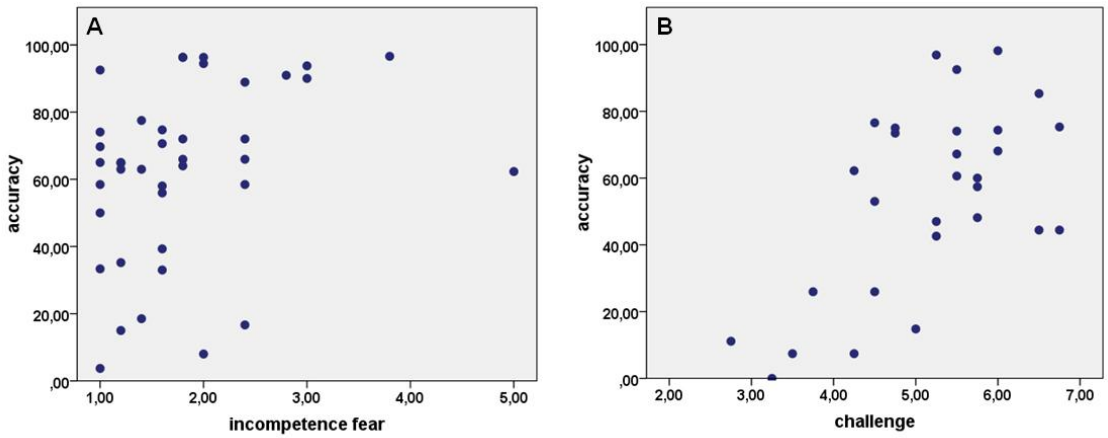


Figure 1: Accuracy in percent correct responses is depicted as a function of motivation ratings. A. incompetence fear B. challenge.

because participants may be highly engaged in the task to prevent a drop in performance. Both, feeling challenged and being afraid to fail may increase the attention allocated to the BCI task. We thus, may speculate that learning an SMR-BCI task is facilitated by increased motivation. In a next step it would be interesting to manipulate motivation for example by offering extra rewards per correct trial in the BCI task. This possibly would also increase attention to the task and allow in turn, for better performances as concentration could be sustained for a longer period of time. In conclusion, we recommend monitoring and if possible enhancing motivation as it may influence BCI performance positively.

5 Acknowledgments

We thank Tobias Kaufmann for his support. This work is supported by the European ICT Programme Project FP7-224631. 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] A. Kübler and K. R. Müller. *An introduction to brain computer interfacing (book chapter).* MIT Press, 2007.

[2] B. Blankertz, C. Sannelli, S. Halder, E. M. Hammer, A. Kübler, K. R. Müller, G. Curio, and T. Dickhaus. Neurophysiological predictor of SMR-based BCI performance. *NeuroImage*, 51(4):1303–1309, July 2010. PMID: 20303409.

[3] F. Nijboer, A. Furdea, I. Gunst, J. Mellinger, D. J. McFarland, N. Birbaumer, and A. Kübler. An auditory brain-computer interface (BCI). *Journal of Neuroscience Methods*, 167(1):43–50, January 2008. PMID: 17399797.

[4] F. Nijboer, E. W. Sellers, J. Mellinger, M. A. Jordan, T. Matuz, A. Furdea, S. Halder, U. Mochty, D. J. Krusienski, T. M. Vaughan, J. R. Wolpaw, N. Birbaumer, and A. Kübler. A P300-based brain-computer interface for people with amyotrophic lateral sclerosis. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 119(8):1909–1916, August 2008. PMID: 18571984.

[5] F. Nijboer, N. Birbaumer, and A. Kübler. 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, 2010. PMID: 20700521.

[6] S. C. Kleih, F. Nijboer, S. Halder, and A. Kübler. Motivation modulates the P300 amplitude during brain-computer interface use. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 121(7):1023–1031, July 2010. PMID: 20188627.

[7] E. V. C. Friedrich, R. Scherer, K. Sonnleitner, and C. Neuper. Impact of auditory distraction on user performance in a brain-computer interface driven by different mental tasks. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, April 2011. PMID: 21511526.

Long-Term BCI Training for Grasp Restoration in a Patient Diagnosed with Cervical Spinal Cord Injury

V. Kaiser[1], A. Kreilinger[1], G. R. Müller-Putz[1], C.Neuper[1,2]

[1]Institute for Knowledge Discovery, Graz University of Technology, Graz, Austria
[2]Department of Psychology, University of Graz, Graz, Austria

vera.kaiser@tugraz.at

Abstract

Increasing the independency and quality of life of cervical spinal cord injured persons by restoring grasping function using a BCI-controlled hand neuroprosthesis is one of the recent goals in research. The aim of this study is to evaluate the changes of motor cortical patterns, which could be used for controlling such a device, in the course of a long-term BCI training in a 20-year-old SCI patient. The patient was regularly trained for 17 months, with different BCI paradigms. The activation patterns of the motor cortex changed over time and got stable after 18 feedback sessions and she reached classification accuracies up to 85 % already at the third session. Although such a performance is acceptable, it is still a long way towards a useful and reliable signal for controlling a hand neuroprosthesis to restore grasp.

1 Introduction

A cervical spinal cord injury (SCI) at a high level leads to a loss of motor, sensory, and autonomous functions and therefore to a life-long dependency on caregivers. Increasing the patient's independency and quality of life by restoring grasping function is one of the goals in research at our lab. Ten years ago, Pfurtscheller et al. [1] already proposed to use brain oscillations to control a hand orthosis with a brain-computer interface (BCI). A strategy to voluntarily modulate brain oscillation is the imagination of limb movements. This motor imagery (MI) induces a relative decrease in band power, known as event-related desynchronization (ERD), which can be measured with electroencephalography (EEG) of the motor cortical areas [2]. In healthy persons these patterns are usually quite prominent and easy to classify. In SCI patients the patterns of motor cortex activation are altered in dependence of the severeness of the paralysis (level of injury; complete or incomplete; time since injury) [3, 4], which makes a training of distinguishable patterns necessary. The aim of the present study is to evaluate the changes of motor cortical patterns in the course of a long-term BCI training using the example of a single SCI patient. We expected to find distinguishable activation patterns over the course of time resulting from regular BCI training.

2 Methods

2.1 Patient

The 20 years old female patient was diagnosed with tetraplegia of C4/C5 (ASIA, A) since a traumatic event in January 2009. Besides the loss of functionality of the lower extremities, as well as the loss of vegetative functions below the lesion, she has lost motor and sensory function of the hand and parts of the arm. Residual muscle activity is restricted to muscles controlling the shoulder and partly to muscles controlling the elbow joint.

2.2 EEG Recordings and Data Processing

The EEG was recorded from 15 channels (C3, Cz, C4, each surrounded by 4 electrodes). Reference and ground electrode were placed on the left and right mastoid. The signals were acquired with a g.USBamp amplifier (Guger Technologies, Austria) with 512 Hz sample rate, 0.5 Hz high-pass and 100 Hz low-pass filter and an additional 50 Hz notch filter. To assess changes in the frequency domain for each class, ERD maps for frequency bands between 4 and 40 Hz were calculated, using overlapping frequency bands with 2 Hz band width and 1 Hz step size [5]. The statistical significance of the ERD values was determined by applying a t-percentile bootstrap algorithm [6] with a significance level of $\alpha = 0.01$. Based on this data 1–3 relevant band power features were selected by checking for relevant ERD patterns in the time-frequency maps. These features were used to train linear discriminant analysis (LDA) classifiers to discriminate between two different classes (2-MI-class BCI: right hand MI vs. feet MI; 1-MI-class BCI: right hand MI vs. rest). Classifiers provided feedback for the training sessions and were updated after every session.

For evaluating performance, classification accuracies were calculated for every feedback session. In case of the 1-MI-class BCI training with continuous feedback the number of true (TP) and false positives (FP) per minute and the ratio between TP and FP per minute were calculated. In addition a t-test for dependent measures was performed to check for significant differences between the number of TP and FP per minute.

2.3 Training Procedure

The training started in October 2009 and is still ongoing. It started with nine screening sessions without feedback (right hand vs. feet MI), twelve 2-MI-class BCI training sessions with feedback (right hand vs. feet MI), again four screening sessions without feedback (right hand MI vs. rest) and 14 1-MI-class BCI training sessions with feedback (right hand MI vs. rest).
A training session usually took place once a week in the patient's home and consisted of 8 runs with 15 trials per class, one trial lasting 7 s. According to the needs of the patient the interval between the sessions was sometimes more than one week (longest interval between sessions 6 weeks) and some sessions were terminated with less than 8 runs (minimum 6 runs).
In the course of the training different kinds of visual feedback were given, due to new developments and to avoid tediousness and keep motivation high (see Figure 1). For the 2-MI-class BCI training at first feedback was given in form of the basket game [7] (see Figure 1 A.). Here, a ball falls downwards with constant speed and its horizontal alignment is controlled by the LDA distance. The task was to guide the ball into a basket on the left (feet MI) or right (right hand MI) bottom of the screen. Due to a clear drop in performance after six sessions, feedback was changed to a fluid cursor (see Figure 1 B.). Here, the classifier distinguished between right hand MI and feet MI by generating an output of positive or negative classlabels with a continuous distance value that was proportional to the probability of the actual classification belonging to one class. The fluid cursor could change color (right hand MI = green, feet MI = blue), intensity (continuous distance value) and direction (right hand MI = right, feet MI = down; integrated classifier output). The task was to move the cursor in the cue-guided direction.
For the 1-MI-class BCI training feedback was given in form of a car, driving along a two-lane street with obstacles appearing on the left lane and coins appearing on the right lane (see Figure 1 C.). The task was to avoid the obstacles and collect as many coins as possible. The car stayed on the left lane if rest was detected and changed to the right lane if right hand MI was detected. This paradigm was a first step towards asynchronous control, as in this case feedback was given constantly, meaning throughout the whole run the car could drive to the left or the right lane, depending on which state was detected (rest or right hand MI).
In addition to the BCI training, the patient was provided with a functional electrical stimulation (FES) device (MotionStim8, KRAUTH + TIMMERMANN GmbH, Germany) to preserve and build up the muscles of the upper extremities. This FES training started after eight months of BCI training and is needed to obtain reasonable grip strength for using a hand neuroprosthesis.

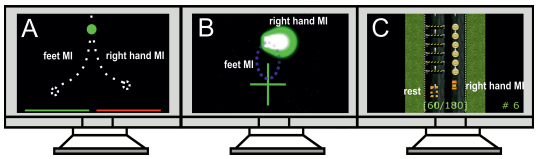


Figure 1: Different feedback types: A. Basket game. B. Fluid Cursor. C. Car game.

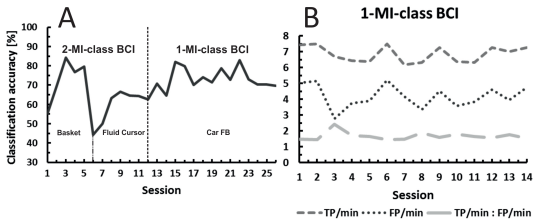


Figure 2: A. Results for 2-MI-class BCI. B. Results for 1-MI-class BCI.

3 Results

3.1 2-MI-class BCI

The patient showed ERD from 15 to 25 Hz over Cz and C4 during MI right from the beginning of the training, but the pattern for feet and right hand MI did not differ from each other. In the course of the screening sessions this pattern changed to ERS at 25 to 30 Hz over C3 for right hand MI and ERS at 10 to 12 Hz over C3 and C4 for feet MI. Both MI were accompanied by an ERD at 20 to 25 Hz over Cz. With these patterns, after five months of screening, a reasonable differentiation of the two classes was possible, so in the following sessions feedback was given. The left part of Figure 2 A. shows the classification accuracies for the 2-MI-class BCI sessions with feedback. The interval between session 5 and 6 was six weeks, which resulted in a clear decrease in classification accuracy, due to a change in the ERD/ERS pattern. The ERS at C3 and C4 did not appear anymore. As the performance did not reach the level it had before this long interval, in further sessions the training paradigm was changed from two MI classes to one MI class (right hand MI) against rest.

3.2 1-MI-class BCI

After four screening sessions for right hand MI against rest a weak ERD from 15 to 20 Hz appeared over C3, Cz and C4. In the course of the following feedback sessions this pattern turned out to be instable and appeared in variable strength at the different channels. After eight feedback sessions (session 19 in Figure 2 A.) an ERD at 15 to 20 Hz appeared over Cz, which stayed stable over the following sessions, only varying in strength. The right part of Figure 2 A. shows the classification accuracies for the 1-MI-class BCI sessions with feedback. Performance varied between 70 and 80 % accuracy. Figure 2 B. shows the true (TP) and false positives (FP) per minute and the ratio between them for all 1-MI-class BCI sessions. The results of the t-test showed that the number of TP per minute ($M = 6.85$; $SD = 0.51$) is significantly higher ($t_{(13)} = 21.88$; $p < .0001$) than the number of FP per minute ($M = 4.18$; $SD = 0.73$). The ratio between TP and FP per minute is quite stable in the course of feedback training.

4 Discussion

BCI training in patients is challenging. Although the patient's SCI was only nine months before start of the BCI training, it took about five months of screening to find distinguishable MI patterns. Due to altered motor cortex activation patterns found in SCI patients [3, 4] it seems to be harder to establish discriminative patterns. After starting feedback training a clear increase in performance, accompanied by a prominent ERS during MI, occurred after only three sessions. After a long interval between session 5 and 6, due to illness of the patient, the performance dropped down to chance level, which shows that the pattern was not stable. Nevertheless, performance increased again in the course of training but did not exceed 85 % by now. Although the performance is acceptable and the number of TP per minute is significantly higher than the number of

FP per minute, it is still a long way towards a useful and reliable signal for controlling a hand neuroprosthesis to restore grasp. Motor cortical activation in SCI patients seems to change with BCI training, but only short-termed. Enzinger et al. [8] reported on an SCI patient who persistently trained MI via BCI over 8 years and established a stable activation pattern, which is similar to the pattern observed with actual movement in healthy controls. To achieve this goal, a regular training, perhaps with shorter session intervals, is needed. As patients do have a life besides supporting BCI research, this is not easy to realize. Longer session intervals, or shorter sessions due to discomfort or a conflicting schedule cannot be avoided. In addition to a regular training, the use of additional information, e.g. error potentials, might increase performance to a reasonable level.
Also the design of feedback is important. Although the patient always reported to be highly motivated, she started to get bored by playing the same game over and over again and got confused by the amount of information of the fluid cursor feedback. An appealing and entertaining feedback, which is adapted to the needs of the patient, is necessary to avoid tediousness and frustration.
Concluding we can state that if we want to design tools to improve motor functions and quality of life in SCI patients, we should be aware that gaining a reasonable control of the BCI takes longer as in healthy subjects.

5 Acknowledgments

We thank University of Glasgow for providing the fluid cursor feedback and the car game. This work is supported by the European ICT Programme Project FP7-224631. 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] G. Pfurtscheller, C. Guger, G. Müller, G. Krausz, and C. Neuper. Brain oscillations control hand orthosis in a tetraplegic. *Neuroscience Letters*, 292:211–214, 2000.

[2] G. Pfurtscheller and F. H. Lopes da Silva. Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical Neurophysiology*, 110:1842–1857, 1999.

[3] K. Gourab and B. D. Schmit. Changes in movement-related beta-band EEG signals in human spinal cord injury. *Clinical Neurophysiology*, in press, 2010.

[4] G. Pfurtscheller, P. Linortner, R. Winkler, G. Korisek, and G. R. Müller-Putz. Discrimination of motor imagery-induced EEG patterns in patients with complete spinal cord injury. *Computational Intelligence and Neuroscience*, 2009:104180, 2009.

[5] B. Graimann, J. E. Huggins, S. P. Levine, and G. Pfurtscheller. Visualization of significant ERD/ERS patterns in multichannel EEG and ECoG datas. *Clinical Neurophysiology*, 113:43–47, 2002.

[6] A. C. Davison and D. V. Hinkley. *Bootstrap methods and their application*. Cambridge University Press, 1997.

[7] G. Krausz, R. Scherer, G. Korisek, and G. Pfurtscheller. Critical decision-speed and information transfer in the "Graz Brain-Computer Interface". *Applied Psychophysiology and Biofeedback*, 28:233–240, 2003.

[8] C. Enzinger, S. Ropele, F. Fazekas, M. Loitfelder, F. Gorani, T. Seifert, G. Reiter, C. Neuper, G. Pfurtscheller, and G. Müller-Putz. Brain motor system function in a patient with complete spinal cord injury following extensive brain-computer interface training. *Experimental Brain Research*, 190:215–223, 2008.

Detecting and Interpreting Responses to Feedback in BCI

M. Perrin[1,2], E. Maby[1,2], R. Bouet[1,2], O. Bertrand[1,2], J. Mattout[1,2]

[1]INSERM U1028,CNRS UMR5292, Lyon Neuroscience Research Center, Lyon, F-69000, France
[2]University Lyon 1, Lyon, F-69000, France

margaux.perrin@inserm.fr

Abstract

Most BCI aim at implementing a real-time and bi-directional interaction between a man's brain and a computer. Current EEG-based systems still suffer from poor performance, including the well-established P300-speller whose application in patients remain difficult [1]. Some recent studies have explored the possible use of EEG responses to the computer action or feedback in order to improve the performance of this closed-loop. On a few subjects, those studies have highlighted the usefulness of this approach, although single-trial detection of errors proved difficult and not always beneficial. In this study, a large group of healthy subjects performed a fully controlled experiment involving our recent OpenViBE implementation of the P300-speller. We addressed two questions regarding the neurophysiological responses to feedback. First, we evaluated a new approach to error detection, assuming stationarity of these responses over trials. Second, we challenged this stationarity hypothesis and studied the modulations of those responses over time. We conclude that it is possible to detect errors with high accuracy, such that even high spelling performance could be further improved. Importantly, we also show that responses to feedback do change in the course of the brain-computer interaction, in ways that reflect learning and attention mechanisms.

1 Introduction

The P300-speller is a well-known BCI paradigm aiming at enabling non-responsive but conscious patients to communicate. However, current performance still need to be improved [1]. A promising avenue is the development of methods to detect errors and apply corrections in real-time. A couple of recent studies have addressed this challenge. In particular, in the context of the P300-speller, Visconti and colleagues have tested several classifiers to discriminate between correct and incorrect trials using feedback-evoked potentials. Results were encouraging but the ensuing benefit to online performance remained fairly limited [2]. We have explored two complementary approaches to overcome this current limitation. On the one hand, we propose a new signal processing approach to detect errors. This approach mimics the methods we use to detect P300 responses online in our implementation of the P300-speller [3]. On the other hand, we investigated the variations of these feedback-related responses, which in the context of cognitive neuroscience studies, have been previously described as reflecting prediction error [4]. We hypothesized that a BCI paradigm would elicit the same type of signals, whose modulations would not only reflect feedback valence but also learning and attention mechanisms. Therefore, we designed a P300-speller based experiment where both error rate and error predictability were manipulated. In addition, we monitored the (true) performance of the subject, which can be seen as a trial-by-trial measure of attention.

2 Methods

2.1 Data Acquisition

Twenty-two healthy adults took part in the experiment (15 men, mean age: 24.2±1.9 years). All of them signed an informed consent approved by the local Ethical Committee. We used a 6×6 matrix made of letters (A–Z), digits (1–9) and an additional symbol (_). Matrix rows and columns

were flashed alternatively for 80 ms, at a pace of 200 ms. A complete cycle of 12 stimulations is referred to as a repetition. We used three repetitions per character to spell. Those parameters were chosen to guarantee an average performance of 80 % accuracy over subjects [3]. The whole experiment included one training followed by three test sessions. The former consisted in spelling 25 random characters and each test session was made of 24 5-letter words. In both cases, target letters were defined prior to the experiment and indicated to the subject at the top of the screen. Subjects were instructed to stare at the target and count how many times it was flashed. The feedback was provided one second after the last flash. This feedback was manipulated so that, on average, each subject obtained errors in 20 % of the trials. Whereas participants believed that their brain activity was causing letter selection, the sequence of errors was fully controlled: one error in 50 % of the words, two errors in 25 % of them and no error in the remaining 25 %. Besides, the true performance of the subject was monitored online.
EEG data were recorded from 32 electrodes placed on standard positions according to the extended 10-10 system and referenced to the nose. Data were digitized at 1000 Hz, bandpass filtered between 0.1 and 20 Hz, down-sampled at 100 Hz and corrected for eye movements.

2.2 Data Analysis

Error Detection To detect responses to error feedbacks, we implemented the same signal processing methods that we use online to detect the P300. 600 ms long epochs starting at feedback onset were extracted. A spatial filter based on the xDAWN algorithm [5] was applied before entering a two-class Naïve Bayes classifier. The spatial filter and classifier parameters were learned from a subset of test session trials. Performance in error detection was then evaluated as a function of subset size.
In each subject and at the group level, we computed the sensitivity and specificity of error detection. We also evaluated the accuracy θ of a correction that would consider the second best guess from the speller, each time an error has been correctly detected. This enables us to predict performance Pc after online correction, and to compare it with initial performance P:

$$Pc = \frac{(1-P) \times \text{Sensitivity} \times \theta}{P \times \text{Specificity}} \tag{1}$$

Modulations of the Neurophysiological Responses to Feedback

By differentiating the EEG responses to good and bad feedbacks, we could observe two central and large deflections that are well documented in the (non-BCI) literature, namely the feedback-related negativity (FRN) and the P300. The former was observed between 250 and 310 ms, and the latter between 380 and 460 ms. We studied the modulations of those two waveforms by attention (true correct vs. true incorrect trials), and predictability (expected vs. unexpected feedbacks). For the latter, we used the two following subsets of trials:

- Responses to a first or unique error in a word (expected) and responses to a second error in a word (unexpected).
- Responses to the last letter of the word, when it was a good feedback, in words containing one error (expected) and words with no error (unexpected)

The main effects and interactions between feedback valence and attention, and between feedback valence and predictability, were thus studied thanks to two separate designs. To account for the unbalance nature of those designs, we used a mixed-effect model based on single-trial data from all subjects and test for significant differences between conditions.

3 Results

Error Detection Figure 1 (A) shows the sensibility, specificity, initial and corrected performance at the group level, as the size of training set increases. Specificity is always very high, above 90 %, while sensitivity is very sensitive to the size of the training set. Based on a training set of 100 letters, Figure 1 (B) depicts the individual performance and values of parameter θ, before and after correction. Importantly, this shows that most subjects would benefit from such a correction, although their initial average accuracy is already very high (73.72±17.85).

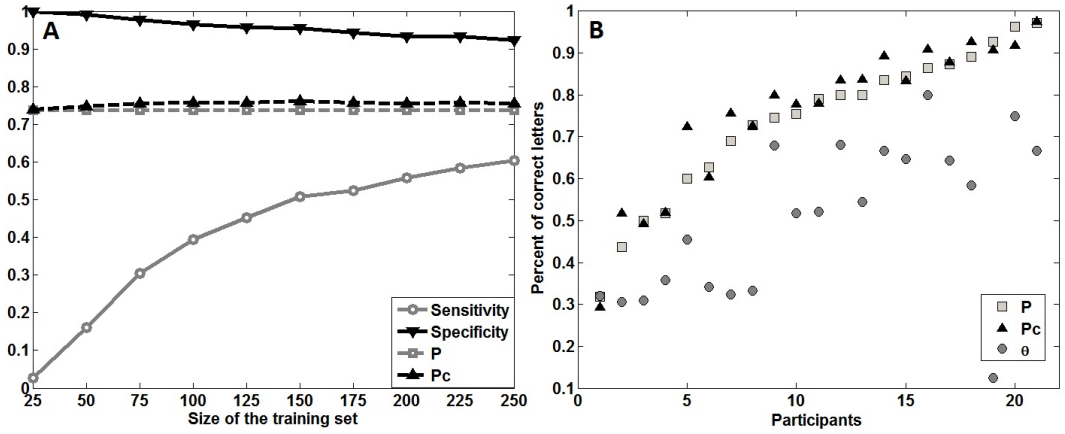


Figure 1: Performance in error detection and correction at the group (A) and subject level (B).

Modulations of the Neurophysiological Responses to Feedback

As shown on Figure 2 (A), we observe a significant interaction between feedback valence and attention as reflected by the true performance (FRN: $F_{1,7730} = 6.37$, $p = 0.012$; P300: $F_{2,7732} = 10.37$, $p < 0.0001$). Enhanced attention yields larger waves, as well as a faster return to baseline. Figure 2 (B) shows the significant interaction between feedback valence and feedback predictability or subject's expectation on the FRN ($F_{1,2516} = 6.11$, $p = 0.014$) and a main effect of feedback predictability on the P300 ($F_{1,2519} = 7.54$, $p = 0.006$). Surprise yields larger FRN and P300 responses, which might reflect prediction error.

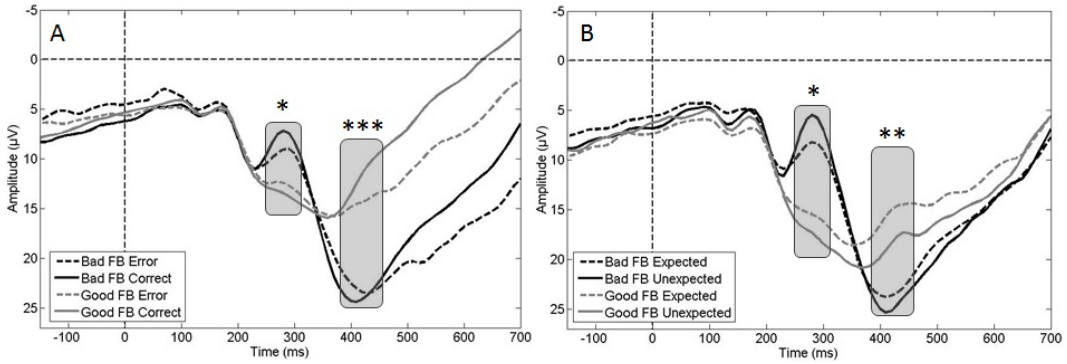


Figure 2: Average time series from a centro-parietal ROI (Cz, Pz, Fc1, Fc2, Cp1, Cp2) of responses to good and bad feedbacks and their modulations by attention (A) and expectation (B). FRN and P300 time windows appear in light grey ($* : p \leq 0.05$, $** : p \leq 0.01$).

4 Discussion

In a large group of healthy subjects, we studied the neurophysiological responses to feedback in a BCI context. Although we relied on three repetitions per character to spell only, the true online performance of the subjects were very high and confirmed our previous results [3].
First, we could see a clear difference between responses to bad and good feedback, dominated by the well-known FRN and P300 waveforms, respectively. Thanks to this strong and reliable difference, we were able to detect single-trial responses to errors with very high specificity and high sensitivity depending on the size of our training set. Importantly, we show that such a correction, implemented online, could improve the spelling performance. This further demonstrates the quality of the signal processing methods we implemented for feature extraction and classification.
In a second offline analysis, we revealed that those feedback responses are not fully stationary but depend on both the attention level of the subject and the predictability of the feedback. This is in line with previous experimental findings and recent theories in electrophysiology and cognitive neuroscience [4, 6]. We also observed that θ, the correction accuracy increases with performance, which might be also related to attention. This might explain why correction remains effective, even in subjects whose initial performance is very high (Figure 1 (A)).

5 Conclusion

Our results show that BCI feedback-related responses relate to known cognitive control signals, and that differences between bad and good feedbacks can be detected in single trials, allowing for automatic correction. Moreover, the modulation of those signals by expectation and ongoing performance can be readily interpreted as reflecting learning and attention mechanisms, respectively. This speaks in favor of new approaches that would incorporate expected individual changes in neurophysiological signals to adapt online and improve BCI performance.

6 Acknowledgments

This work is supported by the French ANR-DEFIS program under grant ANR-09-EMER-002 in the context of the CoAdapt project.

References

[1] E. Maby, M. Perrin, D. Morlet, P. Ruby, O. Bertrand, S. Ciancia, N. Gallifet, J. Luauté, and J. Mattout. Evaluation in a locked-in patient of the OpenViBE P300-speller. *Proceedings of the 5th International Brain-Computer Interface Workshop*, 2011.

[2] G. Visconti, B. Dal Seno, M. Matteucci, and L. Mainardi. Automatic recognition of error potentials in a P300-based brain-computer interface. *Proceedings of the 4th International Brain-Computer Interface Workshop and Training Course*, pages 238–243, 2008.

[3] E. Maby, G. Gibert, P. E. Aguera, M. Perrin, O. Bertrand, and J. Mattout. The OpenViBE P300-speller scenario: a thorough online evaluation. *Human Brain Mapping Conference*, 2010.

[4] G. Hajcak, J. S. Moser, C. B. Holroyd, and R. F. Simons. It's worse than you thought: the feedback negativity and violations of reward prediction in gambling tasks. *Psychophysiology*, 44(6):905–12, 2007.

[5] B. Rivet, A. Souloumiac, G. Gibert, V. Attina, and O. Bertrand. Sensor selection for P300 speller brain computer interface. *Proceedings of ESANN*, 2009.

[6] K. Friston. A theory of cortical responses. *Philos Trans R Soc Lond B Biol Sci*, 360(1456):815–36, 2005.

Modality-Specific Affective Responses and Their Implications for Affective BCI

C. Mühl[1], A. M. Brouwer[2], N. C. van Wouwe[2], E. L. van den Broek[1,3,4], F. Nijboer[1], and D. K. J. Heylen[1]

[1]Human Media Interaction, University of Twente, Enschede, The Netherlands
[2]TNO Behavioural and Societal Sciences, Soesterberg, The Netherlands
[3]Human-Centered Computing Consultancy, Vienna, Austria
[4]Karakter University Center, Radboud University Medical Center Nijmegen, The Netherlands

c.muehl@gmail.com

Abstract

Reliable applications of multimodal affective brain-computer interfaces (aBCI) require a detailed understanding of the processes involved in emotions. To explore the modality-specific nature of affective responses, we studied neurophysiological responses of 24 subjects during visual, auditory, and audiovisual affect stimulation and obtained their subjective ratings. Coherent with literature, we found modality-specific responses in the EEG: parietal alpha power decreases during visual stimulation and increases during auditory stimulation, whereas more anterior alpha power decreases during auditory stimulation and increases during visual stimulation. We discuss the implications of these results for multimodal aBCI.

1 Introduction

Affective brain-computer interfaces (aBCI) aim to provide an intelligent and affective interface, using real-time processing and classification of (single trial) EEG signals. As such, aBCI belong to a new class of affective interfaces. This class relies on the assumption that the EEG reflects affective responses, which will be challenged in the current article.

aBCI studies found EEG signals, especially in the frequency domain, to be informative regarding the affective state [1–3]. Cognitive theories of affect (e.g., the component process theory [4]) suggest that the brain is involved in responses to affective stimulation through both a self-monitoring component and mechanisms of increased cognitive processing of relevant affective stimuli, comparable to the effects of attention [5]. Especially cognitive processes might depend on the modality through which emotional states are induced (e.g. by visual or auditory affective stimuli), because the respective modality-specific sensory processes are supposed to have their own neural substrates, with correlates of their activity in the alpha band [6, 7].

Here, we explore the stimulus-specific cognitive responses during visual, auditory, and audiovisual affective stimulation. According to theories of sensory processing [6], we expect parietal and fronto-central alpha activity to (negatively) correlate to visual and auditory processing, respectively. Specifically, visually induced affect should lead to a decrease of parietal alpha power, whereas auditorily induced affect should lead to a decrease of fronto-central alpha power. Audiovisual stimulation should lead to decreases in both regions.

2 Methods

Participants We collected subjective ratings, and measured EEG from 12 female and 12 male participants (mean age: 28 years, range: 19–39), all but one right-handed.

Condition	Visual Modality (IAPS)		Auditory Modality (IADS)		Audio-visual Modality	
	Valence	Arousal	Valence	Arousal	Valence	Arousal
(1) Unpleasant low arousal	2.58(0.60) *1.99(0.79)*	5.24(0.54) *4.98(1.60)*	3.05(0.51) *2.84(0.81)*	5.81(0.43) *4.55(1.73)*	- *2.28(0.78)*	- *5.08(1.56)*
(2) Unpleasant high arousal	2.26(0.34) *1.97(0.83)*	6.50(0.22) *5.73(1.84)*	2.70(0.51) *2.55(0.77)*	6.79(0.31) *5.32(1.64)*	- *2.02(0.90)*	- *5.82(1.61)*
(3) Pleasant low arousal	7.53(0.44) *6.88(0.70)*	5.26(0.52) *5.24(1.45)*	7.09(0.43) *6.17(0.71)*	5.59(0.39) *4.97(1.50)*	- *6.69(0.81)*	- *5.37(1.50)*
(4) Pleasant high arousal	7.37(0.31) *6.29(0.93)*	6.67(0.38) *5.50(1.60)*	7.19(0.44) *6.28(0.71)*	6.85(0.39) *5.67(1.62)*	- *6.40(0.69)*	- *5.92(1.65)*
(5) Neutral low arousal	4.92(0.54) *4.52(0.64)*	5.00(0.51) *4.41(1.24)*	4.82(0.44) *4.86(0.60)*	5.42(0.42) *4.10(1.29)*	- *4.54(0.60)*	- *4.38(1.38)*

Table 1: Mean (std) ratings for the emotion conditions (norm [8, 9] and *participants'* ratings).

Experimental Setup To manipulate the affective state, a binary division was made between both 40 IAPS [8] and 40 IADS [9] stimuli on both the valence (i.e., pleasant and unpleasant) and arousal (i.e., low and high) dimension. Additionally, a neutral class (i.e., low arousal, neutral valence) with 10 stimuli for each stimulus modality was constructed, see also Table 2 [1]. The mean valence and arousal values of the conditions were matched as good as possible between the emotion conditions, and between modalities. For the audio-visual conditions, visual and auditory stimuli of the same emotion conditions were paired with special attention to match the content of picture and sound (e.g., pairing of "aimed gun" picture and "gun shot" sound).

The stimuli were presented in 3 separate blocks: visual, auditory, audio-visual, each preceded by a resting period of 60 seconds. Their order was balanced (Latin Square) across participants. Within each of these 3 modality blocks, the 5 emotion conditions were presented in pseudo-randomized order to ensure a balancing across participants. Each block was preceded by a 20 second resting period to minimize carry-over effects. Within each block, auditory and/or visual stimuli were presented in a randomized order, each for 6 seconds and separated by 2 seconds. A fixation cross was present in the center of the screen to minimize eye movements.

Before the experiment, participants gave their informed consent and their demographics. They were seated 90 cm away from a monitor and speakers. EEG was recorded with a Biosemi ActiveTwo Mk II system, with 512 Hz sampling frequency. 32 active silver-chloride electrodes were placed according to the 10-20 system. For later artifact rejection, the electrooculogram (EOG) was measured by 2 electrodes attached to the outer canti of the eyes and 2 attached below and above the left eye. Before the start, participants were instructed to avoid movements and to fixate at all times the fixation cross. After the experiment, all stimuli were presented once more to jog their memory while they rated the experienced valence, arousal, and dominance.

Data Extraction and Analysis The EEG data was resampled to 256 Hz, referenced to common average, and high-pass filtered with a 1 Hz FIR filter. EOG artifacts were removed by the AAR toolbox in EEGlab. For each emotion condition and modality, the mean alpha power (8–13 Hz) was extracted for stimulation and resting periods, using Welch's method with 256-point Hanning windows. Subsequently, the power was averaged over the *parietal* (P3,Pz,P4) and *fronto-central* (FC1,Fz,FC2) regions of interest, and a natural log transform was applied. The power during each condition was baselined by subtracting that of the preceding resting period.

For analysis, repeated measures ANOVAs were used. First, the affect manipulation was verified, using the valence and arousal ratings. Second, modality-specific effects on parietal and fronto-central alpha power were analyzed, using the data of the neutral versus the averaged emotion conditions for each region. Where appropriate, Greenhouse-Geisser correction has been applied. Partial eta-squared η_p^2 is reported as effect strength measure.

[1] (1) IAPS: 2141,2205,2278,3216,3230,3261,3300,9120,9253,8230 IADS: 280,250,296,03,241,242,730,699,295,283
(2) IAPS: 2352.2,2730,3030,6360,3068,6250,8485,9050,9910,9921 IADS: 600,255,719,284,106,289,501,625,713,244
(3) IAPS: 1811,2070,2208,2340,2550,4623,4676,5910,8120,8496 IADS: 226,110,813,221,721,820,816,601,220,351
(4) IAPS: 4660,5629,8030,8470,8180,8185,8186,8200,8400,8501 IADS: 202,817,353,355,311,815,415,352,360,367
(5) IAPS: 2220,2635,7560,2780,2810,3210,7620,7640,8211,9913 IADS: 724,114,320,364,410,729,358,361,500,425

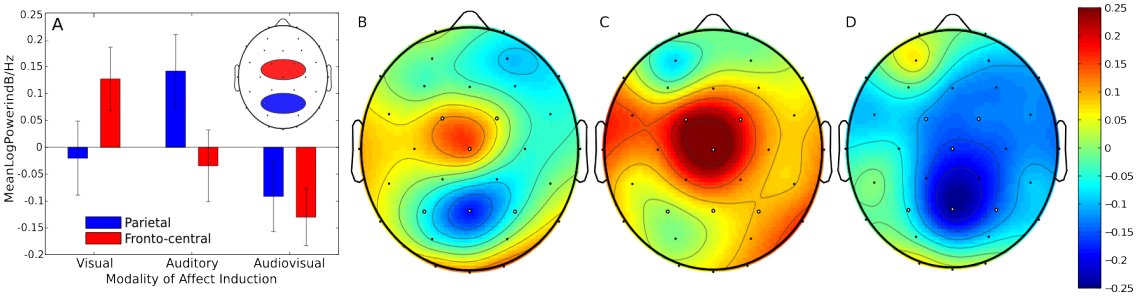


Figure 1: The parietal and fronto-central alpha power (whiskers: SEM) during visual, auditory, and audio-visual affect, averaged over all subjects that entered the respective analyses (A); and the 2nd level contrasts (emotion – neutral) of visual – auditory (B), visual – audio-visual (C), audio-visual – auditory affect (D), showing modality-specific responses averaged over all subjects.

3 Results

A 3(modality) × 5(emotion) ANOVA on the valence ratings showed a main effect of both emotion ($F(4,92) = 246.100$, $p < 0.001$, $\eta_p^2 = 0.915$) and modality ($F(2,46) = 6.057$, $p = 0.005$,$\eta_p^2 = 0.208$). A 3(modality) × 3(valence) ANOVA indicated the successful manipulation of emotional valence ($F(2,46) = 264.100$, $p < 0.001$, $\eta_p^2 = 0.920$), with an effect of modality ($F(2,46) = 7.078$, $p = 0.002$, $\eta_p^2 = 0.235$) due to more positive valence ratings for auditory stimuli. An interaction effect indicates less extreme ratings for auditory stimuli ($F(4,92) = 14.813$, $p < 0.001$, $\eta_p^2 = 0.392$) and, hence, a weaker efficacy. A similar pattern was found in a 3(modality) × 5(emotion) ANOVA on the arousal ratings, showing a main effect of emotion ($F(4,92) = 12.588$, $p < 0.001$, $\eta_p^2 = 0.354$), and of modality ($F(2,46) = 9.177$, $p < 0.001$, $\eta_p^2 = 0.285$). A 3(modality) × 2(arousal) ANOVA showed higher ratings for arousing conditions ($F(1,23) = 27.180$, $p < 0.001$, $\eta_p^2 = 0.542$), and an effect of modality ($F(2,46) = 9.344$, $p < 0.001$, $\eta_p^2 = 0.289$), due to lower arousal for auditory stimuli (see Table 2).

For each EEG analysis we excluded all cases showing outliers ($> 1.5 \times$ *interquartile range*), resulting in $N = 20$ for both regions of interest. The parietal alpha power showed an interaction of emotion and modality ($F(2,38) = 3.373$, $p = 0.045$, $\eta_p^2 = 0.151$). T-tests contrasting the emotion effects (i.e., emotion – neutral) of visual, auditory, and audio-visual conditions, revealed a significant difference between audio-visual and auditory affect ($t = -2.651$, $p = 0.016$), and a trend toward a difference between visual and auditory affect ($t = -2.093$, $p = 0.052$) . The fronto-central alpha power showed a significant interaction ($F(2,38) = 4.891$, $p = 0.013$, $\eta_p^2 = 0.205$). T-tests of the emotion effects showed a significant difference between audio-visual and visual affect ($t = -3.155$, $p = 0.005$), and a trend toward a difference between visual and auditory affect ($t = 1.817$, $p = 0.085$). So, on the one hand, parietal alpha decreases during visual and audio-visual and increases during auditory affective stimulation. On the other hand, fronto-central alpha decreases during auditory and audio-visual and increases during visual stimulation (see Figure 1).

4 Discussion

The responses to visual and auditory affective stimulation (see Figure 1) suggest the activation and deactivation of regions potentially associated with the appropriate and non-appropriate sensory modality, respectively [7]. The marked increases of alpha power over the regions associated with the processing of the non-appropriate sensory modality are in line with the putative role of alpha oscillations in the gating of sensory input [6]. Accordingly, the observed results might be interpreted as general cognitive responses to affective stimuli, rather than primary correlates of affect or feeling.

The current analysis does not allow a conclusion about the emotion-specificity of the observed responses. Especially for aBCI it would be of interest, if such modality-specific responses allow

for a discrimination of valence and arousal. Given the opposing nature of visually and auditorily induced affective responses, a classifier trained on a certain modality will be limited in its capability to generalize its classification to affective states induced via other modalities. Assuming a general cognitive nature of the observed affective responses (e.g., the orienting response [5]), a classifier might be prone to confuse purely cognitive responses with affective responses. This poses a general problem for the community of aBCI: How can affective neurophysiological responses be adequately used for affective state discrimination, given their context-specific nature? Strategies like the restriction to specific contexts or the use of additional contextual information (e.g., the occurrence of external stimulation or peripheral physiological affective responses) might help to deal with the ambiguity of neurophysiological responses. However, context-specificity of correlates of affect demands a critical assessment of the generalizability and specificity of the classified EEG activity.

5 Conclusion

We showed that neurophysiological responses to visual and auditory affective stimulation are differing in terms of parietal and fronto-central activations in the alpha band. This has implications for the generalization and specificity of affect classifiers. Furthermore, these results open up exciting research questions for aBCI, such as: Is it possible to detect the modality of the trigger of an affective response, the object or event that induced a given affective state?

Acknowledgments The authors gratefully acknowledge the support of the BrainGain Smart Mix Programme of the Netherlands Ministry of Economic Affairs and the Netherlands Ministry of Education, Culture, and Science. Further, we thank H. Gürkök for his help improving the article.

References

[1] G. Chanel, J. J. M. Kierkels, M. Soleymani, and T. Pun. Short-term emotion assessment in a recall paradigm. *International Journal of Human-Computer Studies*, 67(8):607–627, 2009.

[2] C. A. Frantzidis, C. Bratsas, C. L. Papadelis, E. Konstantinidis, C. Pappas, and P. D. Bamidis. Toward emotion aware computing: An integrated approach using multichannel neurophysiological recordings and affective visual stimuli. *IEEE Transactions on Information Technology in Biomedicine*, 14(3):589–597, 2010.

[3] S. Koelstra, C. Mühl, M. Soleymani, J. S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras. DEAP: A database for emotion analysis using physiological signals. *IEEE Transaction on Affective Computing. [in press]*.

[4] D. Sander, D. Grandjean, and K. R. Scherer. A systems approach to appraisal mechanisms in emotion. *Neural Networks*, 18(4):317–352, 2005.

[5] P. Vuilleumier. How brains beware: neural mechanisms of emotional attention. *Trends in Cognitive Sciences*, 9(12):585–594, 2005.

[6] O. Jensen and A. Mazaheri. Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Frontiers in human neuroscience*, 4, 2010.

[7] G. Pfurtscheller and F. H. Lopes da Silva. Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical Neurophysiology*, 110(11):1842–1857, 1999.

[8] P. J. Lang, M. M. Bradley, and B. N. Cuthbert. International Affective Picture System (IAPS): Technical manual and affective ratings. Technical report, University of Florida, Center for Research in Psychophysiology, Gainesville, Fl, USA., 1999.

[9] M. M. Bradley and P. J. Lang. The International Affective Digitized Sounds (IADS-2): Affective ratings of sounds and instruction manual. Technical report, University of Florida, Center for Research in Psychophysiology, Gainesville, Fl, USA., 2007.

Predicting Performance in a Hybrid SSVEP/ERD BCI for Continuous Simultaneous Cursor Control

B. Z. Allison[1], C. Brunner[1], S. Grissmann[1],
C. Altstätter[1], I. Wagner[1,2], C. Neuper[1,2]

[1]Institute for Knowledge Discovery, Laboratory of Brain-Computer Interfaces,
Graz University of Technology, Krenngasse 37, 8010 Graz, Austria
[2]Department of Psychology, University of Graz, Universitätsplatz 2/III, 8010 Graz, Austria

allison@tugraz.at

Abstract

Many groups rely on offline screening procedures to estimate users' performance with an online BCI and optimize signal processing parameters. We assess whether an offline screening procedure based on conventional simple ERD and SSVEP BCIs can accurately predict performance with a hybrid ERD/SSVEP. The offline screening procedure was not an effective predictor under the circumstances described here.

1 Introduction

Brain-computer interface (BCI) systems allow users to send information directly through brain signals produced by specific voluntary mental activities. Different BCI approaches are categorized according to the types of mental activities and corresponding brain signals that users perform to send messages or commands. Two common approaches are ERD BCIs [1–3], which rely on changes in synchronous brain oscillations associated with imagined movement, and SSVEP BCIs, in which users focus on rapidly oscillating visual stimuli to elicit corresponding oscillations over visual areas [3, 4].

BCI studies typically report that performance varies considerably across participants. Hence, there have been numerous efforts to predict which participants will perform best with different BCI approaches. Such efforts could help find the right BCI for each user, avoid the inconvenience and disappointment of an unsuccessful BCI session, identify good participants for research studies, and explore causes of poor performance [2, 5, 6]. Many groups attempt to predict performance through a "screening" procedure, during which participants perform tasks typical of an online BCI. For example, participants might imagine hand or foot movements [1–3] or focus on flickering stimuli [3, 4].

Recent research has focused on hybrid BCI systems, which may combine a BCI with another BCI. This creates an emerging question: what are the best ways to predict performance with hybrid BCIs? Here, we explore this question by assessing the predictive value of offline screening procedures, similar to those routinely used with conventional "simple" BCIs, which we extended to a two dimensional hybrid BCI based on simultaneous ERD and SSVEP activity. A secondary goal was to assess both ERD and SSVEP activity in a within-participants design.

2 Methods

Subjects were recruited as follows. One subject (AP3) was chosen because he exhibited exceptional ERD and good SSVEP in prior studies. AP3 had approximately 20 hours of BCI experience, mostly with our prior efforts to combine ERD and SSVEP BCIs for one-dimensional control [3]. This

subject was chosen to assess performance with an elite subject and thereby allow comparison to research efforts that routinely report two dimensional control in only an elite subset of people (e.g. [7]). Three other subjects had a few hours of experience with these earlier BCI systems, but had not exhibited exceptional performance. All other 21 subjects were recruited via word-of-mouth or via advertisements on a website of the Graz University of Technology. None of these subjects had prior experience with any BCI. Ten subjects participated in an online recording session on another day. We recorded from 31 electrode channels distributed over the head, primarily over central areas. We also recorded EOG activity from four monopolar channels. We used a sampling frequency of 256 Hz, bandpass filter of 0.5–100 Hz, and a notch filter at 50 Hz, recorded with three g.USB amplifiers (g.tec). Data were recorded and analyzed using a Matlab Simulink system (Mathworks).

Each participant participated in six runs to train the classifiers. Each run had 40 trials, with 10 trials per class (up, down, left, right). Each offline run lasted about seven minutes, and the whole session of data collection (6 runs) lasted about 45 minutes. This offline session yielded 240 trials (60 trials per class).

The participant viewed a screen with a fixation cross in the center and eight grey boxes. During these offline sessions, the four targets in the corners were never designated as targets. There was also a red LED that oscillated at 8 Hz affixed to the left side of the monitor, and another red LED that oscillated at 13 Hz on the right side of the monitor. Two seconds after each trial began, one of the grey boxes in the top, right, left, or bottom was cued by changing to green. The participant then had to perform a mental activity corresponding to the target. The participant was asked to focus on the 8 or 13 Hz LEDs in response to left or right target cues, respectively. Participants were asked to imagine left hand movement when the top target was cued, and imagine moving both feet when the bottom target was cued. Participants performed the designated mental activity throughout the trial. Each trial lasted 8 seconds and there was a 3–5 second delay between each trial, during which the screen is blank. Participants did not receive feedback.

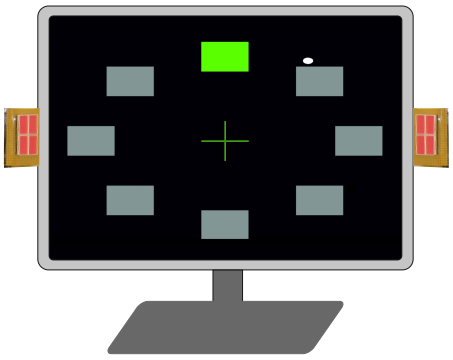


Figure 1: This image shows the monitor display that the participant viewed during the online session. In this example, the top box is the target. One red LED apparatus was affixed to each side of the monitor. The display during the offline sessions was identical, except that only the four boxes on the top, left, bottom, and right were cued.

After the session was complete, we trained two independent LDA classifiers, one for ERD and one for SSVEP. The ERD and SSVEP classifiers each had 120 trials (since there were 60 trials per class). We used 120 ERD trials to compute the CSP filter matrix from 27 electrodes (distributed over the scalp). The ERD classifier used 4 features, which were extracted from the first and last two columns of the CSP filter matrix. The SSVEP classifier used 12 logarithmic band power features (three harmonics around the stimulation frequencies) at O1 and O2. Based on results from this offline screening session, 18 subjects were invited to return to an online session, seven of whom were not available. Hence, 11 subjects participated in the online session.

The online session began with one short offline run to update the bias of the ERD classifier, with

20 trials for the two classes (up and down). Each of the following five runs lasted seven minutes and was divided into a variable number of trials. At the beginning of each run, the participant viewed a display similar to that seen in the screening session, except that there were eight potential target locations (grey boxes, see Figure 1). One of the eight boxes was cued 2 seconds after trial onset, then a white ball appeared on the center of the screen one second later.

The participant tried to move the ball to the target by performing the same mental tasks described during the offline screening. Notably, the online session usually required simultaneous control of both vertical and horizontal movement, and hence simultaneous control of ERD and SSVEP activity. If the ball struck the target, then the trial ended, and the score increased by one. Otherwise, the trial ended twenty seconds after the ball appeared. Striking a nontarget box did not end the trial. This "can't miss" protocol was developed for consistency with comparable BCI studies. There was a 3–5 second delay between trials.

We measured performance in two ways. The first is the number of trials. Better participants should be able to move the cursor to the target very quickly, and hence should complete more trials within each online seven minute run. We also assess the score, since better participants should also succeed in hitting more targets during each run.

3 Results

Based on the offline data, we estimated performance with simple two-class BCIs. The subjects exhibited mean accuracies of 79.5 % and 91.9 % (estimated with 10×10 cross-validation) with the ERD and SSVEP classifiers, respectively (see Table 1). Of the 25 participants, two exhibited SSVEP activity below 70 %, while five were below 70 % with ERD. Hence, these subjects were not contacted for further study. Of the remaining 18 subjects, seven were not available for further research, so eleven subjects participated in the online session.

Table 1 presents performance of these eleven participants with the online BCI, averaged over all five runs. Table 1 also presents offline and offline results with a simulation of random data using all of the same methods applied with the eleven human participants.

	Offline		Online			
	ERD	SSVEP	ERD	SSVEP	Score	Trials
AE9	81.3	96.3	72.3	85.1	8.0	20.4
BR6	70.6	79.5	61.9	94.6	2.8	18.8
BR7	97.8	87.3	93.1	80.2	12.8	21.8
BS3	80.8	99.5	83.9	90.9	16.8	22.4
BT6	75.7	100.0	64.9	94.3	7.6	20.2
BT7	82.8	98.3	64.8	90.0	5.2	19.4
AP3	99.1	90.3	99.9	98.5	29.2	29.8
AO1	75.9	93.5	60.5	87.1	5.6	19.4
BL5	83.3	91.4	83.0	96.1	19	24.2
BG9	77.6	98.0	87.2	97.4	22.6	25.8
BU8	83.4	97.3	85.8	80.6	10.2	21.0
Mean	79.5	91.9	77.9	90.4	12.7	22.1
Rand	58.2	64.2	N/A	N/A	5.8	19.0

Table 1: Results across eleven subjects and one random simulation.

The correlation between offline performance and online score was not significant with $r^2 = 0.35$, based on a multiple regression analysis with two independent variables (ERD and SSVEP accuracy simulated from offline data) and one dependent variable (online score). However, we used the same classification and statistical methods to estimate ERD and SSVEP classification accuracies from the online session. These values were highly correlated with the online score ($r^2 = 0.94$), as determined by a multiple regression analysis with two independent variables, ERD and SSVEP accuracy, and one dependent variable, online score ($F(2, 8) = 65.71$; $p < 0.01$). The weights (beta coefficients) of the ERD and SSVEP accuracy were 0.85 and 0.46, respectively.

4 Discussion

Our results indicate that the BCI screening procedures used here, which were consistent with established procedures for estimating performance with ERD or SSVEP BCIs, are not very effective predictors of successful performance with the hybrid BCI approach described here. One possible explanation is that the signal processing methods used in our screening procedure were inadequate. However, the same methods were effective at predicting performance based on online data. Hence, the poor prediction probably resulted from one or more differences between the offline and online procedures. The online procedure did involve multitasking, whereas the offline screening involved only one task within each trial. Also, the offline procedure did not include feedback. Finally, critical facets of subjects' ERD and SSVEP activity may have changed across days. Future research could evaluate different screening procedures and how well they predict performance with different BCIs. Improved signal processing approaches could also lead to more accurate prediction [2]. Since BCI performance varies within and across different days, future studies should also explore the reliability of predictive assessments over time.

We also found that participants who were unlikely to attain effective control with an ERD BCI would probably be successful with an SSVEP BCI, and vice versa, consistent with our earlier results [3]. In some cases, subjects are unable to use a particular BCI at all, which has been called BCI illiteracy [1, 6]. None of the 25 subjects were illiterate with both ERD and SSVEP, based on accuracy in a two-choice task that was estimated from offline results or (for eleven subjects) online results. Just as with spoken languages, "BCI illiteracy" is probably not universal.

Acknowledgments

The research leading to these results has received funding from the European Union Seventh Framework Programme FP7/2007-2013 under grant agreement 247447. We would like to thank Vera Kaiser for statistic calculations.

References

[1] D. J. McFarland, L. A. Miner, T. M. Vaughan, and J. R. Wolpaw. Mu and beta rhythm topographies during motor imagery and actual movements. *Brain Topography*, 12:177–186, 2000.

[2] B. Blankertz, C. Sannelli, S. Halder, E. M. Hammer, A. Kübler, K. R. Müller, G. Curio, and T. Dickhaus. Neurophysiological predictor of SMR-based BCI performance. *NeuroImage*, 51:1303–1309, 2010.

[3] C. Brunner, B. Z. Allison, C. Altstätter, and C. Neuper. A comparison of three brain-computer interfaces based on event-related desynchronization, steady state visual evoked potentials, or a hybrid approach using both signals. *Journal of Neural Engineering*, 8:025010, 2011.

[4] B. Allison, T. Lüth, D. Valbuena, A. Teymourian, I. Volosyak, and A. Gräser. BCI demographics: how many (and what kinds of) people can use an SSVEP BCI? *IEEE Transactions on Neural Systems Rehabilitation Engineering*, 18:107–113, 2010.

[5] A. Kübler, N. Neumann, B. Wilhelm, T. Hinterberger, and N. Birbaumer. Predictability of brain–computer communication. *International Journal of Psychophysiology*, 18:121—-129, 2004.

[6] B. Z. Allison and C. Neuper. Could anyone use a BCI? In D. S. Tan and A. Nijholt, editors, *Brain-computer interfaces: Applying our minds to human-computer interaction*. 2010.

[7] J. R. Wolpaw and D. J. McFarland. Control of a two-dimensional movement signal by a noninvasive brain-computer interface in humans. *Proceedings of the National Academy of Science USA*, 101:17849–17854, 2004.

Approaching Evaluation and Rehabilitation of Severe Acquired Brain Injuries by Means of a Neurophysiological Screening

M. Risetti[1], J. Toppi[1,2], L.R. Quitadamo[1,3], L. Astolfi[1,2,4], R. Formisano[5], D. Mattia[1]

[1]Neuroelectrical Imaging and BCI Laboratory, Fondazione Santa Lucia, Rome, Italy
[2]Department of Computer Science and Systems, University La Sapienza, Rome, Italy
[3]Department of Electronic Engineering, University of Tor Vergata, Rome, Itay
[4]Department of Physiology and Pharmacology, University La Sapienza, Rome, Italy
[5]Post-Coma Unit, Fondazione Santa Lucia, Rome, Italy

m.risetti@hsantalucia.it

Abstract

Severe Acquired Brain Injury (ABI) may evolve in Disorders of Consciousness (DOC) like Vegetative State (VS) and Minimally Conscious State (MCS) characterized by a lack-of, or minimal responsiveness. The evaluation of these patients is a big challenge, since the clinical examination is provided by a list of items that the patients are not able to perform. An auditory oddball paradigm including duration, deviants and subject own name (SON) presented as a novel, was applied in 11 DOC patients. Preliminary results point out that a neurophysiological screening, based on parameters like Event-Related Potentials and EEG back- ground activity, can be a first step to follow-up the clinical evolution of DOC patients with the intent to give a contribution for a more accurate diagnosis of this patients and to the purpose to start using the Brain Computer Interface (BCI) technology as rehabilitative intervention.

1 Introduction

Survivors from severe Acquired Brain Injury (ABI) often suffer from prolonged disorder of consciousness (DOC) such as Vegetative State (VS) and/or Minimally Conscious State (MCS), that can be acute and reversible or chronic and irreversible [1]. These conditions are characterized by a lack-of or minimal responsiveness, but in some studies islands of consciousness have been described [2]. A neurophysiological screening, based on parameters like ERPs and EEG background activity, can be a first step to follow-up the clinical evolution of ABI patients during their rehabilitation phase. The relevance of such neurophysiological monitoring monitoring is even more evident when considering the potential role of BCI technology in augmenting or even allowing communication [3]. In this study we evaluate some features of the EEG signal such as ERP components in response to an auditory paradigm and the EEG background activity recorded from patients diagnosed as VS and MCS and undergoing a rehabilitation program. We consider this neurophysiological screening as the initial step for elucidating patterns of clinically masked residual cognitive resources required to start using an EEG-based BCI as acommunication tool.

2 Methods

The sample included 11 (6 males and 5 females) severe acquired brain-injured patients (Glasgow Coma Scale ≤ 8 in the acute phase), who were diagnosed as VS ($n = 8$) and MCS ($n = 3$) by means of the JFK Coma Recovery Scale-Revised and were admitted to a rehabilitation ward during their subacute stage. An auditory oddball paradigm [4] previously validated in healthy subjects, including duration, deviants and subject own name (SON) presented as a novel, was applied to the

Group	VS	MCS
1	7.5	10.5
2	8	8
3	10	11.5
4	8.5	9.5
5	7.5	11l
6	7.5	11
7	8	7.5
M	8.14	9.85
SD	0.89	1.57

Table 1: Individual Alpha Frequency peak (IAF) of the VS and MCS patients (IAF peak mean values in red refer to as patients who underwent a second recording session).

VS (mean age: 35,6±14,7) and MCS (mean age: 45,3±16,6) patients, during their rehabilitation period. Patients were presented with 6 blocks (3 passive and 3 active runs) of 500 stimuli in a unique session, while the scalp EEG was recorded and monitored continuously throughout each recording sessions. In the passive runs the patients were asked to just listen to the auditory stimuli, whereas during the "active" runs they were instructed to count the novel stimuli. Potentials were recorded continuously from 10 EEG scalp electrodes (F3, Fz, F4, C3, Cz, C4, P3, Pz, P4, Oz) placed accordingly to the 10-20 International System and amplified by means a commercial EEG system (BrainAmp, Brainproducts GmbH, Germany). The reference were placed on the ear lobes. IThe electro-oculogram (EOG) was recorded from two pairs of electrodes, (one above and below the right eye and the other on the outer canthi of the two eyes), in order to run a semi-automatic procedure for the ocular movement artefacts removal. Electrode impedance was kept below 5 KΩ. Four patients were subjected to a second recording session following their clinical evolution (from VS to MCS), in order to find out if a change in the clinical diagnosis could correspond to a change in the neurophysiological parameters we considered.

3 Results

3.1 Power Spectra Analysis

The digital FFT-based power spectrum analysis was used to estimate the individual alpha frequency (IAF) peak. The VS group was composed by 7 patients (one VS patient was excluded from the analysis because of the contamination of EGG with massive muscular artifacts due to bruxism).The MCS group, included three patients initially diagnosed as MCS and 4 patients who underwent a change in their clinical status from VS (XX) see Table 1 to MCS (Table 1, in bold). We adopted this strategy since we were interested for the first analysis in the possible difference in the IAF related to the different clinical condition (VS and MCS) and not in their evolution. The mean IAF peak value estimated at Pz, was 8,1±0.89 Hz and 9,8±1,57 Hz for the VS and MCS groups, respectively. This IAF difference was statistically significant (two-tailed unpaired t test, $df = 12$; $P < 0.05$).

3.2 N1 to Standard Tones, MMN to Deviant Tones and nP3 to Novel

Mean values of each ERP components at Cz were calculated. N100 was present in all the patients and was significantly delayed in VS (two-tailed unpaired t test, $df = 9$; $P < 0.05$) with respect to MCS group. Mismatch negativity (MMN) was observed in all but one patients (anoxic etiology) and it was never observed without either the N100 component or the N100 deviant. Differences in latency and amplitude of the MMN component in the two patient groups did not reach statistical significance. The novelty P3 was observed in 9 out of 10 patients, 6 diagnosed as VS (3 TBI, 1 ischemic stroke, 2 hemorrhagic stroke), 3 as MCS (1 TBI, 1 ischemic stroke, 1 hemorrhagic stroke)

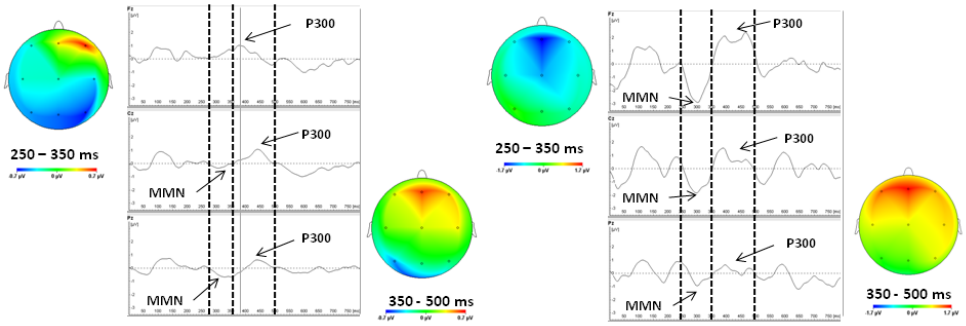


Figure 1: Grand average waveforms recorded at Fz, Cz, Pz of the VS group (on the left side) and of the MCS group (on the right side), in response to their own name. Mean scalp potentials show the distributions of the MMN and of the nP3 responses of the two main groups.

with a significantly larger amplitude in MCS (two-tailed unpaired t-test, $df = 9$; $P < 0.05$) with respect to VS group, see Figure 1. The subject with anoxic etiology showed no objectively detectable positive deflection. For all MCS and one VS patient a change in topography could be observed for the positive wave.

Statistical analysis revealed no significant differences in the mean latency and amplitude values of nP3, component detected during passive and active runs. Nevertheless, in both VS and MCS groups we noted an increment in the mean latency values of the nP3 detected during the active runs with respect to those obtained from the passive runs: maximum peak shift from 335 ms to 395 ms and from 390 ms to 460 ms in the VS group and in the MCS group, respectively An opposite trend was observed for the nP3 mean amplitudes obtained from the active with respect passive runs, with an increment in the MCS group (maximum peak from 1,5 μV to 2,3 μV) and a decrement in the VS group (maximum peak from 1,8 μV to 0,6 μV). Furthermore, we noted again a broader scalp distribution of the positive component at the parietal site in MCS patients, in both the passive and the active runs, in contrast to the frontal topographies of the VS patients, that remains stable in both conditions.

3.3 Neurophysiological Parameters Monitoring: Follow-Up

Four patients were subjected to a second recording session at the time when, the diagnosis evolved from VS into MCS, according to the CRSr score. N100 and MMN were observed in all the patients, although no significant differences were observed in latency and amplitude of the two negative components recored from the two different recording sessions. Similar results were found concerning the fronto-central nP3 occurrence during the second recording session. However, a reduction of the nP3 latency and a clear change in its topography was observed during the second EEG recording session, when the patients were diagnosed as MCS., see Figure 2.

4 Discussion

Previous data have reported that the concomitant presence of a predominant theta-alpha EEG background activity and of cognitive-related components might be relevant in predicting the recovery of consciousness in patients who survive their coma [5]. Our preliminary results replicate and extend those data pointing out that: a) the IAF peak could be useful for a more careful diagnosis of DOC patients; b) among patients with main theta and slow alpha EEG background activity, all subjects with a diagnosis of VS show some “cognitive” ERP components; c) more complex neurophysiological responses were present in individuals with reproducible, albeit inconstant, clinical evidence of awareness of them- selves and/or the environment; d) complex neurophysiological responses in VS patients can be conceived as an indication of possible diagnostic error due to

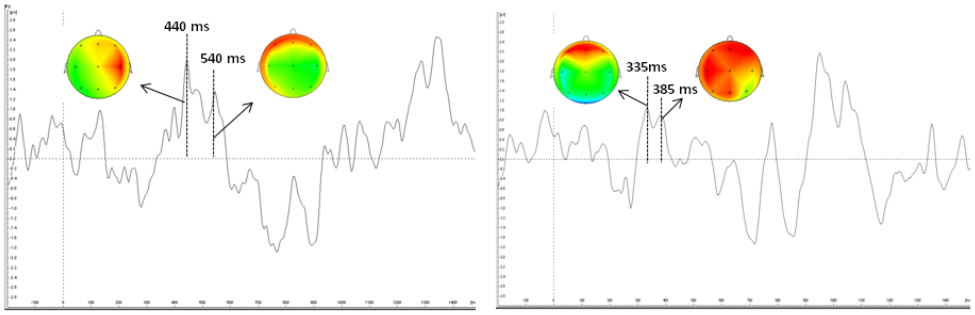


Figure 2: Grand average waveforms and mean scalp potentials at Fz of the 4 patients recorded when they were in VS (on the left side) and in MCS (on the right side), in response to their own name. A clear change in topography can be observed in the second recording session group.

the limits of the clinical evaluation, or as an evidence that isolated thalamo-cortical circuits may be preserved also in VS patients, indicating spared function of some modules, maybe unrelated to conscious experience [6]. The current study suggest also that, for the application of the BCI technology, it could be important to monitor the patients' responses longitudinally.

5 Conclusion

In the light of these preliminary results we suggest that a systematic neurophysiological screening, based on parameters like Event-Related Potentials (ERPs) and EEG background activity, could be useful as a first step to develop a battery of neurophysiological examination procedures for the assessment of the cognitive status of patients with disorder of consciousness during their rehabilitation phase.

Acknowledgments This work is supported by the European ICT Programme Project FP7-247919 (DECODER). 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] J. L. Bernat. Chronic disorders of consciousness. *Lancet*, 367:1181–1192, 2006.

[2] A. M. Owen, M. R. Coleman, M. Boly, M. H. Davis, S. Laureys, and J. D. Pickard. Detecting awareness in the vegetative state. *Science*, 313:1402, 2006.

[3] N. Birbaumer and L. G. Cohen. Brain-computer interfaces: communication and restoration of movement in paralysis. *The Journal of Physiology*, 579:621–636, 2007.

[4] C. Fischer, F. Dailler, and D. Morlet. Novelty P3 elicited by the subject's own name in comatose patients. *Clinical Neurophysiology*, 119:2224–2230, 2008.

[5] C. Babiloni, M. Sarà, F. Vecchio, F. Pistoia, F. Sebastiano, P. Onorati, G. Albertini, P. Pasqualetti, G. Cibelli, P. Buffo, and P. M. Rossini. Cortical sources fo resting-state alpha rhythms are abnormal in presistent vegetative patients. *Clinical Neurophysiology*, 120:719–729, 2009.

[6] B. Kotchoubey, S. Lang, G. Mezger, D. Schmalohr, M. Schneck, A. Semmler, V. Bostanov, and N. Birbaumer. Information processing in severe disorders of consciousness: vegetative state and minimally conscious state. *Clinical Neurophysiology*, 116:2441–2453, 2005.

Exploring Electrophysiological Correlates of Mental Imagery Paradigms Borrowed from fMRI Domain: What Can We Learn for BCI Application?

L. Astolfi[1,2], J. Toppi[1,2], F. Cincotti[1], F. Babiloni[1,2], L. Bianchi[1,3], L. Quitadamo[1,3], M. Risetti[1], D. Mattia[1]

[1]IRCCS Fondazione Santa Lucia, Neuroelectrical Imaging and BCI Lab, Rome, Italy
[2]University of Rome "Sapienza", Rome, Italy
[3]University of Rome "Tor Vergata" Rome, Italy

laura.astolfi@uniroma1.it

Abstract

Recent fMRI studies showed that appropriate mental imagery paradigms were effective in unveiling a significant volitional activity in a number of patients who had been diagnosed Disorders of Consciousness (DOC). The aim of this study was to investigate if such paradigms could be suitable for a translation to the Electroencephalography (EEG) domain. To this purpose, we computed a physiological study on 27 healthy volunteers, by using advanced high resolution EEG methodology. The effect of a previous motor imagery based BCI training was also investigated by comparing the statistical cortical maps obtained for naïve and trained subjects. Our findings indicate that tennis and navigation imagery can be characterized by different frequency content even in subjects who were naïve to previous mental imagery training, while simple movement imagery, like the hand grasping, needs training to elicit significant activations.

1 Introduction

In recent studies [1], the analysis of functional Magnetic Resonance Imaging (fMRI) activation patterns related to a mental activity elicited in response to the instruction to perform a mental imagery task was used to identify residual cognitive function and even conscious awareness in some patients who were assumed to be in Vegetative State (VS). In addition, this strategy was successfully used to establish non-muscular based binary communication with a patient in a diagnosis of VS [2]. Due to its properties in terms of usability, portability and easiness-to-use, EEG provides a flexible and non-expensive instrument to help the assessment of residual conscious awareness in nonresponsive patients, including DOC patients, usable in clinical routine at the patients' bedside. The aim of this study was to investigate whether the mental imagery paradigms which were already shown to be effective in fMRI in some patients who had been diagnosed a VS could be suitable for a translation to the Electroencephalography (EEG) domain. Spectral statistical brain activation maps were obtained from non invasive EEG data recorded from a group of healthy volunteers performing the imagery tasks used with VS patients (imagery of playing tennis and of navigating through a familiar environment). To provide a comparison with a task which has been successfully applied in SMRs-based BCI paradigms, a third imagery task (prolonged hand grasping) was also tested and its activations were compared to those obtained by the two other mental tasks. The translation of such mental tasks from the hemodynamic to the electrical domain constitutes a challenge per se, because it is well known that the spatial resolution of the detected brain signals degrades from the fMRI to the EEG technique, and what is clear and evident in an fMRI deep image becomes blurry and sparse on a scalp EEG. The need of training for subjects

in order to acquire control of a BCI system operated by motor imagery becomes relevant for the deployment of BCI system in DOC patients. On the other hand the two mental tasks of tennis and spatial navigation were able to induce reproducible and robust patterns of blood oxygenation level dependent (BOLD) response without any training [1, 3]. To investigate the effects of BCI training to the activation maps, we compared results obtained from naïve subjects with those obtained from volunteers who had already experienced a motor imagery operated-BCI training.

2 Methods

2.1 Experimental Design

Twenty-seven healthy right-handed volunteers (age: 28.5±6.4 years, 14 females; 13 males) took part in the experiment. Sixteen subjects had already performed a BCI training based on motor imagery of simple hand/foot movements (same number of prior training sessions for all the subjects) while the remaining 11 were naïve to BCI protocols. Subjects were comfortably seated in front of a computer screen and were asked to perform one of the following tasks: to imagine they were in the middle of a tennis court playing tennis (T), moving through the rooms of a familiar environment like their own house (spatial navigation, N), or performing a prolonged grasping of both hands (G) for all the task length. As an idle condition, they were asked just to relax (R). Regarding tennis and hand grasping imagery, subjects were encouraged to perform a kinaesthetic imagery rather than visually "seeing" themselves playing tennis or moving hands.The experiment was divided into 6 runs 24 trials each (randomly ordered). The task length was set to 15 seconds and the inter-trial interval to 2 seconds. EEG potentials were recorded by a 61-channel system by means of an electrode cap. Sampling rate was 200 Hz. Surface EMG was recorded to ensure that participants refrained from contracting their arm muscles during imagery.

2.2 Linear Inverse Procedure and Statistical Spectral Mapping

After preprocessing - band pass filtering (1–45 Hz), eye movements removal by means of Independent Component Analysis (ICA), artifact rejection - a distributed source model of the cortex (based on Colin template, T1-weighted MRI volume, Montreal Neurological Institute), made of about 8000 equivalent current dipoles, was used for the study. High-resolution EEG algorithm [4] was used to estimate the current density strength for each one of the cortical dipoles, by solving the electromagnetic linear inverse problem. From the cortical waveforms, we estimated the Power Spectrum Density (PSD), for each one of the 8 thousands dipoles of the cortical model used. T-test values obtained from comparisons between tennis-rest, grasping-rest and navigation-rest were then mapped, for each subject, on the cortical model in different frequency bands, defined according to Individual Alpha Frequency (IAF). False Discovery Rate [5] correction for multiple comparisons was applied to the statistical tests. In order to extract common features shared by the experimental population, we obtained Grand Average (GA) statistical spectral maps for the three comparisons: tennis vs rest, grasping vs rest and navigation vs rest, for each of the four frequency bands.

3 Results

The Grand Average cortical maps obtained for the statistical contrast between the three imagery tasks and the rest condition indicate that spatial navigation imagery elicits relevant patches of significant activity in the Alpha power spectra, mainly involving the posterior bilateral parietal areas. Conversely, the playing tennis and grasping imagery produced significant power spectra changes within the Beta frequency band, mainly localized over the left posterior parietal areas and left sensorimotor areas.

To investigate the influence of training on the three mental tasks, we performed a second group analysis, by dividing the subjects into two groups: those who had no previous experience in BCI

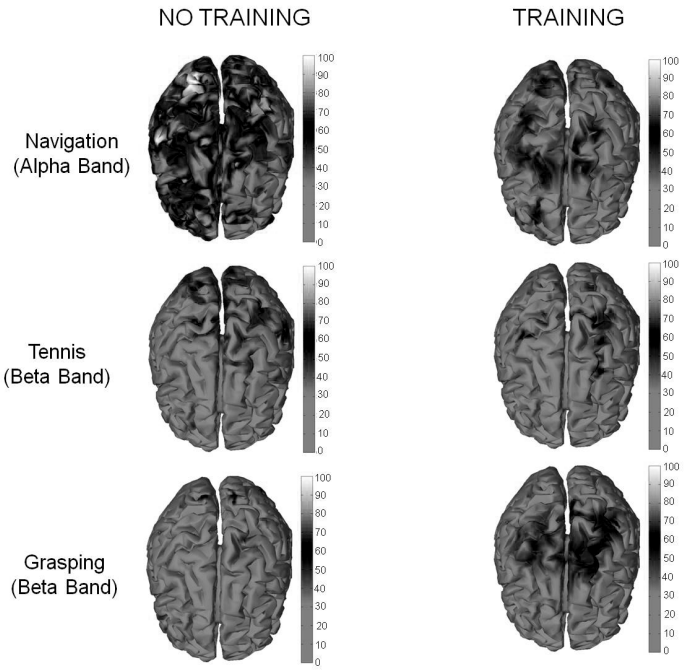


Figure 1: Grand average cortical maps obtained grouping naïve subjects (first column) and trained subjects (second column) for the three comparisons navigation vs rest in Alpha frequency band (first row), tennis vs rest (second row) and grasping vs rest (third row), both in Beta frequency band.

from those who experienced a SMR-based BCI training (about 8 SMR-based BCI sessions, where subjects had to learn to control a up-down cursor motion by imaging hands and feet movements; 45 minute each). For each of the two groups we computed the GA cortical maps, by contrasting tasks with the same procedure described for the whole subject sample. The color of each dipole codes for the percentage of subjects for which the spectral activity of such dipole, elicited during the task, results significantly different from the activity arisen during the rest condition. As expected, we found a remarkable difference between the two groups (Figure 1) in the magnitude of the grasping imagery task-related spectral pattern within the Beta frequency band (third row). Non-trained subjects displayed almost no common relevant spectral changes between the rest condition and the mental imagery, while trained subjects showed robust and reproducible patches of significant activity over the sensorimotor areas. As for the tennis imagery task (second row), the two groups showed no significant difference in the magnitude of spectral power changes in Beta band, that remained distributed mainly over the left posterior parietal cortex. Navigation imagery task (first row) was characterized by consistent Alpha power spectra change, mainly involving the bilateral parieto-occipital areas and pre-frontal areas. The same task performed after an exposure to BCI training was again associated with spectral power change in the Alpha frequency that was confined over the right posterior parietal cortex and the pre-motor areas.

4 Discussion

The temporal resolution of EEG technique allows to point out and differentiate the specific content of the two types of task (motor imagery and spatial navigation imagery) in different frequency bands. Results showed that the motor imagery tasks elicited a significant pattern of activation in the Beta band, while the spectral content of navigation imagery was mainly apparent in the Alpha band. The analysis of the activation patterns obtained in two groups of subjects (naïve

and trained), according to their previous experience in BCI applications, revealed an effect of BCI training on the strength and focus of the cortical activations resulting from the prolonged hand grasping. On the other hand, tennis imagery showed similar pattern of activity within a motor-relevant oscillation frequency (i.e. Beta band) between trained and non-trained group,and navigation imagery is associated with higher magnitude and widespread power spectral changes in non-trained subjects with respect to the trained. The effects on the tennis and spatial navigation imagery could be mainly due to a lower attentive load required by these tasks after an SMR BCI training.Bearing in mind that these differences should be confirmed in a sample of subjects trained for these 2 specific mental tasks (i.e. tennis and navigation imagery), we can speculate that the changes in the oscillatory alpha activity (8–13 Hz) involving the parieto-occipital areas reflect an idling and/or functional inhibition state rather than play a direct role in the neuronal processing required for a given task.

5 Conclusion

To our knowledge, this is the first EEG study where an advanced body of methodology has been applied to disentagle temporal and spatial features of mental tasks that have been already proved under haemodynamical domain to be valuable in detecting cognition in DOC patients. Our findings indicate that tennis and navigation imagery can be characterized by different frequency content even in subjects who were naïve to previous mental imagery training based on BCI technology, while they show the necessity of training to obtain activations from simple movement imagery like the hand grasping. This latter finding is of relevance for a BCI application in detecting residual cognition in DOC patients.

6 Acknowledgments

This work was supported by the European ICT Program FP7-ICT-2009-4 Grant Agreement 247919 DECODER.

References

[1] A. M. Owen, M. R. Coleman, M. Boly, M. H. Davis, S. Laureys, and J. D. Pickard. Detecting awareness in the vegetative state. *Science*, 313(5792):1402, 2006.

[2] M. M. Monti, A. Vanhaudenhuyse, M. R. Coleman, M. Boly, J. D. Pickard, L. Tshibanda, A. M. Owen, and S. Laureys. Willful modulation of brain activity in disorders of consciousness. *New England Journal of Medicine*, 362(7):579–589, February 2010.

[3] M. Boly, E. Balteau, C. Schnakers, C. Degueldre, G. Moonen, A. Luxen, C. Phillips, P. Peigneux, P. Maquet, and S. Laureys. Baseline brain activity fluctuations predict somatosensory perception in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 104(29):12187–12192, July 2007. PMID: 17616583 PMCID: 1924544.

[4] F. Babiloni, F. Cincotti, C. Babiloni, F. Carducci, D. Mattia, L. Astolfi, A. Basilisco, P. M. Rossini, L. Ding, Y. Ni, J. Cheng, K. Christine, J. Sweeney, and B. He. Estimation of the cortical functional connectivity with the multimodal integration of high-resolution EEG and fMRI data by directed transfer function. *NeuroImage*, 24(1):118–131, January 2005.

[5] Y. Benjamini. The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics*, 29(4):1165–1188, August 2001. Mathematical Reviews number (MathSciNet): MR1869245; Zentralblatt MATH identifier: 01829051.

ERPs Contributing to Classification in the "P300" BCI

T. Kaufmann[1], E. M. Hammer[1], A. Kübler[1]

[1] Department of Psychology I, University of Würzburg, Würzburg, Germany

tobias.kaufmann@uni-wuerzburg.de

Abstract

Brain Computer Interfaces (BCI) provide a non-muscular communication channel for people with severe motor impairment. The most commonly used BCI for communication is the so called P300-Speller that received its name from the event-related potential (ERP) P300 which is elicited by the speller (oddball) paradigm. Several researchers reported that it is not only the P300 but also other ERPs that are classified with this type of BCI. This study thus, aims at contributing to the discussion by assessing the ERPs which contribute most to classification in a large sample size of $N = 51$ participants. Our results indicate that almost 30 % of all participants reached highest determination coefficients with the N200, not with the P300. Furthermore it is often a combination of both potentials that is picked up for classification. Therefore the "P300"-BCI is truely a BCI that is controlled by various ERPs, especially the N200 and P300.

1 Introduction

In 1988 Farwell and Donchin introduced a brain-computer interface (BCI) that allowed for non-muscular communication [1]. Users were presented with a matrix of characters that were flashed row/column wise in random order and were asked to focus on the character they intend to spell. The stimulation followed an oddball paradigm as flashing of the target character (odd stimulus) was less likely than flashing of a non-target character. Presentation of such odd stimuli was shown to elicit event related potentials (ERP) in the electroencephalogram (EEG), for example the P300, a prominent, positive deflection approximately 300 ms post stimulus [2–4, for a review]. For that reason, this type of BCI is often calle P300-BCI [5, for a review]. However, the P300 is not the only ERP contributing to classification and successive letter selection and many studies reported on other ERPs in the P300-BCI, e.g. [6–11]. Earlier ERPs were shown to be modulated by attention and paradigm design, thus influencing classification in the P300-BCI [7, 9]. Their amplitudes (absolute values) may even be in the same range as the prominent P300 [9]. Classification of the N200 may not only be performed in the classic P300-BCI. Hong et al. [12] suggested a paradigm to specifically enhance N200 component (N200-speller).
Altogether these studies indicate influence of other ERPs on classification accuracy in the P300-BCI, however, to our knowledge this has never been quantified. Therefore, this study aims at validating the ERPs contributing to classification in a large sample size.

2 Methods

2.1 Participants

$N = 55$ university students with no known neurological disorder and normal or corrected-to-normal vision participated in the study. Data from 4 participants had to be discarded due to noisy EEG recordings. The sample thus, comprised $N = 51$ participants (27 men; mean age 24.3 years, SD = 3.49). All participants signed informed consent and assured they had never previously participated in any BCI study.

2.2 Equipment and Data Acquisition

EEG was recorded from 12 passive Ag/AgCl electrodes with a mastoid ground and a mastoid reference at locations [Fz, FC1, FC2, C3, Cz, C4, P3, Pz, P4, O1, ,Oz, O2]. The signal was amplified with a g.USBamp (g.tec Medical Engineering GmbH, Austria) and recorded at a sampling frequency of 256 Hz using the BCI2000 software [13].

2.3 Procedure

A 6×6 matrix of characters [1] was presented on a computer screen, at a distance approximately 1 m ahead of the participants. The characters were flashed row/column-wise in pseudo-random order and participants were instructed to spell a default word of 10 characters by focusing on the character they intended to spell and counting the number of flashes. Each row and column was flashed 15 times per character, with flash duration of 31.25 ms and an inter stimulus interval of 125 ms.

2.4 Data Analysis

EEG raw data was high pass filtered at 0.1 Hz cut-off frequency and low pass filtered at 30 Hz. Segmentation, averaging and peak detection was performed in Matlab (The Mathworks GmbH, US). For offline analysis of ERP amplitudes all target stimuli with 2 or less non-target stimuli in between were discarded from the data set to avoid overlapping ERPs. Occurrence of such consecutive flashings is unlikely due to the properties of the paradigm; however for observation of ERPs, this correction is necessary to assure a minimum distance of 466 ms between two target stimuli. Hereinafter, determination coefficients were computed for every electrode in a time frame of 600 ms post stimulus. For every participant peak amplitude, latency, polarity and location of the time bin, that yielded maximal R^2 value, were determined.

3 Results

Each participant was assigned to one group according to the polarity of the potential that contributed most to determination of target and non-target trials. These potentials were either the P300 or the N200. Of all participants 70.6 % ($N = 36$) had maximum determination at 335.5 ms at mean, which was at the peak amplitude of the P300. The remaining 29.4 % ($N = 15$) achieved best determination at occipital electrode sites with the N200 at 195 ms on average. The resulting R^2 values show similar range for both groups (see Table 1).

These results were also reflected in the grand average of all participants. On average it was the P300 that could be determined best (e.g., blue arrow at electrode Cz in Figure 1 A). However on occipital electrode sites there was also an increased determination for the N200 (e.g., green arrow at electrode O1 in Figure 1 A).

Single participant data revealed that for some participants the N200 had higher amplitudes than the P300. This led to high determination coefficients around 200 ms but comparatively small values for the P300. Example is provided for participant 20 in Figure 1 B. Other participants, e.g. participant 25 in Figure 1 B, had a pronounced P300 but also large N200 on occipital electrode sites.

4 Discussion

Our results are in line with earlier findings that control of the P300-BCI is not limited to classification of the P300 [7, 9]. With a large sample size of 51 participants we could show that almost 30 % of all participants had higher determination coefficients for the N200 as compared to the P300. Furthermore, in many participants, both ERPs, the P300 and the N200 were taken into account when training a classifier.

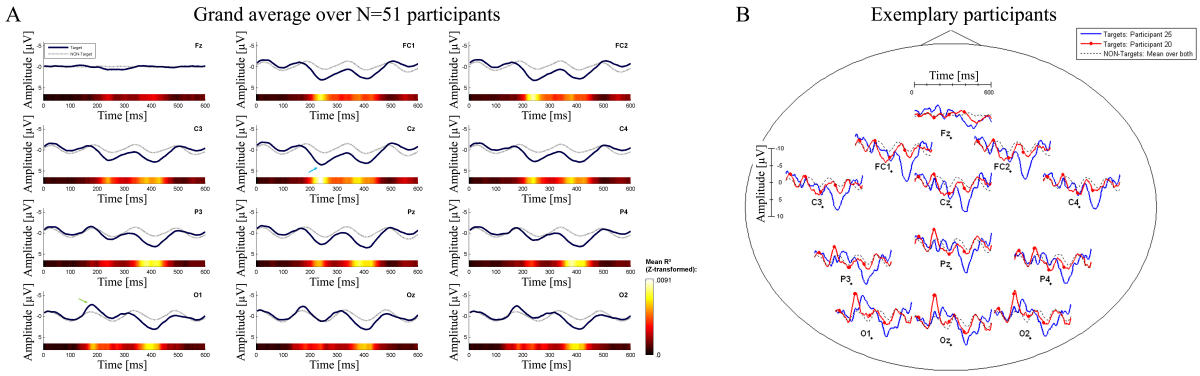


Figure 1: (A) Grand average of $N = 51$ participants. A color bar indicates quality of determination for every time bin at the presented electrode location. P300 could be determined best at Cz (blue arrow), N200 could be determined best at O1 (green arrow). (B) Data from two exemplary participants. Participant 25 is a typical example for large P300 amplitudes at centro-parietal electrode sites, resulting in highest determination coefficiant at electrode Cz. Participant 20 shows large negative peaks and highest R^2 value around 200 ms (N200) on occipital electrode sites.

	Best determination for positive potential (P300)	Best determination for negativ potential (N200)
Sample size[N]	36	15
Sample size[%]	70.6	29.4
R^2 range	.014-.078	.014-.071
Mean amplitude [μV]	5.17	-5.49
Mean latency [ms]	335.5	195.05
Main occurence	70 % centro-parietal	100 % occipital

Table 1: Comparison of parameters according to polarity and time of the most prominent potential.

5 Conclusion

Almost one third of $N = 51$ participants controlling a "P300"-BCI achieved higher discrimination with an N200 at occipital sites than with the P300 at centro-parietal sites. Often a combination of both ERPs is actually classified. Therefore the "P300"-BCI is truely a "P300-N200"-BCI or simply: an ERP-BCI.

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(6):510–523, December 1988.

[2] S. Sutton, M. Braren, J. Zubin, and E. R. John. Evoked-potential correlates of stimulus uncertainty. *Science (New York, N.Y.)*, 150(700):1187–1188, November 1965. PMID: 5852977.

[3] T. W. Picton. The p300 wave of the human event-related potential. *Journal of Clinical Neurophysiology: Official Publication of the American Electroencephalographic Society*, 9(4):456–479, October 1992. PMID: 1464675.

[4] J. Polich. Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 118(10):2128–2148, October 2007. PMID: 17573239.

[5] S. C. Kleih, T. Kaufmann, C. Zickler, S Halder, F. Leotta, F. Cincotti, R. Aloise, C. Herbert, D. Mattia, and A. Kübler. Out of the frying pan into the fire – the P300 based BCI faces real world challenges. *Progress in Brain Research*.

[6] B. Z. Allison and J. A. Pineda. ERPs evoked by different matrix sizes: implications for a brain computer interface (BCI) system. *IEEE Transactions on Neural Systems and Rehabilitation Engineering: A Publication of the IEEE Engineering in Medicine and Biology Society*, 11(2):110–113, June 2003. PMID: 12899248.

[7] B. Z. Allison and J. A. Pineda. Effects of SOA and flash pattern manipulations on ERPs, performance, and preference: implications for a BCI system. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, 59(2):127–140, February 2006. PMID: 16054256.

[8] L. Bianchi, S. Sami, A. Hillebrand, I. P. Fawcett, L. R. Quitadamo, and S. Seri. Which physiological components are more suitable for visual ERP based brain-computer interface? A preliminary MEG/EEG study. *Brain Topography*, 23(2):180–185, June 2010. PMID: 20405196.

[9] M. S. Treder and B. Blankertz. (C)overt attention and visual speller design in an ERP-based brain-computer interface. *Behavioral and Brain Functions*, 6(1):28, 2010.

[10] P. Brunner, A. L. Ritaccio, J. F. Emrich, H. Bischof, and G. Schalk. Rapid communication with a "P300" matrix speller using electrocorticographic signals (ECoG). *Frontiers in Neuroscience*, 5:5, 2011. PMID: 21369351.

[11] S. Frenzel and E. Neubert. Is the P300 Speller Independent? *1006.3688*, June 2010.

[12] B. Hong, F. Guo, T. Liu, X. Gao, and S. Gao. N200-speller using motion-onset visual response. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 120(9):1658–1666, September 2009. PMID: 19640783.

[13] 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 Bio-Medical Engineering*, 51(6):1034–1043, June 2004. PMID: 15188875.

Fundamental Research for Development of P300-Type Brain-Computer Interface Using Air Conduction Sound Stimulation that Localises the Sound Image

M. Chishima[1], A. Nara[2], M. Otani[3], M. Kayama[3], M. Hashimoto[3], K. Itoh[3], Y. Arai[4]

[1]School of Health Sciences, Shinshu University, Nagano, Japan
[2]Central Rehabilitation Service, University of Tokyo Hospital, Tokyo, Japan
[3]Faculty of Engineering, Shinshu University, Nagano, Japan
[4]Department of Electronic Information Engineering, Nagano National College of Technology, Nagano, Japan

mchishi@shinshu-u.ac.jp

Abstract

This preliminary experiment confirmed that it is possible to derive P300 components effective for BCI application using a discrimination task with sound image localization in three directions on the horizontal plane, as auditory stimulations from both the ears through a set of headphones. Discrimination was possible in up to 6 types of auditory stimulation with varying frequencies and sound image directions, and derivation of P300 components effective for BCI application was possible in healthy individuals. Selection with good accuracy was obtained from 5 healthy individuals in the operation experiment using a simplified BCI system and we confirmed the possibility of developing a P300-type BCI using sound image localization for auditory stimulation.

1 Introduction

Many researchers have preformed a development of brain computer interface (BCI) system using visual stimulation [1, 2]. Recently a few studies have been published on an auditory BCI [3–5] in which auditory stimuli are presented via loudspeakers. On the other hand, we examined an auditory stimulant presentaion system through a set of headphones that presents a virtual sound image on the horizontal plane to a listener. The virtual sound image can be presented by utilizing a interaural time differences (ITD) and interaural level differences (ILD) that provides essential cues for sound localization [6]. Through preliminary experiments, we verified whether it is possible to derive P300 components from sound image discrimination with varying parameters of direction information and frequency information. In a group of healthy subjects, we also examined online operational efficiency of an experimental and prototype P300-type BCI system developed using sound image discrimination.

2 Subjects and Methods

Eight healthy subjects provided their informed consent to participation in these experiments, which were performed in accordance with Shinshu University Code of Ethics-school year 2010.

2.1 Construction of an Auditory Stimulation System with Sound Image Localization and Preliminary Experiments

We attempted to derive targeted P300 components using sound image stimulation to have our subjects perceive left and right orientation in the horizontal direction from the median 0° as a reference, using air conduction sound stimulation from both ears through a set of headphones. Subjects listened to sound image stimulation using a set of headphones (Sony MDR-Z700). As shown in Figure 1, a stimulant presentation system was constructed, and auditory stimulation with ITD (interaural time difference) 0.6 ms and ILD (interaural level difference) 10 dB that localised

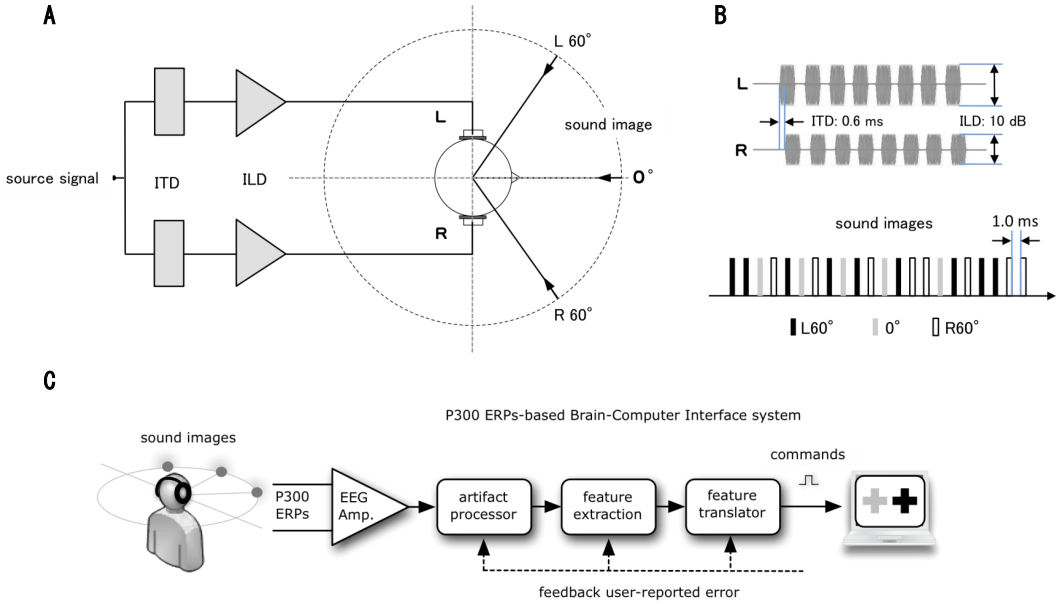


Figure 1: Experimental set-up. (A) Block diagram of the auditory component of the system. 6 types of sound image stimulation with 3 types of image directions (L60°, 0°, R60°) in combination with 2 frequencies (1 kHz, 2 kHz). (B) Location for the sound image stimulation from right (R) and left (L) side. Timeline of events in the protocol. The direction of 60° to the left (L60°) and 60° to the right (R60°) from the median 0° was used. (C) Schematic illustration of the operation experimentaion online.

the sound image in the direction of 60° to the left and 60° to the right from the median 0° was used. Four types of preliminary experiments were performed to derive P300 components through sound image localization stimulation in healthy subjects: The frequency was fixed to 2 kHz and 1 kHz, and 6 types of stimulation localised at median 0° (0°), left 60° (L60°), and right 60° (R60°). The subjects counted the number of the targeted stimuli with various combination of azimuth and frequency as directed.

2.2 Presentation of Sound Image Stimulation and Identification of P300 Components

Auditory stimulation as a localised sound image had rise and fall times of 10 ms, and was articulated using tone burst with a duration of 100 ms, interstimulus intervals of 2.0 s and 1.0 s, and sound pressure level (SPL) of 50 dB. The P300 component was recorded by electroencephalography (MEB-5508; Nihon Kohden Corp.) using 3 small dish electrodes on the median frontal region (Fz), median central region (Cz), and median parietal region (Pz), by following the international 10-20 system. Reference electrodes $A_1 + A_2$ were connected to both earlobes, and the median frontal lobe (Fpz) was grounded. Measurements were taken inside a shielded room with the subject resting in the sitting position. Sound image stimulation was output from a PC using an audio interface (UA-5; Edirol), and the subjects listened with both ears using a set of headphones. The derived EEG signals were stored using the electroencephalograph through Ag-AgCl dish electrodes (NE-121B; Nihon Kohden Corp.) connected to the scalp. Filters were set as band pass filters at 0.5 Hz (low bound) and 50 Hz (high bound). Data were recorded using a data recording board (DAQPad-6016; NI) and a PC at a sampling frequency of 1 kHz with 16-bit quantisation. Bias correction was performed using the average during 100 ms before presentation of the stimulus for each trial as 0 μV, and trials with peak-to-peak amplitude exceeding $\pm 50\,\mu$V were excluded from analysis. Peak-to-peak amplitude between the latent period 250–400 ms was considered as the

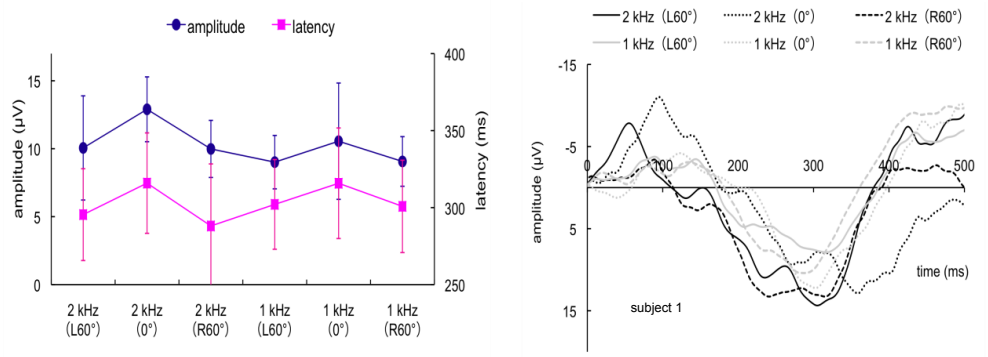


Figure 2: Left: Result of operation experiment. Averaged among 8 healthy subjects, Right: Example for a steady P300 component.

P300 component, and the results from 20 trials for target stimuli were signal-averaged to analyse the target components.

2.3 Operation Experiment Using Simplified P300-Type BCI System with Sound Image Localization

An operation experiment was performed online to examine the operational efficiency of the new BCI system using sound image stimulation with added direction information. 6 types of virtual sound source with varying sound image direction and frequency were randomly presented at an interstimulus interval of 1.0 ms. Figure 1 shows an overview of the system configuration. This simplified BCI system was used to perform an operation experiment online for 5 healthy subjects. Derived P300 components were classified based on selection using a computer, and synchronized averaging was only performed when a peak value for positive electric potential exceeding 10 μV was recorded during the latent period of 250–400 ms. Target components were extracted by discriminant analysis between P300 component information for each subject obtained online for a waveform source that was obtained by adding the waveform components more than 10 times. The waveform components were obtained through target stimulation selection.

3 Results

3.1 Sound Image Discrimination for 6 Types of Stimulation

We performed a P300 component derivation experiment using a total of 6 types of sound image stimulation with 3 types of image directions (L60°, 0°, R60°) in combination with 2 frequencies (1 kHz, 2 kHz). One sound image stimulation (target) was arbitrarily selected from among the 6 types of stimulation which were presented randomly but at the same ratio. A left panel of Figure 2 shows the average and standard deviation of peak-to-peak amplitude and latent period of each sound image stimulation for the 8 subjects. A right panel of Figure 2 shows an example of derived P300 event-related potentials waveforms from one subject. Similar to when 4 types of sound stimulation were used, P300 component derivation was confirmed with all target stimuli selected from sound image stimulations, and there were no clear differences in peak-to-peak amplitude or latent period. As derivation of P300 component is possible from the above 6 types of sound image stimulation, it is confirmed that development of BCI using new auditory stimulation with multiple commands is possible using discrimination tasks with varying sound image directions and frequencies.

3.2 Operation Experimentation Online

An operation experimentation was performed online in 5 healthy subjects. Useful waveform components were derived in all subjects for peak-to-peak amplitude and latent period obtained from each sound image direction presented as target stimulation. For the experiment with 6 types of sound image discrimination, the average selectivity was 88.0 % and the average number of times target stimulation was presented was 10.6 times. The average time required for identifying stimulus was identified was 73.2 s.

4 Discussion and Conclusion

The results of this study suggested the possibility of developing a P300-type BCI using sound image discrimination with auditory stimulation with ITD and ILD, localised in the left and right directions. The auditory stimuli used in the preliminary experiment were tone bursts over a frequency range that is very easy to discriminate with a human being's sound image localization ability. The duration of the stimulus sound was 120 ms, and was presented in random order with constant intervals between presentations of 2.0 s or 1.0 s. This task can be used as an index for stimulus duration that can be discriminated in a short duration of stimuli with good accuracy, as well as the limit for reducing the time to achieve BCI operational efficiency using this mechanism. Reviewing the duration of stimulation sounds that can be discriminated as well as intervals between presentations is necessary. With regard to methods for presentation of auditory stimulation with sound image localization, there is active research related to virtual sound in 3D space, similar to studies on ocular information. With regard to articulation of sound images that is easy to discriminate, this study had the limitation that we did not consider presentation including sound image stimulation in the vertical direction on median plane. More detailed comparative examinations including HRTF filters to take into consideration differences between individuals, as well as number and combination of discrimination task sounds will be required. Selection with good accuracy was obtained from the online operation experiment performed in healthy individuals, but we could not examine device control for specific support of everyday living. It will be necessary to re-examine the feature abstraction algorithm to achieve faster and more efficient multiple commands, and to further examine the possibility of developing BCI using auditory stimulation to proceed to clinical studies.

Acknowledgments

This study was supported in part by a Grant-in-Aid for Scientific Research (C) 19500474 and 22500502 from the Japan Society for the Promotion of Science (JSPS).

References

[1] J. J. Vidal. Toward direct brain-computer communication. *Annual Rev. Biophysics and Bioengineering*, 2:157–180, 1973.

[2] G. Dornhege, J. del R. Millán, T. Hinterberger, D. J. McFarland, and K. R. Müller, editors. *Toward brain-computer interfacing*. MIT Press, Cambridge, MA, 2007.

[3] S. Halder, M. Rea, R. Andreoni, E. M. Hammer, S. C. Kleih, N. Birbaumer, and A. Kübler. An auditory oddball brain-computer interfdace for binary choices. *Clin. Neurophysiol.*, 121(4):516–523, 2010.

[4] M. Schreuder, B. Blankertz, and M. Tangermann. A new auditory multi-class brain-computer inteface paradigm: Spatial hearing as an informative cue. *PLoS ONE*, 5(4):e9813, 1–14, 2010.

[5] F. Nijboer, A. Furdea, I. Gunst, J. Mellinger, D. J. McFarland, N. Birbaumer, and A. Kübler. An auditory brain-computer interface (BCI). *J. Neurosci. Methods*, 167(1):43–50, 2008.

[6] J. Blauert. *Spatial hearing: the psychophysics of human sound localization (revised edition)*, chapter Spatial hearing with multiple sound sources and in enclosed space, pages 201–287. MIT Press, Cambridge, MA, 1997.

Prediction of Visual P300 BCI Aptitude Using Spectral Features

S. Halder[1,2], M. Spüler[2], E. Hammer[3], S. Kleih[3], M. Bogdan[2,4], W. Rosenstiel[2], A. Kübler[3], N. Birbaumer[1,5]

[1]Institute of Medical Psychology and Behavioral Neurobiology, University of Tübingen, Gartenstr. 29, 72074 Tübingen, Germany
[2] Wilhelm-Schickard Institute for Computer Science, University of Tübingen, Sand 13, 72076 Tübingen, Germany
[3] Department of Psychology I, University of Würzburg, Marcusstr. 9-11, 97070 Würzburg, Germany
[4] Computer Engineering, University of Leipzig, Johannisgasse 26, 04103 Leipzig, Germany
[5] Ospedale San Camillo, Istituto di Ricovero e Cura a Carattere Scientifico, Via Alberoni 70, 30126 Venezia, Italy

sebastian.halder@uni-tuebingen.de

Abstract

Brain-computer interfaces (BCIs) enable patients with amyotrophic lateral sclerosis (ALS) or other motor impairments to control devices for communication and assistive purposes. Previous studies have shown that the aptitude to control a BCI without prior training varies from person-to-person. Reliable prediction of this aptitude would reduce the amount of time needed for the selection of the optimal paradigm for a particular user. In this study we show that it is possible to predict P300 BCI aptitude from a resting state electroencephalography (EEG) measurement recorded on a different day than the P300 BCI measurement.
Forty healthy participants took part in a single visual P300 EEG BCI measurement. This measurement was used to assess the aptitude of the participants. On a different day than the P300 BCI measurement a two-minute resting EEG was recorded from which spectral features (band power and frequency) were extracted. The predictive power of P300 BCI aptitude of these features was evaluated using correlation.
A negative correlation between the frequency of the spectral peak with the maximal amplitude in the resting EEG and P300 BCI aptitude of $r = -0.72$ ($p < 0.000001$) was found on posterior electrodes. Positive correlations between frequency and P300 BCI aptitude were found on frontal electrodes ($r = 0.36, p < 0.05$). Significant correlations between amplitude and P300 BCI aptitude could not be found.
Predictors of visual P300 BCI aptitude extracted from resting EEG are of particular interesting with respect to the application of BCIs to non-responsive patients. The effort to set up the experiment is minimal and the conveyance of instructions is not required.

1 Introduction

Brain-computer interfaces (BCIs) can be controlled using signals recorded directly from the user's brain. This enables completely paralyzed people, affected by e.g. amyotrophic lateral sclerosis (ALS), the use of communication devices or other applications such as internet browsers [1]. Most BCIs are based on electroencephalography (EEG) due to the availability of economic and portable recording devices. The most common signals extracted from the EEG for BCI control include slow cortical potentials (SCPs) [2], sensorimotor rhythms (SMR) [3] and the P300 event-related potential (ERP) [4]. In this study we evaluated a method for the prediction of P300 BCI aptitude because of the wide-spread use of this control signal for communication with ALS patients [5].

The P300 is commonly elicited by flashing rows and columns of a 6×6 letter matrix. The user selects a symbol by focusing his or her attention on the row and the column containing the symbol intended for selection. Recently SMR BCI aptitude has been shown to be predictable based on spectral features extracted from resting state EEG [6]. A relationship between resting state EEG and P300 ERP has also been shown to exist [7]. Based on these observations we decided to analyze the predictive power of spectral features extracted from the resting state EEG using the method originally developed for prediction of SMR BCI aptitude from [6] on P300 BCI aptitude.

2 Methods

2.1 Participants

Forty healthy participants (21 male, 19 female, mean age 25.8 years, SD 8.46 years, range 17–58) took part in the study which was approved by the Ethical Review Board of the Medical Faculty, University of Tübingen. Each participant was informed about the purpose of the study and signed informed consent prior to the experimental session.

2.2 Experimental Design

Participants were seated in a comfortable chair approximately 1 m away from a 43 cm diameter digital computer screen on which the flashing rows and columns of the visual P300 BCI were presented.
The resting state EEG was always recorded on a different day than the P300 BCI session. It consisted of ten periods with a duration of 15 s with alternating instructions "relax with eyes open" and "relax with eyes closed". The periods with eyes open were used for the analysis. This data was recorded using 119 EEG Ag/AgCl electrodes of a 128-channel cap and sampled at 1000 Hz with a bandpass filter between 0.5 and 200 Hz.
The visual P300 BCI experiment was conducted using a 5×5 matrix containing all letters of the latin alphabet excluding the letter Z. Each row and each column flashed 15 times per letter selection. One flash had a duration of 80 ms, followed by 160 ms inter-flash interval. Between letter selections pauses of 2.4 s were used to determine and present the classification results. In total, a single letter selection required 38.4 s. The participants spelled the words BRAIN and POWER three times. The first two words were used to calibrate the classifier. Therefore, feedback was only provided for words three to six. The accuracy measured in correctly selected letters of these four words was used to define the aptitude of the participants. The EEG was recorded using 67 Ag/AgCl electrodes in a 128-channel cap with 500 Hz and a notch filter at 50 Hz. Online classification was performed using stepwise linear discriminant analysis (SWLDA). To prevent a ceiling effect in the performance data (most participants reached 100 % accuracy) the visual P300 BCI data was reclassified offline with three flashes per selection using the same classification weights as online.

2.3 Performance Prediction

The power spectral density (PSD) of the resting state EEG data from 2–35 Hz was calculated using Welch's method. Spectral features were extracted from all channels of the resting state EEG recordings using the method described in [6]. This method is based on the estimation of the individual noise floor of each participant and the calculation of the difference between the alpha and the beta peak amplitudes and the noise floor. The maximum amplitude of these two peaks and the frequency of the maximal peak in Hz are used as features for the prediction of P300 BCI performance.

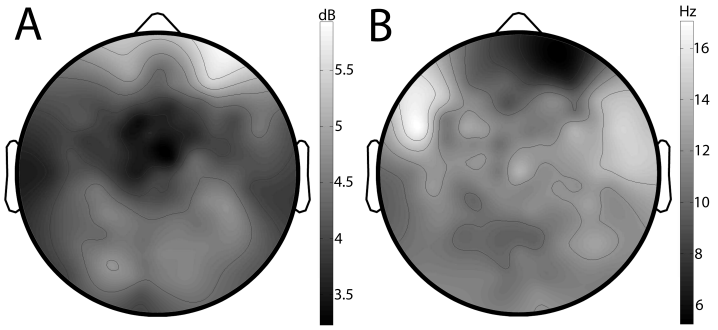


Figure 1: Spatial distribution of mean peak amplitudes (A) in dB and frequencies (B) in Hz of all participants during rest.

3 Results

3.1 BCI Performance

The forty participants reached a mean online accuracy of 94.5 % (SD 14.7, range 35–100). Offline classification using three flashes of each row and column per selection lead to a mean accuracy of 71.8 % (SD 25.1, range 0–99).

3.2 Aptitude Prediction

Mean dB values of maximal spectral peak amplitudes and their frequency in Hz are shown in Figure 1. E.g. mean amplitude during rest on electrode Cz was 3.8 dB with a frequency of 13.5 Hz and on electrode Oz at 4.5 dB with a frequency 10.7 Hz. Correlations between P300 BCI aptitude and peak amplitude were found between $r = -0.18$ and $r = 0.25$, but none were significant (see Figure 2 (A)). In contrast correlations between aptitude and the frequency of the maximum peak were found between $r = -0.72$ and $r = 0.36$ that were significant (see Figure 2 (B)). The strongest correlation between aptitude and frequency was found on channel CCP6 ($r = -0.72$, $p < 0.000001$; see Figure 2 (C)).

4 Discussion

It has been shown that alpha rhythm amplitude is related to P300 amplitude and alpha rhythm frequency to P300 latency [7]. In our data we could not find significant relations between the EEG resting state peak amplitude and the P300 BCI performance but instead between the frequency of the resting state peak and performance. If the observation made in [7] holds for ERPs elicited by a P300 BCI paradigm this would indicate a stronger effect of latency on P300 BCI performance than amplitude.

5 Conclusion

The fact that resting EEG data from a different day contains such strong predictors of P300 BCI aptitude is an indication of the stability of the presented BCI aptitude prediction method. Further investigations will have to be conducted to better understand the nature of the connection between the spectral peak frequency and P300 BCI aptitude.

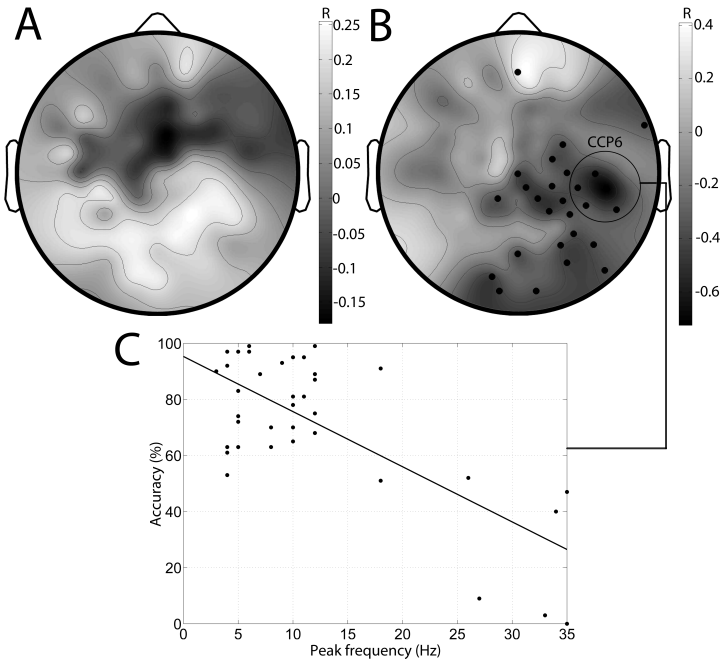


Figure 2: Distribution of correlation between P300 BCI aptitude and peak amplitude (A) or peak frequency (B). Electrodes at which the correlation between aptitude and predictive measure was significant are marked with a black dot ($p < 0.05$). There are no significant correlations between peak amplitude and aptitude. An exemplary correlation between aptitude and peak frequency ($r = -0.72, p < 0.000001$) is shown for electrode CCP6 (C).

References

[1] M. van Gerven, J. Farquhar, R. Schäfer, R. Vlek, J. Geuze, A. Nijholt, N. Ramsey, P. Haselager, L. Vuurpijl, S. Gielen, and P. Desain. The brain-computer interface cycle. *J Neural Eng*, 6(4):041001, Aug 2009.

[2] N. Birbaumer, N. Ghanayim, T. Hinterberger, I. Iversen, B. Kotchoubey, A. Kübler, J. Perelmouter, E. Taub, and H. Flor. A spelling device for the paralysed. *Nature*, 398(6725):297–8, Mar 1999.

[3] A. Kübler, F. Nijboer, J. Mellinger, T. M. Vaughan, H. Pawelzik, G. Schalk, D. J. McFarland, N. Birbaumer, and J. R. Wolpaw. Patients with ALS can use sensorimotor rhythms to operate a brain-computer interface. *Neurology*, 64(10):1775–7, May 2005.

[4] L. A. Farwell and E. Donchin. Talking off the top of your head: toward a mental prosthesis utilizing even-related brain potentials. *Electroencephalogr Clin Neurophysiol*, 70:510–523, 1988.

[5] F. Nijboer, E. W. Sellers, J. Mellinger, M. A. Jordan, T. Matuz, A. Furdea, S. Halder, U. Mochty, D. J. Krusienski, T. M. Vaughan, J. R. Wolpaw, N. Birbaumer, and A. Kübler. A P300-based brain-computer interface for people with amyotrophic lateral sclerosis. *Clin Neurophysiol*, 119(8):1909–16, Aug 2008.

[6] B. Blankertz, C. Sannelli, S. Halder, E. M. Hammer, A. Kübler, K. R. Müller, G. Curio, and T. Dickhaus. Neurophysiological predictor of SMR-based BCI performance. *Neuroimage*, 51(4):1303–9, Jul 2010.

[7] J. Intriligator and J. Polich. On the relationship between background EEG and the P300 event-related potential. *Biol Psychol*, 37(3):207–218, 1994.

Exploring the Resonant Properties of Alpha-Band in Covert Spatial Visual Attention

P. Horki[1], C. Pokorny[1], C. Neuper[1,2], G. R. Müller-Putz[1]

[1]Institute for Knowledge Discovery, BCI Lab, Graz University of Technology, Graz, Austria
[2]Department of Psychology, University of Graz, Graz, Austria

petar.horki@tugraz.at

Abstract

Previous studies demonstrated that individually selected EEG channel locations lead to improved classification in covert spatial attention experiments. However, it is unclear whether an advantage, i.e. an increase in accuracy, can be gained in the selection of stimulation frequencies. We hypothesised that individual stimulation frequencies in the alpha-band may improve performance compared to fixed stimulation frequencies outside the alpha-band. To test this hypothesis, we performed offline classification for both modalities in a 2-class covert spatial attention experiment with a wearable visual stimulation unit.

1 Introduction

The so called steady-state visual evoked potentials (SSVEP) can be elicited by presenting visual stimuli at a rate higher than 6 Hz [1]. In the context of the covert spatial attention (i.e. attending to targets in space outside the foveal vision), left/right spatial attention can be detected by extracting the SSVEPs elicited by the stimuli [2].

Previous studies demonstrated that individually selected EEG channel locations lead to improved classification both in covert and in overt experiments [3,4]. One question left open in covert experiments is whether an advantage, i.e. an increase in accuracy, can be gained in the selection of stimulation frequencies. Another potential improvement in covert spatial attention experiments is to analyse the SSVEP not only at the stimulus frequency, but also at its harmonics [5].

The resonant properties of the alpha-band are well known, but not fully understood. Our hypothesis is that individual stimulation frequencies in the alpha-band may lead to an enhanced performance in covert spatial attention paradigm compared to fixed stimulation frequencies outside the alpha-band. We test this hypothesis by calculating the offline accuracy for both modalities in a 2-class paradigm. Furthermore, to account for patients suffering from severe motor impairment we use a wearable stimulation unit.

2 Methods

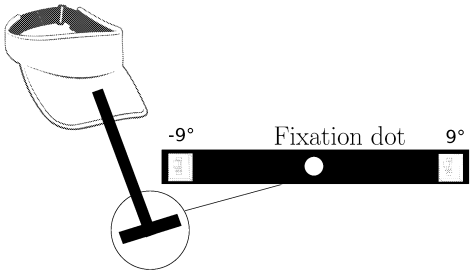
2.1 Stimulation Unit, Subjects and EEG Recording

In the initial screening measurement individual stimulation frequencies were selected. To this end, stimuli were delivered via a single white LED (3.2 mm×1.6 mm) flickering at 9 individual alpha-band frequencies centered arround the subject specific α-peak (1 Hz radius, 0.25 Hz resolution, random order), with a duty cycle of 50 %. In an additional stimulus condition the LED was constantly illuminated for the duration of the trial.

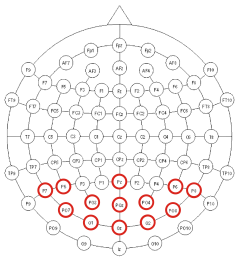
In the covert spatial attention experiment a wearable visual stimulation unit was used. To this end, two bright white LEDs (3.2 mm×1.6 mm), a battery power supply and a microcontroller board were integrated into a golf cap(Figure 1). The LEDs, connected to the microcontroller board, were programmed independently to flicker at desired frequencies.

Ten healthy subjects (9 male and 1 female) participated in the experiment. They reported no history of neurological illness and had normal or corrected-to-normal vision. None of the subjects participated in this or any other similar, i.e. evoked potentials based, experiment before.

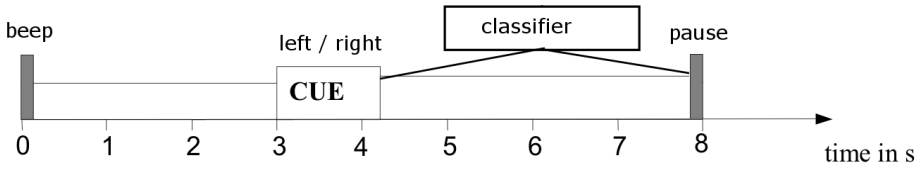
EEG was recorded by thirteen active Ag/AgCl electrodes (g.USBamp, g.tec Guger Technologies, Graz, Austria) positioned over the occipito-parietal cortex (Figure 1). Vertical and horizontal EOG was recorded with three active electrodes. In detail, two EOG electrodes were placed on the outer canthi of the eyes and one was placed superior to the nasion. Reference and ground electrodes were placed at the left earlobe and AFz channel position, respectively. The EEG amplifier was set up for a band-pass filter of 0.5 to 100 Hz. The notch filter (50 Hz) was on and the sampling rate was $fs = 512$ Hz.



(a) Wearable stimulation unit.



(b) EEG recording.



(c) Experimental paradigm.

Figure 1: a) Wearable visual stimulation unit; b) EEG electrode positions (international 10-20 system); c) Experimental paradigm used in the covert spatial attention experiment.

2.2 Experimental Paradigm

For the screening measurement subjects were seated about 0.75 m in front of the eye-level monitor. The cue-based experiment consisted of five runs containing 40 trials each, separated by breaks to avoid fatigue. Subjects were instructed to covertly attend the flickering light, placed midline upwards at eccentricity of 9 ° relative to the fixation point, according to the cue-based paradigm.

Each trial lasted 7 s and was defined as follows: 1. At the beginning of each trial ($t = 0$ s), a fixation cross was presented at the center of the screen, and remained visible until the end of the trial. 2. From $t = 0$ to 2 s, the participants looked at the fixation cross. 3. A short tone at $t = 2$ s indicated to the participants that they should focus their attention on the flickering light without moving their eyes. 4. At $t = 7$ s the fixation cross dissapeared, indicating the end of a trial.

For the covert spatial attention experiment subjects were seated in a comfortable armchair. They were instructed to covertly attend to flicker stimuli of the wearable stimulation unit in the left or right visual field while maintaining gaze at a fixation dot (Figure 1). This cue-based experiment consisted of 8 runs containing 40 trials each, separated by short breaks of about 1–2 min to avoid fatigue. In one half of the runs, i.e. in the first 4 or in the last 4 runs, the flickering frequencies were set to 8 Hz (left) and 13 Hz (right). In the other half of the runs, the flickering frequencies were set to the individual values obtained from the previous screening measurement (two strong SSVEP frequencies were selected). The starting frequencies of the flickering lights were counterbalanced.

Each trial lasted 8 s and was defined as in Figure 1 c). Each flickering light was randomly indicated (verbal instructions) 20 times within each run resulting in 80 trials for each of the

subj	AY3	BV1	BV2	BV5	BV6	BV9	μ
$acc_{8,13}$	60 %	62 %	66 %	68 %	66 %	67 %	65 %
acc_{α}	63 %	59 %	67 %	67 %	69 %	66 %	65 %
f_L	10.50	10.75	9.00	10.75	9.50	11.50	
f_R	9.50	9.25	8.50	10.25	10.00	10.75	
α	10.00	10.25	9.25	11.00	9.75	11.00	

Table 1: Classification accuracies obtained for within the alpha-band (individual) and for 8 and 13 Hz stimulation frequencies. Also shown are the individual stimulation frequencies and the corresponding alpha peak.

SSVEP frequencies. During the whole experiment the subjects did not receive any feedback.

2.3 Data Processing

Data from subjects were visually inspected and trials containing EMG artifacts were discarded. Trials with horizontal or vertical EOG amplitude exceeding $\pm 50\,\mu$V were also excluded from the analysis.

Single-trial EEG epochs were derived in association with each SSVEP cue. To detect the SSVEP frequencies a power spectrum analysis was performed on each of the EEG channels and for each condition separately. The SSVEP response was analyzed at the corresponding flicker frequency, as well at its second and third harmonic.

2.3.1 Feature Extraction

The frequency components were computed by estimating the power density spectrum (4*fs-point FFT, rectangular 4 s window) of the EEG signal for the activation interval between one second after a cue and the end of a trial. The three spectral components around the target frequency (e.g. for 8 Hz the components 7.75, 8.0 and 8.25 Hz) were averaged. The harmonic sum at each stimulation frequency and at each channel was computed [4].

2.3.2 Classification

Fisher's linear discriminant analysis method (LDA) was used as a classifier. An exhaustive search was conducted to determine optimal bipolar derivations for each subject. All combinations of two bipolar channels were classified using four dimensional features constructed from SSVEP magnitudes extracted from both channels. A 10×10 cross-validation was applied to estimate the accuracy for the activation interval. The classification accuracy calculated from these features is presented within the results section.

3 Results

EEG data from four subjects were excluded from the analysis due to excessive EOG and EMG artifacts, because of which more than 50 % of trials was discarded. For the remaining six subjects, the average trial rejection rate was 28 %.

In Table 1 the alpha-band stimulation frequencies with strongest SSVEP response, considering the first three harmonics, together with the corresponding alpha-band peak for each subject are shown. The most important resonant frequency was around 10 Hz.

Also presented in Table 1 is a summary of the accuracies calculated for 8 and 13 Hz stimuli, as well as for individual stimuli. In both modalities the mean accuracy was 65 % for six subjects.

4 Discussion and conclusion

No significant difference could be found between classification accuracies obtained with standard and individual stimulation frequencies. Thus, the hypothesis of an improvement of classification accuracy by individually selecting stimulation frequencies was not confirmed. Four out of six subjects analysed reached significant (greater than 65 % [6]) accuracy in both modalities. The invariance of classification accuracy may partially be attributed to the modest performance of participants. It is reasonable to assume, that further training sessions will lead to improved performance.

A novelty in our approach is the use of a wearable LED stimulation unit, previously used only in overt SSVEP experiments [7]. By using a wearable stimulation unit, patients suffering from severe motor impairment can be accounted for. Also, LEDs allow for precise and flexible delivery of visual stimuli. One possible improvement may be to increase the size of the LEDs and to reduce the flicker intensity. Also, colored LEDs can be used.

One disadvantage of SSVEP is that they may cause fatigue during prolonged usage. However, a previous study [8] already demonstrated that by using an imagery-based "brain switch", the flickering lights can be turned "on" and "off", thereby reducing the rate of unintended activations. Similarly, the alpha-band modulation induced by covert spatial attention could be used to activate the flickering lights only when needed, and to deactivate the flickering lights during the resting period.

Acknowledgments

We are grateful to J. Faller for providing the phototransistor circuit used to assess the technical performance of the LED stimuli, G. Clauzel for helpful advice regarding the stimulation unit construction, and to V. Kaiser, J.Wagner and T. Solis-Escalante for their input and suggestions. This research was supported by the European ICT Programme Project FP7-247919 (DECODER).

References

[1] D. Regan. *Human Brain Electrophysiology: Evoked Potentials and Evoked Magnetic Fields in Science and Medicine*. Elsevier, New York, 1989.

[2] S. P. Kelly, E. C. Lalor, C. Finucane, G. McDarby, and R. B. Reilly. Visual spatial attention control in a independent brain-computer interface. *IEEE Trans Biomed Eng*, 2005.

[3] S. P. Kelly, E. C. Lalor, R. B. Reilly, and J. J. Foxe. Visual spatial attention tracking using high-density SSVEP data for independent brain-computer communication. *IEEE Trans Neural Syst Rehabil Eng*, 2005.

[4] G. R. Müller-Putz, E. Eder, S. C. Wriessnegger, and G. Pfurtscheller. Comparison of DFT and lock-in amplifier features and search for optimal electrode positions in SSVEP-based BCI. *J Neurosci Meth*, 2008.

[5] G. R. Müller-Putz, R. Scherer, C. Brauneis, and G. Pfurtscheller. Steady-state visual evoked potential (SSVEP)-based communication: impact of harmonic frequency components. *J Neural Eng*, 2005.

[6] G. R. Müller-Putz, R. Scherer, C. Brunner, R. Leeb, and G. Pfurtscheller. Better than random? A closer look on BCI results. *Int J Bioelectromagnetism*, 10:52–55, 2008.

[7] T. Lüth and A. Gräser. Wearable stimulator for SSVEP-based brain-computer interfaces. In *Proc. 11th IEEE Intl Conf Rehabil Rob*, 2009.

[8] G. Pfurtscheller, T. Solis-Escalante, R. Ortner, P. Linortner, and G. R. Müller-Putz. Self-paced operation of an SSVEP-based orthosis with and without an imagery-based "brain switch": a feasibility study towards a hybrid BCI. *IEEE Trans Neural Syst Rehabil Eng*, 2010.

Correlating SSSEP Screening Results with BCI Performance

C. Breitwieser[1], C. Pokorny[1], V. Kaiser[1], C. Neuper[1,2], G. R. Müller-Putz[1]

[1]Institute for Knowledge Discovery, BCI Lab, Graz University of Technology, Graz, Austria
[2]Department of Psychology, University of Graz, Graz, Austria

c.breitwieser@tugraz.at

Abstract

Relative band power (BP) increase during a steady-state somatosensory evoked potentials (SSSEP) screening measurement and classification results from an attention modulation paradigm based on SSSEP have been correlated to determine if high relative BP values imply good BCI classification results. Vibro-tactile stimulation has been applied to the thumb and the middle finger of the right hand of 14 participants to elicit SSSEPs. Person specific resonance-like frequencies have been determined in a screening session. Two frequencies for simultaneous stimulation used in a focused attention paradigm were selected based on the screening. In them focused attention paradigm, participants switched their attention either to the thumb or the middle finger. Recorded data was classified offline using a lock-in amplifier system (LAS) and a linear discriminant analysis (LDA) classifier. Classification results were compared with the relative band power increase during the screening session to determine if those values are correlated in some way. No significant correlation between the relative band power increase and the classification results could be observed.

1 Introduction

Various strategies to establish a Brain-Computer Interface (BCI) are used nowadays. Examples are BCIs based on sensory motor rhythms [1] or systems utilizing evoked potentials like P300 [2] or steady-state visual evoked potentials (SSVEPs) [3]. In general, using steady-state evoked potentials, external stimulation causes an amplitude increase in the electroencephalogram (EEG) in the range of the external stimulation frequency. An amplitude in- or decrease can be modulated by shifting attention to or from a respective target stimulus, in case of SSVEP a flickering light. The same principle has already been successfully applied to set up a BCI utilizing steady-state somatosensory evoked potentials (SSSEPs) [4], where SSSEPs have been elicited by applying vibro-tactile stimulation to the index fingers of both hands. Participant specific stimulation frequencies were obtained via a simple screening process, as every person seems to react with different resonance-like frequencies to tactile stimulation [5]. Classification accuracies between 70 % and 80 % could be reached.
However, a relation between the classification results and the relative BP increase was not analyzed. Therefore, it is the goal of this work to investigate, whether a high relative BP increase during a screening measurement already indicates classification accuracies above chance for BCIs.

2 Methods

The fingertips of thumb and middle finger of the right hand from 14 healthy persons (7 male, 7 female, mean age: 25.64 ± 2.6 years) were stimulated using C2 tactors [Engineering Acoustics, Inc., Casselberry, FL, USA]. A 200 Hz sine modulated with a rectangular signal as mentioned by Müller-Putz et al. [4] was utilized for stimulation. A screening measurement to determine the participants resonance-like frequency and a focused attention paradigm have been conducted.

EEG was recorded with 2.4 kHz using 48 Ag/AgCl electrodes with linked references, the ground electrode was placed at the participants nose. All impedances were kept below 5 kΩ, measurements were done in a shielded room.

2.1 Measurement Procedure

To examine person specific resonance-like frequencies, the participants thumb was stimulated with frequencies from 13–35 Hz in 2 Hz steps in a screening measurement. Resonance-like frequencies were assumed to be similar on all fingers [6]. Participants were distracted by performing a mathematical task to avoid shifting attention unintentionally to the stimulated finger. During the screening, every frequency was stimulated randomly 60 times for 2 s. Reference periods without stimulation were inserted before every twelfth stimulation. Two frequencies for simultaneous stimulation to be used in the focused attention paradigm were selected based on the screening measurements. Therefore, FFT maps as shown in Figure 1(a) were calculated based on the screening data. The bipolar channel FC3-CP3 was selected for manual inspection based on FFT power spectra. Two stimulation frequencies showing a high amplitude increase were selected as shown in Figures 1(b) and 1(c). The focused attention paradigm was divided into multiple trials. Every trial consisted of a 3–3.5 s reference followed by a 4–4.5 s focused attention period. During both periods stimulation was applied simultaneously with frequencies selected in the prior screening. Participants had to shift their attention to a target finger, presented on a computer monitor. During reference subjects were instructed to merely look at the blank screen. Every class (target finger to focus on) was repeated 80 times.

2.2 Correlation of Relative Band Power and Classification Results

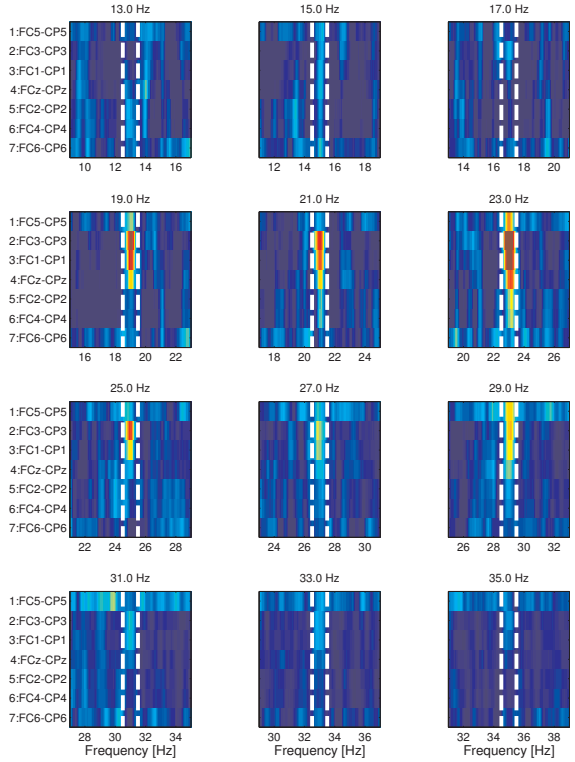
Relative band power (BP) tuning curves (as shown by [5]) were calculated relating the BP during stimulation with a single frequency to the BP of the respective reference interval. Recorded data was manually inspected, trials consisting EMG artifacts were discarded. Data recorded during the focused attention paradigm was offline classified using the amplitude output of a lock-in amplifier system (LAS) [7] as a feature and a linear discriminant analysis (LDA) as a classifier. Classification results were 10×10 cross-validated. The Pearson product-moment correlation coefficient (PPMCC) between the relative BP increase for selected frequencies for the focused attention paradigm and the maximum accuracy for the respective class were investigated.

3 Results

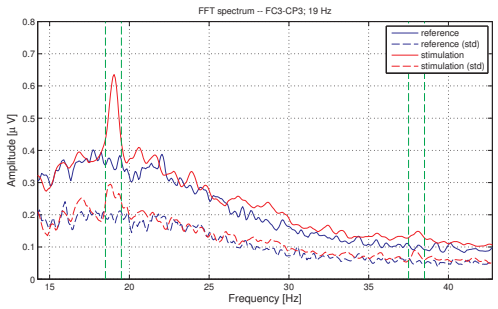
All shown figures presented are related to participant s1, which was randomly chosen. An exemplary relative BP tuning curve for different bipolar channels is shown in Figure 2. The emergence of a tuning curve over channels covering the left hemisphere is visible. Table 1 shows the maximum accuracy and the relative BP increase during the screening paradigm. The mean accuracy over all participants was 66,8 % ($\pm 5,7$) for shifting attention to the thumb against reference and 66,6 % ($\pm 5,1$) for shifting attention to the middle finger, also against reference. The thumb was stimulated with the first selected frequency and is related to rBP 1, the middle finger with the second frequency and is thus related to rBP 2. No significant correlation between either the relative BP increase (rBP) and the maximum accuracy (ACC) for both classes could be found (thumb: rBP–ACC: r = .09, n.s.; middle finger: rBP–ACC: r = .26, n.s.).

4 Discussion

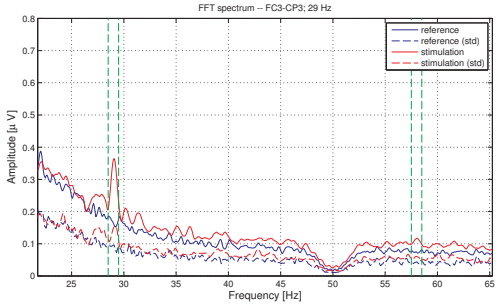
Within this work we determined two participant specific stimulation frequencies based on FFT calculations. Those frequencies were applied in a focused attention task to form an offline BCI. The maximum achieved classification accuracy was subsequently correlated with the relative BP increase during the screening session.



(a) FFT maps used do determine stimulation frequencies with highest SSSEP response. Every plot shows an FFT map during stimulation. The x-axis shows the frequency, and the y-axis shows seven bipolar channels over the primary sensory and motor regions. Blue colored regions indicate a low amplitude, yellow and red colored sections show an increased amplitude. Applied stimulation frequencies are highlighted with white dashed lines.



(b) Power spectrum of bipolar channel FC3–CP3 during stimulation with 19 Hz used for manual inspection. The red lines represent the frequency response and its standard deviation during stimulation, the blue lines during reference period. The stimulation frequency and its 2^{nd} harmonic are highlighted with green dashed lines.



(c) Power spectrum of bipolar channel FC3–CP3 during stimulation with 29 Hz. Despite a different stimulation frequency, this image is similar to Figure 1(b).

Figure 1: FFT maps and power spectra from participant s1.

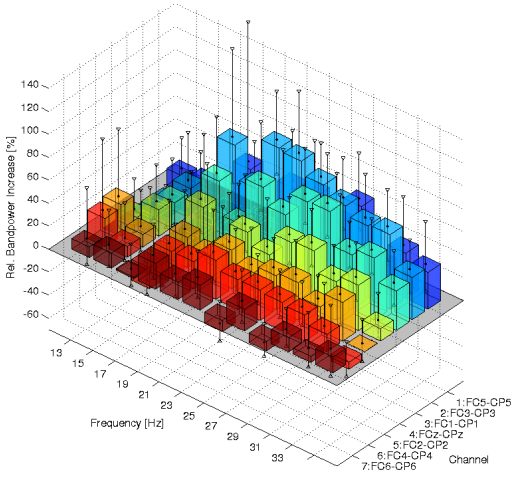


Figure 2: Tuning curves showing the relative BP increase of seven bipolar channels for participant s1. The x-axis shows the stimulation frequencies, the y-axis the channels and the z-axis the rel. BP increase. The 95 % confidence interval is plotted with vertical lines.

	s1	s2	s3	s4	s5	s6	s7
rBP 1	80	75	300	190	100	550	140
rBP 2	75	140	450	380	240	600	210
thumb	74	72	64	76	69	70	59
mid. fi.	73	64	65	78	64	69	66
	s8	s9	s10	s11	s12	s13	s14
rBP 1	175	75	75	150	100	75	175
rBP 2	100	50	250	100	300	175	175
thumb	73	64	67	63	61	65	58
mid. fi.	72	65	66	58	68	64	61

Table 1: Relative BP increases (in %) for person specific stimulation frequencies (rBP 1 and rBP 2) during screening for all participants. The maximum classification accuracy (also in %) for shifting attention on the thumb or the middle finger during the focused attention paradigm are also shown.

Broad person specific tuning curves from 19 Hz to 33 Hz could be observed, similar to the findings from Müller et al. [5]. These effects are in contrast to a narrow tuning curve with a person independent maximum at 21 Hz reported by Tobimatsu et al. [8]. The reasons for these differences are still an open question, however Tobimatsu et al. used a different experimental setup.
Due to a small number of 14 participants, the PPMCC showed no significant results. But considering Table 1 it can be seen that participant s1, showing a relative BP increase lower than 100 % reached classification results above 70 %. Furthermore, the participant achieving maximum accuracy (s4) did not have the highest relative BP increases over all participants. Thus it seems that it is not possible to conclude good or bad performance just from a screening measurement and exclude people out of this reasons from further measurements. Online measurements are planned as a next step, where participants should perform online control. Investigations towards an optimal channel selection are scheduled. Furthermore it is planned to additionally including higher harmonics into the classification procedure, as this approach already significantly increased the performance of SSVEP BCIs [7]. Any effect due to this planned modifications to the correlation of relative BP increase and classification results will be observed.

5 Acknowledgments

This work is supported by the European ICT Programme Project FP7-247919. 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] G. Pfurtscheller and C. Neuper. Future prospects of ERD/ERS in the context of brain-computer interface (BCI) developments. In Christa Neuper and Wolfgang Klimesch, editors, *Event-Related Dynamics of Brain Oscillations*, volume 159 of *Progress in Brain Research*, pages 433 – 437. Elsevier, 2006.

[2] L. A. Farwell and E. Donchin. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalogr Clin Neurophysiol*, 70(6):510–523, Dec 1988.

[3] 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, Jan 2008.

[4] G. R. Müller-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.

[5] G. R. Müller, C. Neuper, and G. Pfurtscheller. Resonance-like frequencies of sensorimotor areas evoked by repetitive tactile stimulation. *Biomed Tech (Berl)*, 46(7-8):186–190, 2001.

[6] C. Breitwieser. Are steady-state somatosensory evoked potentials suitable for biometric use? Master's thesis, Graz University of Technology, 2009.

[7] G. R. Müller-Putz, E. Eder, S. C. Wriessnegger, and G. Pfurtscheller. Comparison of DFT and lock-in amplifier features and search for optimal electrode positions in SSVEP-based BCI. *Journal of Neuroscience Methods*, 168:174–181, 2008.

[8] S. Tobimatsu, Y. M. Zhang, and M. Kato. Steady-state vibration somatosensory evoked potentials: physiological characteristics and tuning function. *Clinical Neurophysiology*, 110:1953–1958, 1999.

Do User-Related Factors of Motor Impaired and Able-Bodied Participants Correlate with Classification Accuracy?

E. V. C. Friedrich[1,2], R. Scherer[1], J. Faller[1], C. Neuper[1,2]

[1]Institute of Knowledge Discovery, Graz University of Technology, Austria
[2]Department of Psychology, University of Graz, Austria

elisabeth.friedrich@uni-graz.at

Abstract

This study investigated (1) if able-bodied and motor impaired participants differ in the evaluation of BCI tasks with respect to the quality of imagery, ease and enjoyment and (2) if prior experience and task evaluation correlate with classification accuracy. This article includes data from three studies that included word association, mental subtraction, spatial navigation and motor imagery as control strategies for a mental-imagery based brain-computer interface (BCI). Task evaluation and prior experience were surveyed with self-reports. Classification was based on common spatial patterns and Fisher's linear discriminant functions. The results showed that motor impaired participants enjoyed performing the mental tasks generally less than able-bodied participants. Task evaluation and classification accuracy was especially low in the motor imagery task for the motor impaired users in contrast to the able-bodied participants. Although prior experience and task evaluation showed some correlations with classification accuracy, it was not possible to explain the classification results with these two factors. To conclude, motor imagery tasks might not be the best choice for motor impaired users for BCI control and therefore, more efforts should be made to find user-appropriate tasks for individuals with motor disabilities.

1 Introduction

It is well known that not every user can learn to control a mental imagery-based brain-computer interface (BCI). However, user-related factors that influence the ability to achieve BCI control are by far unclear. Therefore, this study examined correlations of prior task experience and task evaluation with classification accuracy. This article includes data from three studies that aimed to investigate BCI performance when using different mental tasks for BCI control. The tasks word association, mental subtraction, spatial navigation and motor imagery were included as control strategies for able-bodied as well as motor impaired participants. Especially, for motor impaired people, who might use a BCI in their every-day life, user's experienced quality of imagery, task ease and enjoyment are crucial factors. Therefore, we investigated if the mentioned mental tasks are suitable for motor impaired participants and if there are differences between able-bodied and motor impaired users with respect to task evaluation and classification accuracy. Furthermore, it was researched if the task evaluation or prior experience in task related domains showed correlations with classification accuracy at all in a large sample including participants from all three studies.

2 Methods

2.1 Participants

Study 1 included 9 able-bodied participants (20–32 years, 9 female) who participated in 4 sessions without feedback [1]. Study 2 included 12 motor impaired participants (20–57 years, 7 female)

who participated in 2 sessions without feedback. The participants were diagnosed with spinal cord injury (7 people, laison at C3–C5, ASIA A–C) or stroke (hemorrhagic stroke, including 3 individuals with locked-in syndrome due to brainstem stroke). The experimental paradigm and the EEG recordings from 30 electrode positions were the same in study 1 and 2. Study 3 included 14 able-bodied participants (20–35 years, 7 female) who participated in 2 screenings without feedback and then 8 feedback sessions. The EEG was recorded from 29 electrode positions [2].

2.2 Experimental Paradigm

In all studies, the participants were asked to perform the on the screen indicated mental task for 7 s while staying motionless. The mental tasks occurred in randomized order and included: word association (generate as many words as possible that begin with the presented letter), mental subtraction (perform successive elementary subtractions by a presented fixed number), spatial navigation (imagine navigating through a familiar house) and motor imagery (imagine repetitive self-paced movements of the own hand).

2.3 Self-Reports

In all studies, participants indicated how often they performed verbal, arithmetic, spatial, motor tasks or mental training in their daily-life activities on a 5-point rating scale before the first session (e.g. How often did/do you make activities in your daily life that falls in the category 'motor', such as sports, handcraft, sewing, mechanics, etc). Additionally, users evaluated each of the mental tasks on a 5-point rating scale concerning the quality of imagery (1 = no image at all, you only 'know' you are thinking of the object and 5 = perfectly clear and as vivid as normal vision), the ease (1 = very exhausting and full concentration needed and 5 = very relaxing and possible to perform also during major distractions such as activated television, visit of friends or in the traffic) and the enjoyment (1 = no fun at all and very frustrating and 5 = a lot of fun and not frustrating at all) after the session.

2.4 Classification

The classification of the first session without feedback in all three studies was based on common spatial patterns (CSP) and Fisher's linear discriminant functions (LDA) with majority voting [1]. To quantify the quality of detection, for each task true positive detection rates (TPRs) were computed by dividing the number of correctly classified samples with the same time lag from cue onset by the number of trials. Reported TPRs are peak values within the imagery period of a trial.

3 Results

3.1 Comparison of Task Evaluation Between Able-Bodied and Motor Impaired Participants

Data from the first screening session from study 1 (able-bodied) and 2 (motor impaired) were analyzed. Generally, the word association and spatial navigation task were rated easier ($F_{3,57}, = 6.1$, $p < 0.01$) and more enjoyable ($F_{2.4,45} = 6.5$, $p < 0.01$) than the mental subtraction task and more vivid to imagine ($F_{2.1,41} = 7.4$, $p < 0.01$) than both other tasks. However, motor impaired users evaluated the tasks significantly less enjoyable than able-bodied users ($F_{1,19} = 29.6$, $p < 0.01$). Especially, the word association and motor imagery task was rated less enjoyable by motor impaired users ($F_{3,57} = 2.9$, $p < 0.05$). Figure 1 shows that motor impaired participants evaluated the motor imagery task worst in enjoyment and quality of imagery in comparison to the other tasks and in contrast to the able-bodied participants. Furthermore, also the TPR was lowest in the motor imagery task for the motor impaired in contrast to the able-bodied users, Table 1.

	N	Word association	Mental subtraction	Spatial navigation	Motor imagery
Able bodied	9	85.33(2.03)	83.56(2.69)	81.67(2.26)	85.11(2.63)
Motor impaired	12	78.33(3.19)	77.33(2.17)	74.00(2.29)	73.33(2.85)

Table 1: Number of participants (N) and mean TPR (standard error) for the different mental tasks in the first session for able-bodied (study 1) and motor impaired (study 2) participants.

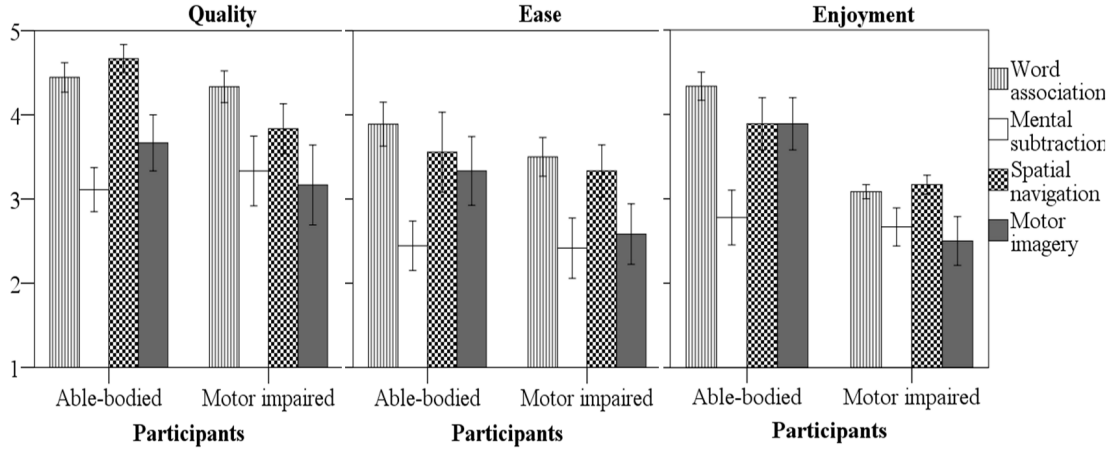


Figure 1: Evaluation of the different mental tasks in respect of quality of imagery, ease and enjoyment on a 5-point rating scale with standard errors 1 = no image at all/very exhausting/very frustrating; 5 = perfectly clear and vivid/very relaxing/very enjoyable) for able-bodied (study 1) and motor impaired (study 2) participants.

3.2 Correlation Between Task Evaluation and Prior Experience and TPR

Data from the first session including 35 participants from all studies were analyzed. The self-reports of experience were positively correlated with the respective evaluation of tasks ($\rho = 0.34$, $p < 0.05$). Especially, the indicated frequency of arithmetical activities showed a significant positive correlation with the general evaluation of the mental subtraction task ($\rho = 0.52$, $p < 0.01$). The frequency of mental training tended to correlate positively with the TPR of the spatial navigation task ($\rho = 0.3$, $p < 0.1$), whereas the experience in motor tasks tended to correlate negatively with the respective TPR ($\rho = -0.32$, $p < 0.1$). The evaluation of the task ease correlated significantly positive with the TPR of the word association task, the motor imagery task and with the mental subtraction task by trend, Table 2. On the contrary, the ease of the spatial navigation task correlated negatively with the TPR. The quality of imagery did not show any significant correlations with classification accuracy. The rated enjoyment of tasks trended to correlate positively with the TPR in the motor imagery task, Table 2.

4 Discussion

The results suggest that motor imagery might not be the best choice for BCI control for motor impaired users in contrast to able-bodied participants. Although studies have shown that motor impaired users are still able to control a BCI with motor imagery [3], selective motor imagery defects in patients with locked-in syndrome in contrast to other mental tasks were reported [4]. Motor impaired users evaluated the motor imagery task towards 'rather frustrating' and reached

TRP	Word association	Mental subtraction	Spatial navigation	Motor imagery
Quality of imagery	-0.179	-0.066	-0.056	0.168
Ease	0.358*	0.324(*)	-0.390*	0.373*
Enjoyment	0.051	0.134	-0.019	0.301(*)

* $p < 0.05$
(*) $p < 0.1$

Table 2: Correlations between task evaluation in respect of the quality of imagery, ease and enjoyment and the TPR of each mental task including 35 participants from all three studies. Significant correlations are marked with *. Significant correlations by trend are marked with (*).

least TPR in this task. A positive correlation between task evaluation and TPR in the motor imagery task was also shown over a large sample including participants from all three studies. Second, prior experience in mental training and more experience in a rather difficult task (e.g. mental subtraction) could be beneficial for BCI performance and task evaluation, respectively. Furthermore, the correlations of classification accuracy with task ease and - at least for the motor imagery task - enjoyment indicate that there is a relationship between BCI usability and performance. Therefore, future studies should additionally focus on improvements of BCI usabiliy. However, the correlations between TPR and self-reports were not consistent and involved also negative correlations. Thus, an initial screening with EEG measurement and task classification is necessary for performance prediction [5]. To conclude, motor imagery tasks might not be the best choice for motor impaired individuals to use a BCI for communication and control. Therefore, more efforts should be made to find user-appropriate tasks for individuals with motor disabilities and to make BCIs generally more enjoyable for those who really need them.

5 Acknowledgments

We thank Ursula Costa, Eloy Opisso and the Institut Guttmann for their support. This work is supported by the Brainable (ICT-2009-4-247447) EU project, the NeuroCenter Styria and the European ICT Programme Project TOBI (FP7-224631). 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] E. V. C. Friedrich, R. Scherer, D. Skliris, and C. Neuper. Mental strategies boost classification accuracy for non-invasive EEG-based brain-computer interface. *BBCI Workshop*, 2009.

[2] E. V. C. Friedrich, R. Scherer, K. Sonnleitner, and C. Neuper. Impact of auditory distraction on user performance in a brain-computer interface driven by different mental tasks. *Clinical Neurophysiology*, 2011.

[3] C. Neuper, G. R. Müller-Putz, A. Kübler, N. Birbaumer, and G. Pfurtscheller. Clinical application of an EEG-based brain-computer interface: a case study in a patient with severe motor impairment. *Clinical Neurophysiology*, 114:399–409, 2003.

[4] M. Conson, S. Sacco, M. Sara, F. Pistoia, D. Grossi, and L. Trojano. Selective motor imagery defect in patients with locked-in syndrome. *Neuropsychologia*, 46:2622–2628, 2008.

[5] B. Blankertz, C. Sannelli, S. Halder, E. M. Hammer, A. Kübler, K. R. Müller, G. Curio, and T. Dickhaus. Neurophysiological predictor of SMR-based BCI performance. *NeuroImage*, 51:1303–1309, 2010.

Data Driven Neuroergonomic Optimization of BCI Stimuli

M. Tangermann[1], J. Höhne[1], M. Schreuder[1], M. Sagebaum[1,2], B. Blankertz[1], A. Ramsay[3], R. Murray-Smith[3]

[1]BBCI Lab, Berlin Institute of Technology, Berlin, Germany
[2]IDA, Fraunhofer First, Berlin, Germany
[3]School of Computing Science, University of Glasgow, Glasgow, Scotland

michael.tangermann@tu-berlin.de

Abstract

Neuroergonomic design of Brain-Computer Interface (BCI) experiments can be realized as a data driven optimization of stimuli. The goal of this process is to increase the number and information content of class-discriminant features of the EEG for the BCI task at hand. While existing electrophysiological literature indicated the influence of confounding variables on e.g. P300 latency and amplitude by group studies and grand average statistics, the BCI performance can be boosted by a large amount when optimized stimuli are explored and designed for individual users and then used in single-trials. The potential of this design principle is shown in an offline analysis for the example of a visual ($n = 8$) and an auditory ($n = 5$) ERP study with healthy subjects, where the optimization of stimulus parameters leads to both a decrease in classification errors and an increase in speed.

1 Introduction

Brain-Computer Interfaces (BCI) are still a very difficult approach to controlling computers, with many people in target user groups unable to learn to use them at all. The ratio of people who cannot use BCI depends on the BCI technology applied, but common problems are caused by the slow and inherently uncertain control of BCI applications due to the low signal-to-noise ratio of the commonly used signals of the electroencephalogram (EEG). There is no proprioceptive feedback, what feedback there is is delayed and the interaction principles and mental tasks in BCI are often unnatural.

The transfer of BCI systems out of the research lab and into the real world will only succeed if the usability of BCI systems can be increased. To bridge the gap, this paper underlines the necessity to follow neuroergonomic [1] design principles in order to optimize BCI stimuli. Ideally, the full process of optimizing a BCI paradigm includes (a) investigations in neuroscience (to understand the covariates of human stimulus perception in BCI paradigms), (b) the modelling of these covariates in order to create improved stimulus methods, and (c) the design of BCI interaction paradigms that exploit the stimulus advantages but respect the restrictions imposed by slow and uncertain control. Steps (a) and (b) can be addressed by data-driven approaches. Step (c) requires careful application of human-computer design which takes these constraints into account (e.g. [2]) in a way that is appropriate for the specific target users.

Basic research in the field of electrophysiology revealed early the sensitivity of EEG phenomena to many covariates. Some examples are the influence of the task difficulty or the target-to-target distance on ERP, pitch and amplitude on auditory ERP, age on the lateralization of sensory motor rhythms (see e.g. [3, 4]).

For a number of reasons, it is difficult to transfer these results to the field of BCI. First, the prevailing scientific methods were group studies and evaluation by grand average statistics. In contrast, single trial analysis and performance optimization for individual subjects are the

predominant methods in BCI research. Second, these studies often concentrate on a single aspect like the the latency or amplitude of the P300 component. In contrast, for BCI the amount of class discriminant information (between target and non-target stimuli) is of importance and it is typically distributed over several ERP components. Third, the varied range of the investigated confounds in older studies is partially out of the practical range for BCI. However, this basic research is stimulating for the field of BCI by providing entry points to conduct specialized studies.

Following neuroergonomic design principles, interesting new experimental paradigms have recently been created for BCI. On the area of visual ERP paradigms the *adjacency problem* and *repetition blindness* have been mitigated by optimized stimulation codes [5, 6] that modify the sequence and layout of highlighting patterns. The *crowding effect* (which limits the usability if gaze control is restricted) has been tackled by centered presentation of stimuli [7]. In the very recent field of auditory BCIs, spatial encoding has been introduced by [8, 9] to improve the discriminability of tones.

This contribution would like to emphasize the importance of sophisticated and individual stimulus design by presenting two brief examples of successful neuroergonomic optimization for BCI. For an auditory ERP paradigm, a thorough screening from relatively long to extremely short stimulus offset asynchrony (SOA) intervals is investigated. For a visual ERP paradigm, the effects of a variety of basic visual transformations are compared. In both cases, changes in BCI performance upon variation in the stimuli were tracked.

2 Methods

To examine the influence of stimulation parameters onto ERP paradigms, results from two studies are examined.

Auditory Oddball Study: In a two-tone auditory oddball paradigm, the influence of SOA (50, 75, 125, 175, 225, 275, 400 and 1000 ms) is investigated. Five healthy volunteers (three male, all non-smokers) participated in a single session of this offline study. A simple two-tone auditory oddball paradigm with counting task was implemented, with 83 % of non-target tones. In eight block-randomized conditions the SOA steps were compared. For each condition and subject, 1296 epochs (216 targets and 1080 non-targets) were available. EEG was recorded from 61 wet Ag/AgCl electrodes placed at symmetrical positions. The band-pass filtered data (0.5 Hz to 40 Hz) was epoched between -150 ms and 1000 ms relative to each stimulus onset. Binary classification of target and non-target epochs was performed using a (linear) Fisher Discriminant Analysis (FDA) with shrinkage regularization. The mean potentials in hand-selected intervals (based on r^2 values) of epochs were taken as features. The binary classification error was estimated by $[5 \times 5]$ cross-validation.

Photobrowser: In a (6×6) matrix layout of 36 photos, six different visual highlighting effects for rows and columns are compared:

- ▶ *Brightness* increase: Photos were increased to high brightness and back.
- ▶ *Inversion*: Photos were stepwise inverted to a color negative and back.
- ▶ *Masking*: Photos were overlaid with a grid of high contrast lines (B&W, magenta).
- ▶ *Rotation*: Photos were rotated clock-wise by 10 degrees and back.
- ▶ *Scaling*: Photos were scaled to 110 % size and back.
- ▶ *Combination*: Photos underwent a combination of the above effects (excluding *Inversion*).

The duration of each effect was 100 ms with 100 ms pause between two highlighting events (SOA of 200 ms). Taking the discretization by the screen update rate of 60 Hz into consideration, an effect was present during 5 to 7 frames. Eight healthy volunteers (three male, all non-smokers) participated in a single session of this offline study. A visual oddball paradigm with counting task was implemented, with 83 % non-targets. The six conditions were presented in a block-randomized order. For each condition, 1800 epochs (300 targets and 1500 non-targets) were available. EEG signals were recorded and processed as described above.

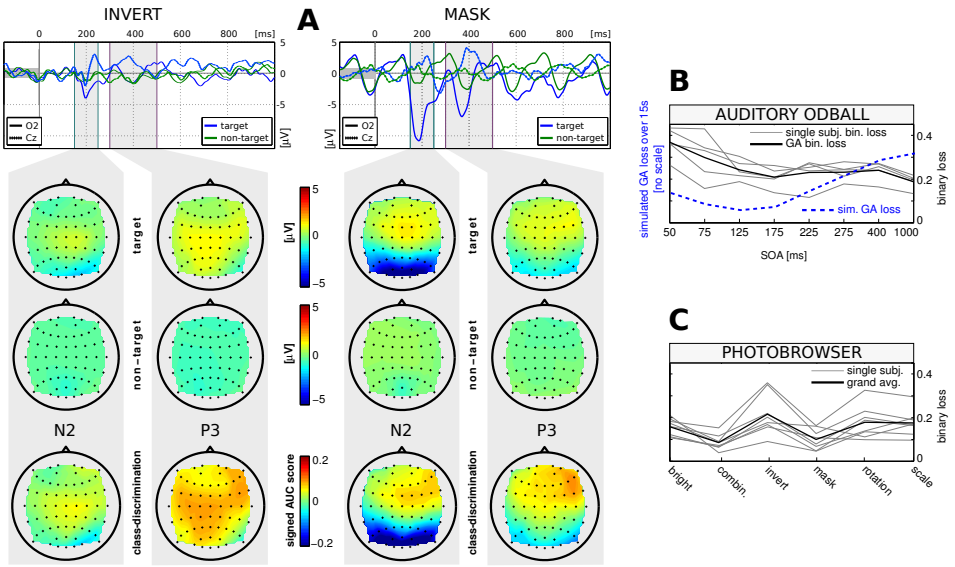


Figure 1: **A**: N2 and P3 for two visual stimulation conditions (one subject). The top row shows ERP traces of electrodes O2 and Cz for *Inversion* and *Masking*. Two middle rows: ERP maps for targets and non-targets. Lower row: spatial distribution of class-discriminant information. **B** and **C**: Estimated class-wise balanced classification loss. **B**: Binary loss (single epochs for single subjects and grand average (GA)) and simulated GA loss for 15 s trial duration for auditory SOA conditions ($n = 5$). **C**: Binary loss for visual highlighting conditions ($n = 8$).

3 Results

The class discriminant information contained in the EEG epoch of a single stimulus remains the major factor for a quick BCI. For this reason the loss derived from the binary classification task of target vs. non-target epochs is used in the following. In addition, an analysis of ERP location and time structure illustrates the influence of a neuroergonomic stimulus design.

An analysis for both studies reveals that the majority of subjects shows typical ERP components. For short SOA values however, early components (P1, N2) dominate the target and non-target ERP responses. Later components (e.g. P3a and P3b) are decreased in amplitude by recurring strong N2 of following stimuli. Even slight changes of the SOA can lead to a complete elimination of class-discriminant late ERP components. This explains the non-monotonic changes of the binary loss for successive SOA values in Figure 1 B for some subjects. For different stimulation conditions, drastic changes of ERP components within each subject were observed. To illustrate this, the ERPs of conditions *Inversion* and *Masking* are compared in Figure 1 A.

Judging by the binary loss estimated by cross validation, the optimal design of stimulation parameters has the potential to improve the individual performance for both study designs. The left plot in Figure 1 C shows that for the Photobrowser study the grand average classification loss based on single epochs can be reduced by ∼50 % compared to the standard visual highlighting effect in the form of a brightness increase. The improvement is strongest for the rather complex combination effect (reduction of loss from 0.16 to 0.085), but even the simple masking effect improves the performance by nearly 40 % (to ∼0.1). If optimized individually, only slightly stronger classification gains are possible. Figure 1 A compares the effect of a brightness highlighting and a mask highlighting. The plot in Figure 1 B shows that a reduction of SOA intervals from 1000 ms to 175 ms or even 125 ms results in minor increases of the grand average binary loss for the auditory oddball paradigm. At first glance, these results seem to favor long SOA values. However, shorter SOA values strongly improve the transmitted information per time (e.g. in terms of simulated loss over 15 s trial duration), especially if BCI applications have a reduced overhead in terms of inter-trial pauses or time needed for result feedback. In this light, the short SOA conditions

SOA_{75}, SOA_{125} and SOA_{175} are to be preferred on average. Condition SOA_{50} shows a strongly increased binary loss on the grand average, such that it is not competitive despite even shorter trial durations. Please note the individual sensitivity even for neighboring SOA steps.

4 Discussion

The data driven optimization of BCI stimuli according to neuroergonomic principles leads to more class-discriminative ERP responses of the EEG. As interpretable models of the underlying functions of perception and focused attention do not come "for free" from these data driven methods, it is difficult to describe the exact causes of the resulting performance improvements.

The improvements were exemplified in an auditory and a visual ERP study not only for the grand average of subjects, but - even more emphatically - when optimizing the stimuli for individuals. In particular, performance improvements were largest in a visual ERP paradigm for a novel visual masking effect that outperforms the predominantly used highlighting methods (rotation, scaling and change of brightness). For an auditory ERP paradigm, short SOA values of 75 to 175 ms showed best results, but the best SOA for an individual is relatively sensitive to fine-tuning. While one of the stimulus types could even be optimized across subjects (visual highlighting effect), the other (SOA) should clearly be selected individually for users.

Acknowledgments

The authors appreciate the financial support by the European ICT Prog. Proj. FP7-224631.

References

[1] R. Parasuraman and G. Wilson. Putting the brain to work: neuroergonomics past, present, and future. *Humfactors*, 50(3): 468, 2008.

[2] J. Williamson, R. Murray-Smith, B. Blankertz, M. Krauledat, and K. R. Müller. Designing for uncertain, asymmetric control: interaction design for brain-computer interfaces. *Int J Hum Comput Stud*, 67(10): 827–841, 2009.

[3] J. Polich and A. Kok. Cognitive and biological determinants of P300: an integrative review. *Biol Psych*, 41(2): 103–146, 1995.

[4] A. Sailer, J. Dichgans, and C. Gerloff. The influence of normal aging on the cortical processing of a simple motor task. *Neurology*, 55(7): 979, 2000.

[5] 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*, 21, 2009.

[6] J. Jin, B. Z. Allison, E. W. Sellers, C. Brunner, P. Horki, X. Wang, and C. Neuper. Optimized stimulus presentation patterns for an event-related potential EEG-based brain-computer interface. *Med Biol Eng Comput*, 2010.

[7] L. Acqualagna, M. S. Treder, M. Schreuder, and B. Blankertz. A novel brain-computer interface based on the rapid serial visual presentation paradigm. *Conf Proc IEEE Eng Med Biol Soc*, 1: 2686–2689, 2010.

[8] 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.

[9] J. Höhne, M. Schreuder, B. Blankertz, and M. Tangermann. Two-dimensional auditory P300 Speller with predictive text system. *Proc IEEE Eng Med Biol Soc*, 1: 4185–4188, 2010.

Can Severe Acquired Brain Injury Users Control a Communication Application Operated through a P300-Based Brain-Computer Interface?

A. Riccio[1], F. Leotta[1], S. Tiripicchio[1], D. Mattia[1], F. Cincotti[1]

[1]Fondazione Santa Lucia, IRCCS Rome, Italy

a.riccio@hsantalucia.it

Abstract

In this study a P300 based prototype for communication and control was evaluated by three Acquired Brain Injury (ABI) users. The prototype was developed by merging an already existing assistive technology commercial solution (Qualiworld) with a BCI system and it allows the user to access to text entry and internet applications. The three ABI users were challenged with four complex tasks and the results revealed good performance levels in standard communication and internet tasks (mean = 73.79 %, SD = 9.4). Two end-users completed the four tasks and reached a performance of respectively 69 % and 76 % on average. One of the end-users did not finish one of the tasks (internet task) and the mean of accuracy for the three completed tasks was 73 %.

1 Introduction

Different neurological disease may lead to severe motor disabilities.

Brain Computer Interface (BCI) represents a promising technology to provide severely motor-disabled people with an alternative channel for sending commands to the external world that does not rely on muscular activity [1]. A list of potential BCI users can include individuals with severe disability because of disorders such as Amyotrophic Lateral Sclerosis (ALS), cerebral palsy, brainstem stroke, spinal cord injuries, muscular dystrophies or chronic peripheral neuropathies. Acquired Brain Injuries (ABIs), which include both Traumatic Brain Injuries (TBI) and strokes, are among the major causes of disability worldwide and affected users could benefit by the BCI.

Despite of this, the few BCI studies conducted over potential users with a P300-based BCI for communication and control, during the last years, were mostly conducted on ALS users ([2–6]). Only in two recent studies two ABI users were included in the sample for testing a visual P300-based BCI: Piccione at al. [7] included one stroke patient in the testing group of a visual P300 based BCI, while Hoffman [8] included in the testing sample one TBI user. In previous studies a system in which the BCI stimulation overlays a commercial software application was described and tested by healthy volunteers ([9, 10]) and by four potential users with degenerative diseases [11]. Both the healthy and patients groups achieved high level (> 70 %) of performance. As Stroke and TBI can result in severly motor disabilities and in a locked-in state and as the development of the medical techniques are improving ABI patients estimated life, they could be potential BCI users. In the present study the same system was evaluated in terms of effectiveness (accuracy) by three ABI patients whose cognitive abilities were previously evaluated.

	Subject A	Subject B	Subject C
Syndrome Aethiology	Hemorrhagic stroke	Ischemic stroke	Hemorrhagic stroke
Age	41	46	24
Months since event	129	24	8
Speech deficit	Disartric	Mutacic	Mutacic
Movement impaired	Tetraparethic	Tetraplegic	Paraplegic
		Alert	
Cognitive Deficit	Vigilance	Vigilance	Divided attention
		Divided attention	

Table 1: Clinical features of participants.

2 Methods

2.1 Prototype

The QualiWorld (QW) platform [12] allows the access to many applications (text editor, web browser etc.), while replacing standard mouse and keyboard by a variety of computer access solutions (Auto-Scan mode, mouse alternatives, gesture recognition). The prototype exploits the BCI2000 P300 signal processing pipeline [13] as brain transducer to access to the QW application.The stimulation is overlaid to the graphical interface of the QualiWorld application suite: the stimuli are directly superimposed on the buttons available on the interface and the stimulation is obtained by means of using flashing dots (color can be selected by the user; 125 ms SD, 125 ms ISI) ([10, 11]).

2.2 Participants and Procedure

Three acquired brain injury end-users were enrolled in the testing of the prototype. They were post-comatose users who had suffered a hemorrhagic stroke (user A and user C) and an ischemic stroke (user B). A neuropsychological screening was performed: the users were screened for verbal and visuo-spatial working memory and for the intensity (alertness and sustained attention) and selectivity aspects of the attention (selective and divided attention). In the Table 1 the patients' clinical profile is reported. None of the users had working memory problems. They presented vigilance (user A and user B) alert (user B) and divided attention (user B and user C) deficits. They were challenged with four complex tasks performed in 4 daily sessions: "word processing in copy mode", "word processing in free mode", "email sending", "internet browsing". In order to cope with the variability of the number of the stimuli available on the screen, the calibration comprised the selection of a total of 20 pre-set letters from four different matrices with different sizes (5 letters per matrix) and was performed during the first session. The "word processing in copy mode" task consisted of writing three pre-established words. The "word processing in free mode" task consisted of the following steps: i) selection of the "word-function" on the graphical user interface (GUI); ii) a selection to open the text-matrix and finally iii) writing an individually chose sentence requiring 20 selections. As for the the email sending task, the users were asked to write a five letters word and send it as an email. This task required the following: the users had to select the e-mail sending application (i.e. select the "send by" icon and the "email" icon), then, a predefined address within an address book had to be selected (i.e "to" icon, "address book" icon, "predefined address") and finally, he had to confirm the address and to send the email ("add" icon, "ok" icon and "send" icon). The internet task consisted of selecting the internet browser application, choosing two specific links of the webpage and performing some browser operations. In particular they were asked to zoom-in the web page previously selected (choosing the predisposed icon), to scroll it down, then to select one more link, zoom-out the page and scroll it down. The total number of selections (without errors) was 21 for the copy spelling, 22 for the free spelling, 17 in the email and 15 in the internet task. The errors had to be corrected in the last three tasks.

	Copy Mode		Free Mode		Email		Internet	
	Accuracy	n. sel	Accuracy	n. sel	Accuracy	n. sel	Accuracy	n. sel
Subject A	85.71 %	21	80 %	36	67 %	31	nc	nc
Subject B	76 %	21	70.5 %	51	55.8 %	43	63 %	38
Subject C	80 %	21	70 %	45	78 %	28	85.7 %	21
Average	80.57 %	21	73.5 %	44	66.93 %	34	74.35 %	29
	± 4.8		± 5.6		± 11.1		± 16.5	

Table 2: Accuracy: (%; number of correct selections/number of total selections of the 3 users challenged in the 4 tasks. N. sel: total number of selection needed to accomplish the tasks.

2.3 Data Acquisition

The EEG data were acquired with an 8-channels cap (Electro-Cap International) with electrodes placed according to the International 10-20 system (Fz, Cz, Pz, Oz, P3, P4, PO7, PO8). All 8 channels were referenced to the right earlobe, and grounded to the right mastoid. Impedances were kept below 10 kΩ. The EEG signals were recorded with a g-tec USB amplifier, bandpass filtered between 0.1 and 30 Hz, and sampled at 256 Hz. All aspects of data collection and experimental procedure were controlled by the BCI2000 system [13]. Data acquired during calibration in the first session were loaded into an off-line analysis software based on Matlab environment, for feature extraction. The Stepwise Linear Discriminant Analysis (SWLDA) was used for signal classification [11].

2.4 Data Analysis

The group of users was too small for inferential statistical analysis, therefore individual data will be reported descriptively. The accuracy was calculated for each task as Correct Response Rate (CRR, number of correct selections/number of total selections).

3 Results

Despite their attention deficit all the three ABI end-users were able to accomplish the four tasks except the end-user A who did not terminate the “internet browsing” task.

The average of the overall performance was 73.79 % with a standard deviation of 9.4. The accuracy and the number of selections needed for each task are reported in the Table 2. The performance of the user A was high in the “copy mode spelling” task and “free mode spelling” task and had a drastic decline in the more complex “email sending” and “internet browsing” tasks, probably because of the vigilance deficit of the user. Even the performance of the user B was lower in the in the email and internet tasks than in the two spelling tasks and her general performance was lower in comparison to the other users. The user B had more attention deficit than the others, in particular alertness, vigilance and divided attention deficit. The user C, who had only a slight divided attention deficit, did not result compromise in any other cognitive domain and did not suffer decline in performance due to increasing of attention load.

4 Discussion

For the first time a P300-BCI was integrated with an assistive technology commercial software and was tested with several tasks by severly disabled end-users whose attention and memory capabilities were previously screened. These preliminary findings revealed good performance levels (> 70 %) on average and showed that attention deficits should not prevent in proposing the use of visual P300-based BCI as one input channel to operate AT device. Although with caution due to the small sample size, we can preliminary drawn the conclusion that specific attention deficit

might affect BCI performance in specific applications. This should be taken into account when designing BCI systems for daily life use.

Acknowledgments

This work is supported by the European ICT Programme Project FP7-224631 (TOBI). This paper only reflects the authors' views and funding agencies are not liable for any use that may be made of the information contained here in.

References

[1] A. Kübler and K. R. Müller. *An introduction to brain computer interface*. 2007.

[2] E. W. Sellers and E. Donchin. A P300-based brain-computer interface: initial tests by ALS patients. *Clinical Neurophysiology*, 2006.

[3] F. Nijboer, E. W. Sellers, J. Mellinger, M. A. Jordan, T. Matuz, A. Furdea, S. Halder, U. Mochty, D. J. Krusienski, T. M. Vaughan, J. R. Wolpaw, N. Birbaumer, and A. Kübler. A P300-based brain–computer interface for people with amyotrophic lateral sclerosis. *Official Journal of the International Federation of Clinical Neurophysiology*, 2008.

[4] F. Nijboer, N. Birbaumer, and A. Kübler. 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(0), 2010.

[5] G. T. Townsend, B. K. La Pallo, C. Boulay, D. J. Krusienski, G. E. Frye, C. K. Hauser, N. E. Schwarts, 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*, 2010.

[6] E. M. Mugler, C. A. Ruf, S. Halder, M. Bensch, and A. Kler. Design and implementation of a P300-based brain-computer interface for controlling an internet browser. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2010.

[7] 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. *Clin Neurophysiol*, 117(3):531–537, March 2006.

[8] U. Hoffmann, J. M. Vesin, T. Ebrahimi, and K. Diserens. An efficient P300-based brain-computer interface for disabled subjects. *Journal of Neuroscience Methods*, 167(1):115 – 125, 2008. Brain-Computer Interfaces (BCIs).

[9] A. Riccio, F. Leotta, L. Bianchi, F. Aloise, F. Cincotti, and D. Mattia. Evaluation of a P300 overlaid stimulation for controlling an assistive technology software. *TOBI Workshop II: Translational issues in BCI development: user needs, ethics, and technology transfer.*

[10] A. Riccio, F. Leotta, L. Bianchi, F. Aloise, C. Zickler, E. J. Hoogerwerf, F. Cincotti, and D. Mattia. Workload measurement in a communication application operated through a P300-based BCI. *Journal of Neural Engineering.*

[11] C. Zickler, A. Riccio, F. Leotta, S. Hillian-Tress, S. Halder, E. Holz, A. H. George, P. Staiger-Salzer, E. J. Hoogerwerf, L. Desideri, D. Mattia, and A. Kübler. BCI and assistive technology: BCI as input channel in a standard assistive technology ICT-software. *Under Submission.*

[12] QualiLife s.a. *QualiWorld user manual.*

[13] G. Schalk. BCI2000: A general-purpose brain-computer interface (BCI) system. *IEEE Trans. Biomed. Eng.*, 51(6):1034–1043, 2004.

Offline Comparative Classification of Hand Movement Direction from Non-Invasive EEG

G. Clauzel[1], C. Neuper[1,2], G. R. Müller-Putz[1]

[1]Institute for Knowledge Discovery, Laboratory of Brain-Computer Interfaces, Graz University of Technology, Graz, Austria
[2]Department of Psychology, University of Graz, Graz, Austria

guillaume.clauzel@tugraz.at

Abstract

In the event of a spinal cord injury, the motor cortex is usually left unharmed. In this situation a by-pass of the lesion in order to restore part of the patient's mobility would be very helpful. A first step in this direction would be to decode hand movements using a non-invasive brain-computer interface. In the present study, a 8-class 2-dimensional center-out task is carried out, with a final average accuracy of 65 %, with best performance at 83 %.

1 Introduction

While continuous decoding of hand position is at the focus of attention, discretization and classification of the hand movement direction with electroencephalogram (EEG) have only been recently studied. Hammon and al. [1] was able to classify target reaching in a 3-dimensional space with an average accuracy of 63 % for a 3-class setup, and to classify a 4-class 2-dimensional center-out reaching task using a tactile screen with an average accuracy of 57 %. Meanwhile, Waldert and al. [2] classified a 4-class 2-dimensional center-out reaching task using a joystick with an average accuracy of 55 %. The present study aims at improving the actual classification rate, so as to create a viable 8-class classifier. Being able to discriminate between 8-class with a hand movement would create a natural way to control a hand neuroprosthetics or any device acting on a 2-dimensional output, like a computer mouse or wheelchair. It would also enable the user to focus on other BCI mental strategy in a hybrid BCI. More intuitive, implying less time consuming training and hopefully less resources consuming concentration, such a device would greatly improve the use of a BCI by disabled people.

2 Methods

2.1 Subjects

Ten healthy subjects (6 females, aged 23.6 std 2.4) participated in the experiment. None of them had prior experience with EEG measurements and BCI experiments. They were all clearly informed about the procedure, and one subject asked to stop the recording for personal reasons. All subjects were right-handed, and their right-handedness was confirmed using a hand dominance test [3, 4]. Results show an average right-handedness of 13.3 %, with a standard deviation of 5.0 % (results are consistent based on other combined studies involving 1171 right-handed subjects).

2.2 Experimental Setup

The participants performed a 8-class 2-dimensional center-out reaching task. A joystick was tuned with an extension of the control column to 40 cm, increasing the arm displacement and thus enhancing the patterns. This joystick was used to control a cursor on the screen, and to move it to one of the 8 targets which were placed at regular intervals on a circle around the center of the screen. The EEG was measured using 31 electrodes spread over the scalp, with an emphasis on the primary motor cortex. The signal was amplified using a g.USBamp (g.tec medical engineering GmbH, Austria) amplifier, with inside filters set to a 8^{th} order Chebyshev bandpass filter between 0.5 Hz and 100 Hz, and a 4^{th} order notch filter at 50 Hz. The EEG signal was then acquired and synchronized with the joystick trajectories using the TOBI Signal Server v0.1 ([5], www.tobi-project.org/download). MatLab R2010a and Simulink (The MathWorks Inc., Massachussets, US) were used to save the signal and run the paradigm.

2.3 Paradigm

After an initial resting period of 2 s, one of the 8 targets (which were continuously displayed in red) would turn green, instructing the participant to move the cursor to the target. Once the highlighted target was reached, it turned red again, and the participant moved the joystick back to the origin position, and then prepared for the next trial. After 4 s, the highlighted target would turn red even if the participant did not reach it. Participants were asked to move the joystick without overshooting the target, and to favor straight movement to the target rather than fast and inaccurate movement. Then a random break lasting between 0.5 s and 2.5 s occured before the next trial. The experiment had 512 trials, split in 8 runs of 64 trials, with the appearance of each class uniformly being randomly generated. Each participant had a break between runs for as long as desired, averaging five minutes.

2.4 Signal Processing

Artifact processing was carried out in two phases: first, the EEG signal was carefully inspected, and each trial contaminated with artifacts was marked for removal. Secondly, the trajectories were analyzed, and trials where the participants did not reach the target in a straigthful movement were discarded. Trials where the movement was not finished after the 4 s window were also discarded. An average of 19±9 trials per class were removed for each participant.

After the artifact removal process, a general algorithm was created, and no subject specific optimization was performed.

First of all, the discrete joystick position signal was linearized, and all trials were synchronized at 10 % of the maximum absolute distance from the center. Two sets of synchronization were created, one at 10 % during the reaching phase, and the other one at 10 % during the return phase.

The EEG signal was then filtered in both the forward and the reverse directions with a 4^{th} order butterworth bandpass filter between 8 Hz and 30 Hz, in order to remove every eye-related information. The common spatial pattern algorithm (CSP) [6] was then applied to the 31 input channels, over a sliding window of 1 s. The variance of the first two and last two output channels was computed over 0.25 s and smoothed over 0.125 s. For each pair of classes, a classifier based on Fisher's linear discriminant (LDA) was trained on these 4 features with a 10×10 cross-validation, and the CSP and LDA weights at the best classification time on the test set were selected.

The outputs of the 28 binary classifiers were then used to create a unique 8-class classifier based on the majority of votes.

Participant	P1	P2	P3	P4	P5	P6	P7	P8	P9	Mean	Std
Acc FM [%]	56.7	59.5	61.6	83.3	61.8	68.1	72.9	59.3	63.0	65.1	8.4
Acc BM [%]	60.3	58.1	58.5	81.7	53.0	51.7	72.5	48.4	52.9	59.7	10.8

Table 1: Overall accuracies for each subject, at the beginning and the end of the movement. Corresponding chance level is at 24.6 % for an average number of trials of 45 per class.

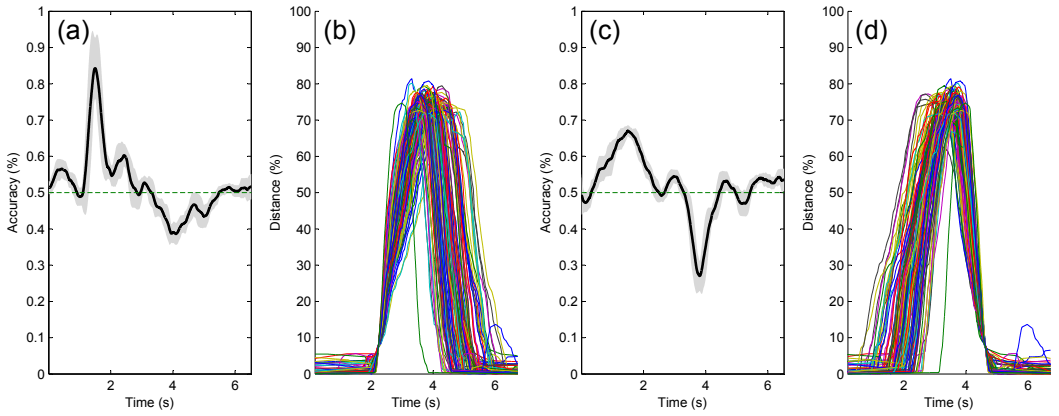


Figure 1: Synchronizations and classifications accuracy for subject P9, class down-right against class up-left. The classifier used was trained with data synchronized at 10 % of the maximum distance on the forth movement (b), and then applied on both the forth movement (b) and the return movement (d). Panels (a) and (c) represent the average classification accuracies and the standard deviations, while panels (b) and (d) show the joystick distance from the center for every trials.

3 Results

As it can be seen in Table 1, the obtained average classification accuracy is around 65.1 % on the forth movement, and 59.7 % for the return movement. These results are above the confidence limit of 24.8 % (for a chance level corrected by the average number of trials of 45.2, with a significance level of 1 % [7]). It is worth noting that the accuracies are similar at the beginning and the end of the movement, showing that the cue had no impact on the classification. The reduction of 5.5 % of accuracy is consistent with the fact that participants were paying less attention on the way back to the center position, and hence less planning and coordination was required. Lowest classification accuracy was for subject P9 at 48.4 %, while the highest classification was for participant P4 with an accuracy above 80 % for both synchronization values.

4 Discussion

Ability of the classifiers to classify the directions is illustrated in Figure 1 where a classifier is trained to discriminate between 2 classes diametrically opposed (e.g. left against right, or down against up). This classifier is trained with synchronization on the forth movement (Figure 1 (b)) and then applied on the signal synchronized on the return movement (Figure 1 (d)). Since the class 1 return direction is the same as the class 2 forth direction, and vice versa, it can be seen that the classifier correctly classifies the return movement as belonging to the opposite class.

Also, it can be seen that maximum accuracies are achieved during planning of the movement.

This is consistent with the fact that the cursor is controlled through a joystick, where the user has first to plan the trajectory, and then correct the ouput of the joystick. For a comparison, Hammon and al. [1] controlled the cursor directly with the hand and obtained similar accuracies during planning and execution of the reaching movement.

Compared with previously cited studies, the present results improve both the classification rate and increase the number of classes from 4 to 8. Moreover, since this experiment was the first contact with BCI experiments for all of the participants, accuracies can be expected to increase with an adequate training. Final results show that a 2 degrees of freedom classifier based on natural movements of the hand is possible with a BCI. Moreover, the 2 degrees of freedom can be used at the same time through the diagonal directions, simplifying the control. Such a system would easily find application to control a computer by imitating a mouse cursor, or to help real-world motion with the control of a wheelchair. For communication, a high accuracy 8-class selection system would enable faster spelling.

5 Conclusion

Results with accuracies as high as 80 % for a 8-class problem show that an acute classification of hand movement directions is possible, therefore enabling new and reliable application for BCIs.

6 Acknowledgments

This work is supported by the European ICT Programme Project FP7-224631. 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] P. S. Hammon, S. Makeig, H. Poizner, E. Todorov, and V. R. de Sa. Predicting reaching targets from human EEG. *IEEE Signal Processing Magazine*, 25(1):69–77, 2008.

[2] S. Waldert, H. Preissl, E. Demandt, C. Braun, N. Birbaumer, A. Aertsen, and C. Mehring. Hand movement direction decoded from MEG and EEG. *The Journal of Neuroscience*, 28:1000–1008, 2008.

[3] H. J. Steingrüber and G. Lienert. Hand-Dominanz-Test. *Verlag für Psychologie Hogrefe*, 1971.

[4] I. Papousek and G. Schulter. Quantitative assessment of five behavioural laterality measures: Distributions of scores and intercorrelations among right-handers. *LATERALITY-HOVE-*, VOL 4:NUMBER 4, pages 345–362, 1999.

[5] C. Breitwieser, A. Kreilinger, C. Neuper, and G.R. Müller-Putz. The TOBI Hybrid BCI – The data acquisition module. TOBI Workshop 2010 Graz, 2010.

[6] H. Ramoser, J. Müller-Gerking, and G. Pfurtscheller. Optimal spatial filtering of single trial EEG during imagined hand movement. *IEEE Transactions on Rehabilitation Engineering*, 8:441–446, 2000.

[7] G. R. Müller-Putz, R. Scherer, C. Brunner, R. Leeb, and G. Pfurtscheller. Better than random? A closer look on BCI results. *International Journal of Bioelectromagnetism*, 10:52–55, 2008.

Neurofeedback of Fronto-Parietal Gamma-Oscillations

M. Grosse-Wentrup[1]

[1]Max Planck Institute for Intelligent Systems, Spemannstr. 38, 72076 Tübingen, Germany

moritzgw@ieee.org

Abstract

In recent work, we have provided evidence that fronto-parietal γ-range oscillations are a cause of within-subject performance variations in brain-computer interfaces (BCIs) based on motor-imagery. Here, we explore the feasibility of using neurofeedback of fronto-parietal γ-power to induce a mental state that is beneficial for BCI-performance. We provide empirical evidence based on two healthy subjects that intentional attenuation of fronto-parietal γ-power results in an enhanced resting-state sensorimotor-rhythm (SMR). As a large resting-state amplitude of the SMR has been shown to correlate with good BCI-performance, our approach may provide a means to reduce performance variations in BCIs.

1 Introduction

Although research on brain-computer interfaces (BCIs) has seen remarkable progress in recent years, a substantial percentage of subjects remains incapable of utilizing a BCI [1]. Furthermore, subjects often display a large variation in performance over the course of an experimental session [2]. These factors limit the utility of BCI systems, and hinder a successful commercialization of this technology. Understanding and eliminating across- as well as within-subject performance variations arguably constitutes one of the most relevant problems in research on BCIs.

Recent studies have provided first important insights into the neuro-physiological causes of performance variations in motor-imagery BCIs. In particular, empirical evidence has been presented that the amplitude of the sensorimotor-rhythm (SMR) at rest is a good predictor of subsequent BCI-performance [3]. This suggests that in order to perform well, subjects first need to generate a strong SMR, i.e., a high amplitude of electromagnetic oscillations over sensorimotor-areas in the μ- (8–14 Hz) and β-range (20–30 Hz). Localized attenuation of the SMR by means of motor-imagery may then be used to convey a certain intention. Concurrently, our group has provided evidence that suggests a role of γ-range oscillations (≥ 40 Hz) in determining subject-specific levels of BCI-control. In particular, we have presented empirical evidence for an inhibitory modulation of the SMR by γ-range oscillations originating in frontal- and parietal areas [4, 5]. This effect may have a large impact on the design of future BCI-systems, as we found group-average classification accuracies in a two-class BCI to vary by up to 22.2 % depending on the state of fronto-parietal γ-power [2].

In this work, the hypothesis is tested that teaching subjects to attenuate fronto-parietal γ-power results in an enhanced SMR - a mental state that is likely to result in good BCI-performance. Online beamforming was employed to train three healthy subjects in intentional modulation of fronto-parietal γ-power. Two subjects acquired significant control. The third subject had to be discarded due to muscular artifacts. In agreement with the initial hypothesis, intentional attenuation of fronto-parietal γ-power resulted in a significant enhancement of the SMR. None of the subjects reported the use of motor-imagery for modulating fronto-parietal γ-power. These results establish that it is possible to learn how to intentionally control fronto-parietal γ-power, and that this new skill may be used to induce a state of mind that is beneficial for BCI-performance.

2 Methods

2.1 Neurofeedback Paradigm & Experimental Data

Each session of the feedback paradigm consisted of one resting-state baseline- and three training blocks. In the baseline block, the subject was instructed to relax for five minutes with eyes open while watching a grey fixation cross on a screen at a distance of approximately 1.5 m. This resting-state data was then used to learn a beamformer for estimating fronto-parietal γ-power and to compute the baseline mean and standard deviation of fronto-parietal γ-power (as described below in the section on data processing). In each of the subsequent training blocks, the subject received real-time feedback on the state of fronto-parietal γ-oscillations by means of a white ball displayed on the screen. Specifically, log-bandpower of fronto-parietal γ-oscillations was mapped to the vertical position of the ball on the screen. Here, the center of the screen corresponded to the mean bandpower observed during the baseline, and the upper and lower borders of the screen corresponded to $\pm$ 2 standard deviations. The vertical position of the ball was updated every 40 ms, while its horizontal position was fixed to the center of the screen.

Each training block consisted of twenty trials in pseudo-randomized order, in which the subject was instructed to try moving the ball to either the upper or lower border of the screen (subsequently termed conditions "Up" and "Down"). For this purpose, two grey rectangles were placed centrally on the upper and lower border of the screen, which changed their color to yellow in order to indicate the current target. The current target turned green whenever the subject managed to position the ball over it. The end of a trial was indicated by changing the color of both rectangles back to grey and hiding the ball. Each trial lasted for 60 s and was preceded by a baseline, randomly varying in length between 3.5 and 4.5 s.

During each session, a 121-channel EEG was recorded at 500 Hz using a QuickAmp amplifier with built-in common average reference BrainProducts GmbH, Germany). Electrodes were placed according to the extended 10-20 system. Three healthy subjects participated in this study (S1, S2, and S3). The first subject performed three training sessions on different days, corresponding to a total training time of three hours. This subject was a member of the BCI-lab with experience in motor-imagery. For this subject, only the results of the last training session are reported. The remaining two subjects performed one training session on a single day, with the third subject only completing two of the three training blocks. The second and third subject were naive to BCIs.

2.2 Data Processing

To learn a beamformer for estimating fronto-parietal γ-power, the resting-state baseline data was first temporally filtered between 55 and 85 Hz using a third order Butterworth filter. The spatial covariance matrix of this data was then used in conjunction with the topography shown in Figure 1 a) to compute a linearly constrained minimum variance (LCMV) beamformer [6]. The topography and frequency band chosen here correspond to the fronto-parietal γ-oscillations that were found to negatively correlate with motor-imagery performance in [2,5]. This beamformer was then used to spatially filter the resting-state baseline data. Subsequently, log-bandpower in the 55–85 Hz range was computed by a FFT in conjunction with a Hanning window, using a sliding window of 5 s length in steps of 40 ms. The resulting mean and standard deviation were used to calibrate the feedback procedure. During the actual training sessions, the same beamformer and sliding-window procedure were employed for estimating log-bandpower of fronto-parietal γ-oscillations and providing feedback in real-time. Stimulus-presentation and real-time data-processing was performed with the BCI2000-framework [7] and its extension BCPy2000 (http://bci2000.org/downloads/BCPy2000/).

3 Results

To assess the capability of each of the three subjects to intentionally control fronto-parietal γ-power, we divided the difference of mean ball position across conditions "Up" and "Down" by the

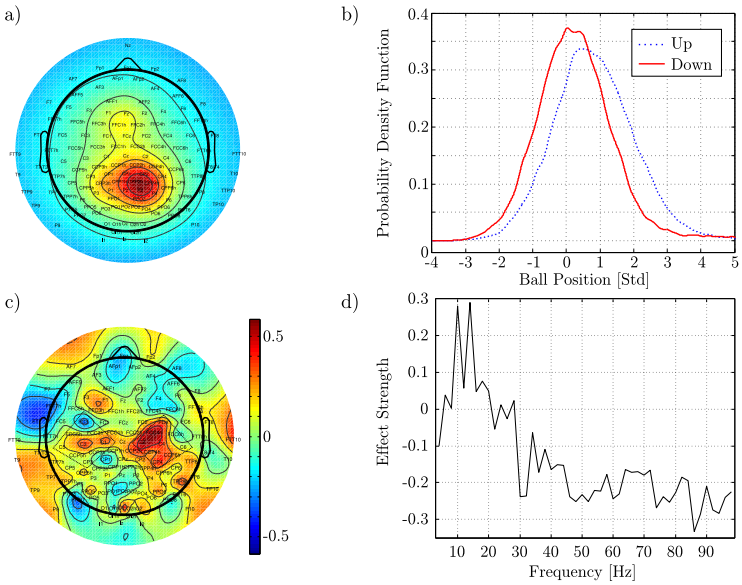


Figure 1: a) Topography used for beamforming. b) Estimated probability densities [8] of ball position (in standard deviations relative to resting-state baseline) for conditions "Up" and "Down" (group average). c) Topography of effect size d_{SMR} between 8–14 Hz (group average). d) Effect size d_{SMR} across spectral bands, averaged across electrodes over sensorimotor areas (dashed lines in c) (group average).

mean of the standard deviation of each condition. This resulted in effect strengths $d_\gamma^{S1} = 0.7273$, $d_\gamma^{S2} = 0.3419$, and $d_\gamma^{S3} = 0.3723$. Based on random permutation tests with 10.000 iterations, these effect strengths were found to be sufficient for rejecting the null-hypothesis of zero effect strength with $p_\gamma^{S1} < 1e^{-4}$ ($N = 60$), $p_\gamma^{S2} = 0.0005$ ($N = 60$), and $p_\gamma^{S3} = 0.0078$ ($N = 40$).

As γ-range oscillations may be caused by muscular artifacts, it is crucial to investigate whether observed changes in γ-power may have been confounded by systematic changes in muscle tone. While non-cortical components of the EEG can not be completely eliminated, their influence may be attenuated by artifact correction procedures, e.g., based on Independent Component Analysis (ICA). Here, the SOBI algorithm was employed to decompose the recorded EEG of the training blocks of each subject into independent components (ICs) [9, 10]. IC topographies and spectra were then manually inspected, and questionable ICs were discarded. The remaining ICs were reprojected to the scalp electrodes. To determine whether an observed modulation of fronto-parietal γ-power may have been caused by artifactual components, effect sizes before and after artifact correction were compared. Recomputation of effect sizes after artifact correction resulted in $d_\gamma^{S1} = 0.7578$, $d_\gamma^{S2} = 0.0249$, and $d_\gamma^{S3} = 0.3051$, i.e., a slight increase for the first subject, a strong decrease for the second subject, and a mild decrease for the third subject. Effect sizes S1 and S3 remained large enough to reject the null-hypothesis of zero effect strength with $p_\gamma^{S1} < 1e^{-4}$ ($N = 60$) and $p_\gamma^{S3} = 0.0241$ ($N = 40$). Subject S2 did not show a significant effect anymore ($p_\gamma^{S2} = 0.4351$, $N = 60$). This indicates that S2 employed muscle activity for cursor control. The data of subject S2 was hence discarded.

To test the hypothesis that intentional control of γ-power modulates the SMR, the raw EEG of each training session was spatially filtered using a Laplacian setup [11]. Then, for each trial and electrode log-bandpower was computed in 2 Hz frequency bins ranging from two to 98 Hz. Group-average effect size was computed across 18 electrodes covering sensorimotor-areas in frequency bins from 8–14 Hz. These electrodes (enclosed in dashed lines in Figure 1 c)) and frequencies were selected a-priori. This resulted in an effect size of $d_{\mathrm{SMR}} = 0.2088$, which is sufficient to reject the null-hypothesis of zero effect size with $p = 0.0069$ ($N = 100$). To illustrate the modulation of the SMR by γ-control, Figure 1 c) displays the topography of effect size between 12–14 Hz,

with positive values corresponding to increased μ-power due to an attenuation of fronto-parietal γ-power. Note that μ-power is enhanced primarily over right sensorimotor cortex. Figure 1 d) shows the effect strength averaged over sensorimotor areas across spectral bands. Only the μ-range displays an increase in bandpower due to intentional attenuation of fronto-parietal γ-power.

4 Discussion & Conclusions

The results presented in this work demonstrate that it is possible to learn how to control the power of fronto-parietal γ-oscillations, and that this new skill can be used to modulate the magnitude of the SMR in a manner beneficial for BCI-performance. Importantly, none of the subjects reported the use of motor-imagery. As such, the effect reported here is likely due to a modulation of sensorimotor cortex by fronto-parietal areas independently of motor-imagery. It remains to be seen which percentage of subjects may benefit from this effect, and whether similar results as reported here can be achieved with locked-in patients.

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(2):145–147, 2003.

[2] M. Grosse-Wentrup. Fronto-parietal gamma-oscillations are a cause of performance variation in brain-computer interfacing. In *Proceedings of the 5th International IEEE EMBS Conference on Neural Engineering (NE 2011)*. IEEE, 2011.

[3] B. Blankertz, C. Sannelli, S. Halder, E. M. Hammer, A. Kübler, K. R. Müller, G. Curio, and T. Dickhaus. Neurophysiological predictor of SMR-based BCI performance. *NeuroImage*, 51(4):1303–1309, 2010.

[4] M. Grosse-Wentrup, B. Schölkopf, and J. Hill. Causal influence of gamma rhythms on the sensorimotor rhythm. *NeuroImage*, 56(2):837–842, 2011.

[5] M. Grosse-Wentrup and B. Schölkopf. Gamma-power in a fronto-parietal network predicts motor-imagery performance. (under review).

[6] B. D. van Veen, W. van Drongelen, M. Yuchtman, and A. Suzuki. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Transactions on Biomedical Engineering*, 44:867–880, 1997.

[7] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw. BCI 2000: A general-purpose brain-computer interface (BCI) system. *IEEE Transactions on Biomedical Engineering*, 51(6):1034–1043, 2004.

[8] Z. I. Botev, J. F. Grotowski, and D. P. Kroese. Kernel density estimation via diffusion. *Annals of Statistics*, 38(5):2916–2957, 2010.

[9] A. Belouchrani, K. Abed-Meraim, J. F. Cardoso, and E. Moulines. A blind source separation technique using second-order statistics. *IEEE Transaction on Signal Processing*, 45(2):434–444, 1997.

[10] A. Delorme and S. Makeig. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1):9–21, 2004.

[11] D. J. McFarland, L. M. McCane, S. V. David, and J. R. Wolpaw. Spatial filter selection for EEG-based communication. *Electroencephalography and Clinical Neurophysiology*, 103:386–394, 1997.

Generating Artificial EEG Signals to Reduce BCI Calibration Time

F. Lotte[1]

[1]INRIA Bordeaux Sud-Ouest, Talence, France

fabien.lotte@inria.fr

Abstract

One of the major limitations of Brain-Computer Interfaces (BCI) is their long calibration time. This is due to the need to collect numerous training EEG trials for the machine learning algorithm used in their design. In this paper we propose a new approach to reduce this calibration time. This approach consists in generating artificial EEG trials from the few EEG trials initially available, in order to augment the training set size in a relevant way. The approach followed is simple and computationally efficient. Moreover, our offline evaluations suggested that it can lead to significant increases in classification accuracy when compared with existing approaches, especially when the number of training trials available is small. As such, it can indeed be used to reduce calibration time.

1 Introduction

In recent years, research efforts in Brain-Computer Interfaces (BCI) have led to the development of numerous BCI prototypes which have highlighted many promising applications, both in medical and non-medical domains. However, BCI still suffer from several limitations which make these prototypes usually far from being usable in practical applications [1]. One of these limitations is that many examples of the user's EEG signals must be recorded in order to calibrate the BCI, which is both inconvenient and time consuming. Currently, the best online BCI systems require a calibration time of about 20 minutes [2], hich is still far too long. As a comparison, nobody would indeed use a computer mouse if it required a 20-minute-long calibration before each use.

A few recent works have proposed approaches to reduce [3–5] or suppress [6–8] this calibration time. However, these approaches require either data from numerous other users whose EEG signals have been previously recorded, or a large data set of past trials from the same user. Naturally, such data cannot be always available and we would ideally like the BCI to be efficient with very little training data even for a first-time BCI user. Therefore, there is a need for a method that can reduce the BCI calibration time without requiring numerous EEG data previously recorded. In this paper we propose such a method. It is a simple and efficient approach which consists in generating numerous artificial EEG signals from a few signals recorded from the user. Such new trials can then be used to augment the available training set.

2 Methods

BCI calibration times are long due to the need to collect numerous training EEG trials for the subsequent machine learning algorithms used in their design. This problematic need for large amounts of training data is not unique to the BCI field. It can also be found in other fields in which machine learning is involved, although the problem might be more severe for BCI. In these fields, one proposed approach dealing with this issue was to generate numerous artificial training data from the few actual training data available, and use it to augment the training

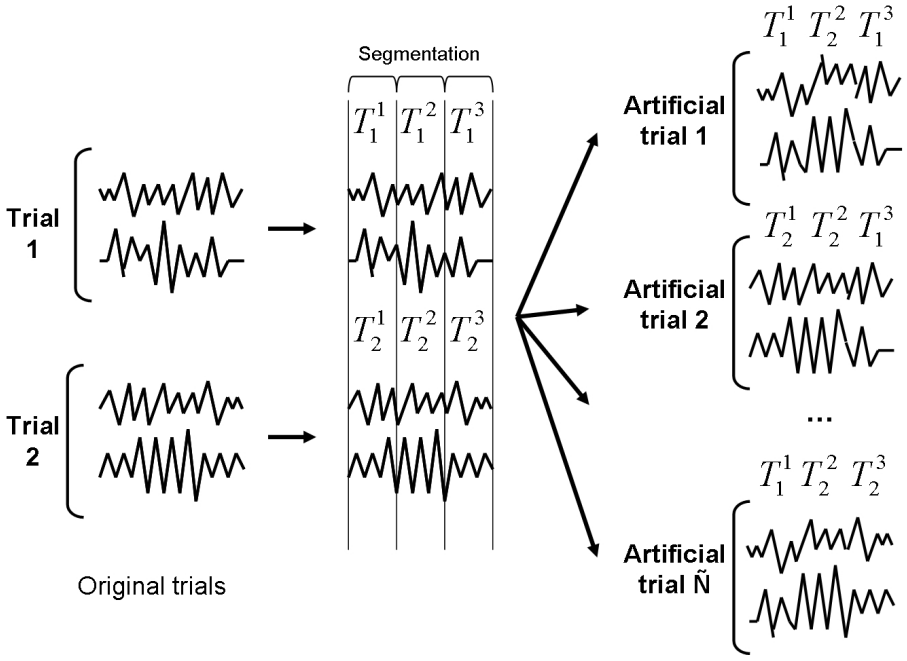


Figure 1: Principle of artificial EEG data generation for BCI design.

set. This has been shown to lead to increased classification accuracies in fields such as speech or hand-writing recognition [9, 10]. In this paper, we propose to follow a similar approach in order to generate artificial EEG data based on the method used in [9] for speech recognition. With this approach, we first divide each training EEG trial into several segments, and then generate new artificial trials as a concatenation of segments coming from different and randomly selected training trials. More formally, let us denote as $\Omega = \{T_i\}, i \in [1, N]$ the set of N EEG trials that are available for training. $T_i \in \Re^{C \times S}$, with S the number of samples in a trial, and C the number of channels. The first step consists in dividing the signals (from each channel) of each training trial T_i into K consecutive and non-overlapping segments $T_i^k \in \Re^{C \times S/K}, k \in [1, K]$. Then, from these segments, we can generate a new artificial trial $\tilde{T}_j$ as $\tilde{T}_j = [T^1_{R_1} T^2_{R_2} \dots T^K_{R_K}]$, where $[AB]$ denote the concatenation of the samples from segment A and B, and R_k is a randomly selected integer from $[1, N]$. The whole process is schematized in Figure 1. This simple approach enables us to generate a large number of new trials, different from the original ones, but still relevant and likely to be similar to future trials, as they were made from parts of real trials and have the same temporal structure. By adding these artificial trials to the original training trials, we can fill the feature space in a relevant way and hence ease the training of the subsequent machine learning algorithms. We thus expect this approach to give good classification performances even when only few EEG trials are initially available.

3 Evaluation and Results

We evaluated this approach offline on data set 2a from BCI competition IV, provided by the Graz group [11]. This data set comprises EEG signals from 9 subjects who performed left hand, right hand, foot and tongue Motor Imagery (MI). The EEG signals were recorded using 22 EEG channels. For the purpose of this study, only EEG signals corresponding to left and right hand MI were used. EEG signals were band-pass filtered in the 8–30 Hz frequency band using a 5th order Butterworth filter. Features were extracted using the Common Spatial Patterns (CSP) [12] algorithm (we used the filters corresponding to the 3 largest and smallest eigenvalues) from the 2-second time window starting 0.5 s after the cue. Features were classified using Linear Discriminant

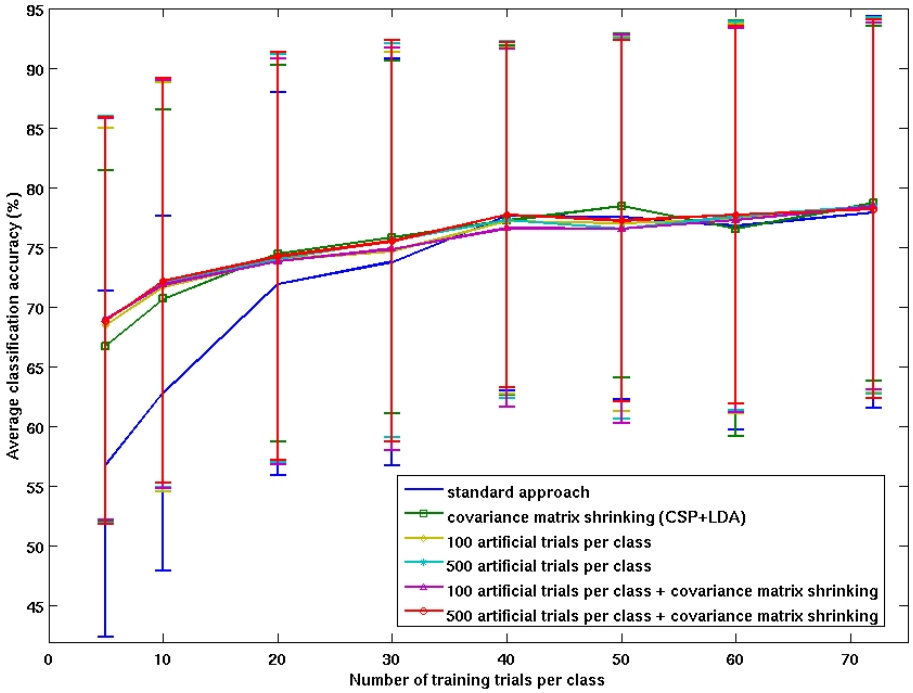


Figure 2: Classification accuracy (%) as a function of the number of training trials, for different approaches (standard approach, covariance matrix shrinking, artifical data generation).

Analysis (LDA).

We evaluated our approach on the test sets available (each comprising 72 trials per class), while using only the N first training trials from each class ($N \in [5, 10, 20, 30, 40, 50, 60, 72]$) to train CSP and LDA. In this context, we compared different approaches:

- The standard approach which directly trains CSP and LDA on the initially available training trials. This is what is done in most current BCI.
- Our approach based on Artificial Data Generation (ADG) as described above. We generated either 100 or 500 artificial EEG trials per class (using K = 10 segments) and added them to the original training set.
- An approach which uses automatic Covariance Matrix Shrinking (CMS) for CSP and LDA, based on Ledoit and Wolf's method, as we did in [4]. This approach is specifically designed to obtain better estimate of covariance matrices with small training sets.
- A last approach which uses both ADG and automatic CMS for CSP and LDA.

As ADG involves random number generation, we repeated the ADG procedure 10 times, and averaged the classification accuracies hence obtained for each repetition. The average classification accuracies obtained for each method and training set size are reported on Figure 2. Results show that both ADG and CMS are more efficient than the standard approach, particularly when the number of training trials available is small (< 40). Globally, this difference is statistically significant ($p << 0.0001$ with a paired t-test). Using ADG, the accuracy obtained with only 10 trials is the same as that obtained with the standard approach when using 20 trials. Hence, as compared to the standard approach, ADG can effectively reduce the calibration time by 2. Using both CMS and ADG together leads to a slight improvement in accuracy as compared to using ADG alone, although this improvement is not statistically significant. When comparing ADG with CMS, it appears that when the number of training trials is really small (< 20), ADG leads to

higher classification accuracies. For a larger number of training trials, ADG and CMS appear to be rather similar. Overall, only ADG combined with CMS (with 500 artificial trials) leads to an accuracy that is significantly higher than that obtained by CMS alone ($p < 0.05$). Overall, these results suggest that artificial EEG trials generation is a simple and efficient approach to calibrate BCI systems with few training trials, thus effectively reducing the calibration time.

4 Conclusion

In this paper we have proposed a new approach to reduce the calibration time of BCI that are based on machine learning. This approach consists in generating artificial EEG trials from the few EEG trials initially available in order to augment the training set size in a relevant way. This is achieved by first dividing each available training EEG trial into several segments, and then by generating new artificial trials as a concatenation of segments coming from different and randomly selected training trials. This approach is simple, easy to implement and computationally efficient. Moreover, our offline evaluations suggested that it can indeed lead to significant increases in classification accuracy as compared to existing approaches, especially when the number of training trials available is small. As such, it can indeed be used to reduce the calibration time. Moreover, this approach is independant of the machine learning algorithms used subsequently and could therefore become a new tool in the repertoire of BCI designers. Future work will be dedicated to study more advanced methods to perform artificial trial generation, possibly based on data distortion and analogy as in [10], or exploiting a-priori knowledge on the mental tasks used to drive the BCI.

References

[1] B. Blankertz, M. Tangermann, C. Vidaurre, S. Fazli, C. Sannelli, S. Haufe, C. Maeder, L. Ramsey, I. Sturm, G. Curio, and K. R. Müller. The Berlin Brain-Computer Interface: non-medical uses of BCI technology. *Frontiers in Neuroprosthetics*, 5, 2010.

[2] B. Blankertz, G. Dornhege, M. Krauledat, G. Curio, and K. R. Müller. The non-invasive Berlin Brain-Computer Interface: fast acquisition of effective performance in untrained subjects. *NeuroImage*, 37(2):539–550, 2007.

[3] H. Lu, H. L. Eng, C. Guan, K. N. Plataniotis, and A. N. Venetsanopoulos. Regularized common spatial pattern with aggregation for EEG classification in small-sample setting. *IEEE Trans. Biomed. Eng.*, 57(12):2936–2946, 2010.

[4] F. Lotte and C. Guan. Learning from other subjects helps reducing brain-computer interface calibration time. In *ICASSP'2010*, pages 614–617, 2010.

[5] S. Lu, C. Guan, and H. Zhang. Unsupervised brain computer interface based on inter-subject information and online adaptation. *IEEE Trans. Neural Syst. and Rehab.*, 17(2):135–145, 2009.

[6] S. Fazli, C. Grozea, M. Danóczy, B. Blankertz, F. Popescu, and K. R. Müller. Subject independent EEG-based BCI decoding. In *NIPS 22*, pages 513–521, 2009.

[7] F. Lotte, C. Guan, and K. K. Ang. Comparison of designs towards a subject-independent brain-computer interface based on motor imagery. In *EMBC'2009*, pages 4543 – 4546, 2009.

[8] M. Krauledat, M. Tangermann, B. Blankertz, and K. R. Müller. Towards zero training for brain-computer interfacing. *PLoS ONE*, 3(8):e2967, 2008.

[9] R. Chakraborty and U. Garain. Role of synthetically generated samples on speech recognition in a resource-scarce language. *ICPR*, pages 1618–1621, 2010.

[10] H. Mouchère, S. Bayoudh, E. Anquetil, and L. Miclet. Synthetic on-line handwriting generation by distortions and analogy. In *IGS2007*, pages 10–13, 2007.

[11] M. Naeem, C. Brunner, R. Leeb, B. Graimann, and G. Pfurtscheller. Seperability of four-class motor imagery data using independent components analysis. *J. Neur. Eng.*, 3:208–216, 2006.

[12] 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.

Comparison of Feature Stages in a Multi-Classifier BCI

I. J. Cester[1]*, A. Soria-Frisch[2]

[1]Current address: Neurochemistry Group, Physiological Science I, University of Barcelona, Spain
[2]Starlab Barcelona S.L., Spain

ivan.cester@ub.edu | aureli.soria-frisch@starlab.es

Abstract

The present communication is focused on the comparison of two different approaches for the realization of the feature extraction stage in a Brain-Computer Interface (BCI) system. For this purpose we compare the usage of Common Spatial Patterns (CSP) and of Wavelet Analysis (WA). Due to the large dimensionality of the wavelet feature space, we further compare the performance of two selection procedures: Analysis of Variance (ANOVA) and Genetic Algorithms (GA). The inclusion of such data-driven feature selection stages avoids the application of a priori hypothesis on the most relevant frequency bands. An extensive performance analysis is given for supporting the outperformance of the WA approach over CSP, and of the GA one over ANOVA.

1 Introduction

A generic block diagram of a Brain-Computer Interface (BCI) system includes stages for data acquisition, signal processing, feature extraction and feature translation [1]. Often the feature extraction stage presents an additional stage for feature selection. Their combination has been denoted herein as feature stage. If a system makes use of a classifier as feature translator, the resulting system can be analyzed from a pattern recognition perspective. Pattern recognition systems for BCI have been extensively reviewed in [2].

The work described herein is focused on motor imagery, where control commands are generated by classifying the electroencephalogram (EEG) streams into diverse classes. We have focused our research work in the classification of the data set IIIa generated within the BCI Competition III. It includes 60 channels of EEG signal streams corresponding to cued motor imagery (multi-class) with 4 classes, i.e. movement of left hand, right hand, foot, and tongue, that are delivered for classification. The data correspond to three different subjects labeled as k6b, k3b and l1b. The former has 90 trials per class and the other two 60; half of them are used for training and the other half for testing as defined in the BCI competition[1].

The paper is focused on the comparison of two different ways of undertaking the feature extraction stage in a pattern recognition system. The system that is used as reference includes a multi-classifier approach within a bagging structure. It was presented by G. Xiaorong et al.[2] and obtainined the second place in the classification of data set IIIa at the BCI Competition III . We chose it as a starting point because the winning system was not described with enough detail. Moreover it is worth mentioning that multi-classifier approaches are presented as a promising technology for BCI in [2]. Xiaorong and colleagues' s multi-classifier system presents a feature extraction stage based on the One-Versus-Rest (OVR) methodology [3]. OVR is a multi-class

*Ivan Cester realized the research works described herein when working in Starlab Barcelona S.L. They have been realized partially funded within AsTeRICS, a STREP collaborative project supported under the EU 7th Framework Program (Grant agreement number: FP7-ICT-2009-4-247730).

[1]Data set description in http://www.bbci.de/competition/iii/desc_IIIa.pdf.

[2]System description in http://www.bbci.de/competition/iii/results/graz_IIIa/GaoXiaorong_desc.pdf.

generalization of the Common Spatial Patterns (CSP), which are widely used within the BCI community [4].

The paper compares the results of Xiaorong's group based on CSPs with these attained by realizing the feature extraction stage through the utilization of wavelets. Wavelets have been already used in BCI research [5]. Since this type of joint time-frequency analysis presents a very large dimensionality, we additionally consider a feature selection stage. Hence we perform the comparison of two different alternatives for feature selection: Analysis of Variance (ANOVA) [6] and Genetic Algorithms (GA) [7]. This constitutes the novel aspect of the presented methodology.

The motivation for using these data-driven approaches is exploring whether the avoidance of an a priori selection of frequency bands can improve the motor imagery system performance. The a priori selection of bands and electrodes underlay approaches like Xiaorong et al.'s one and similar systems. It is based on the fact that motor imagery is known to trigger event related desynchronization (ERD) of the mu (10–12 Hz) and beta (20–25 Hz) rhythm in related motor cortex areas. This physiological information drives the feature extraction [8]. It is worth mentioning that our purpose when using a feature selection approach is not to manage re-adjusting the selected features online during BCI operation to tackle inter-trial variability. This will not be possible by using GA, which present a large training time. Moreover both approaches are thought to be used off-line just to select the wavelet coefficients, i.e. frequency bands and electrode locations, that offer a better inter-class discrimination capability without making an a priori selection.

We mainly attain the improvement of classification performance by the comparison of the two feature selection approaches. Nevertheless ANOVA and GA present several methodological differences. Therefore one possible conclusion of the paper is therefore the suitability of using a one-stage approach, i.e. ANOVA, vs. this of using a multi-stage one, i.e. GA when dealing with the high-dimensional wavelet spaces.

The paper organization can be summarized as follows. Section 2 presents the compared feature extraction stages Section 3 the results obtained and finally some conclusions are drawn in Section 4.

2 Feature Extraction Approaches under Comparison

The reference system as presented by Xiaorong and colleagues uses a laplacian filter and OVR, which has been implemented herein as described in [3]. In the implementation we first optimize the performance of the individual classifiers applying a 10 cross-fold validation in order to select the best parameters for each classifier type. The output dimension of the OVR methodology, characterized by the parameter **m**, is also optimized with respect to the classification rate. The classifier of the reference system is based on a multi-classifier ensemble formed by a support vector machine, a linear discriminant, and a K-nearest neighbor classifiers. A bagging procedure is additionally conducted over the multi-classifier scheme. We maintain the same classification methodology for all compared feature stages.

The work we present in this paper avoids the general physiological priors for feature extraction mentioned in the former section. We do not look for particular bands and/or electrode locations, we first use an extensive time-frequency analysis of the 60 channels raw EEG signals. This has been realized by applying a Morlet wavelet, recommended for the analysis of EEG signals [9].

The pre-processing and joint time-frequency feature extraction process applied to the data is given as follows: 1. Decimate the raw signal from 250 to 62.5 Hz, using a Chebyshev Type I filter with cutoff frequency 25 Hz. 2. Trials with signals over $\pm 100\,\mu$V are considered artifacts and rejected. 3. Signal of each channels is referenced to the mean signal of all channels, aka as common average reference spatial filter. 4. Wavelet analysis: Time frequency information from the signals is extracted through a complex Morlet wavelet with bandwidth parameter $Fb = 1$ and a wavelet center frequency $Fc = 1.5$. A total of 30 coefficients corresponding to 1–30 Hz bandwidth are obtained for each of the 60 channels. That is a total of 1800 coefficients for each time sample.

Since the number of the extracted wavelet coefficients is rather large, a feature selection stage should be better applied. Hence we compare the performance of the analysis of variance (ANOVA) [6] and of Genetic Algorithms (GA) [7] for the implementation of this stage.

2.1 Feature Selection through ANOVA and OVR

A one-way ANOVA is applied to each coefficient for each frequency and time samples through the trials of the training set. This analysis determines the significance of each coefficient to explain the variability between classes. As expected, the most significant frequencies are found around mu (10–12 Hz) and beta (20–25 Hz) rhythm for most of the channels, but with slight differences between subjects and scalp locations.

The classification is performed with the last 3 seconds of the trial, when the imagery task is performed by the subject. The 250 coefficients with lower mean of significance p over time in the whole trial have been selected as the most reactive; this corresponds to mean p values under 0.00008, 0.21 and 0.20 for subjects k3b, k6b and l1b respectively.

In this first approach, after the ANOVA pruning of the feature set, the OVR algorithm projects the selected features into a subspace where the variability among classes is incased and the intra-class variability minimized [3]. The tuning of the optimal number of coefficients selected with the ANOVA has been attained by setting a fix value for the dimensions of the class spaces where OVR matrices was projecting (parameter **m**).

2.2 Feature Selection through GA and OVR

This second feature selection approach applies a genetic algorithm [7] as an alternative to the ANOVA. The used fitness function seeks for a subset of coefficients that minimizes the error in the classification of a validation set. Hence the training set has been split in 2 subsets; 70 % of the trials have been used to train the LDA and the rest to validate its generalization performance.

The employed GA uses a population of 200 individuals represented through binary vectors. It is parameterized with 0.9 cross-over and 0.1 mutation probabilities and elitism of 4 individuals per generation. This parameters are standard in GA research. The OVR projection is optimized and applied again after the GA one.

2.3 Time Domain Fusion

It is known that it is important to find a trade-off between the classification performance and the rate of commands per second that can be obtained with a BCI system. In this context, an averaging fusion among samples can be used to obtain the prevalent output in a time interval. We apply this last step when giving the experimental results.

3 Results

In the following paragraphs we compare the results obtained with all the approaches. Table 1 summarizes the results obtained with the 3 approaches when classifying sample per sample. Feature selection of the wavelet coefficients through GA gave very good results, increasing notably the performance of the system for subject k3b compared to both the OVR and ANOVA approaches.

Subject	OVR			ANOVA+OVR			GA+OVR		
	m	κ train	κ test	m	κ train	κ test	m	κ train	κ test
k3b	[4,4,3,4]	0.65	0.34	[4,4,3,4]	0.97	0.53	[16,11,16,16]	0.97	0.63
k6b	[2,2,2,2]	0.48	0.11	[3,2,3,5]	0.99	0.20	[15,15,15,15]	0.97	0.23
l1b	[4,4,3,4]	0.44	0.12	[4,5,4,4]	0.98	0.38	[10,12,16,14]	0.97	0.42

Table 1: Performance for each subject after optimization of OVR parameter **m**: each component correspond to the output dimensionality of each of the 4 class spaces where the trials are projected.

The results by applying a fusion in the time domain show that the performance in the classification is enhanced as the decision rate decreases (see Figure 1). The test kappa index increases about 20 % for all subjects when making a single decision per trial e.i. 1 decision each 3 seconds.

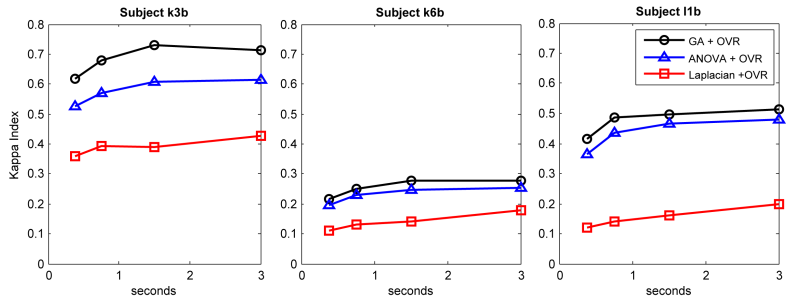


Figure 1: Test κ index vs. decision rate for the 3 subjects with the 3 approaches. Laplacian + OVR refers to the reference methodology (see Section 2).

4 Conclusion

The classification performance of three different feature sets have been compared, the reference system based on OVR with two wavelet-based approaches. These last ones outperforms OVR. We additionally compare GA and ANOVA for the feature selection stage, being GA the best option for this particular dataset in terms of classification performance.

The generalization performance can be improved. In this context it is worth mentioning two facts. First these are results obtained when classifying the epochs sample by sample. Second there are two classes that are specially difficult to classify. Moreover it is worth mentioning that the evaluation procedure herein and the one used in the competition seem to differ, what would explain the differences in the given performances. The performance of the OVR system is given herein for the sake of showing the relative improvement and not as an absolute value.

References

[1] A. Kachenoura, L. Albera, L. Senhadji, and P. Comon. ICA: a potential tool for BCI systems. *IEEE Signal Processing Magazine*, 25(1):57–68, 2008.

[2] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi. A review of classification algorithms for EEG-based brain-computer interfaces. *J. Neural Engineering*, 4(2), 2007.

[3] W. Wu, X. Gao, and S. Gao. One-Versus-the-Rest (OVR) algorithm: An extension of common spatial patterns (CSP) algorithm to multi-class case. *Proc. IEEE 27th Int. Conf. Engineering in Medicine and Biology Society.*, pages 2387–2390, 2005.

[4] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K. R. Müller. Optimizing spatial filters for robust EEG single-trial analysis. *IEEE Signal Proc. Mag.*, 25:581–607, 2008.

[5] L. Qin and B. He. A wavelet-based time-frequency analysis approach for classification of motor imagery for BCI applications. *J. Neural Engineering*, 2(4):65–72, 2005.

[6] C. P. Doncaster and A. J. H. Davey. *Analysis of Variance and Covariance. How to Choose and Construct Models for the Life Science.* Cambridge University Press, 2007.

[7] D. E. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning.* Addison-Wesley Pub. Co., Inc., Reading, Massachussets, 1989.

[8] G. Pfurtscheller, C. Brunner, A. Schlögl, 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.

[9] C. S. Herrmann, M. Grigutsch, and N. A. Busch. EEG oscillations and wavelet analysis. In *Event-related potentials: a methods handbook*, pages 229–259. MIT Press, 2005.

Assessment Framework of Functional Brain Networks During Covert Motor Performance After Stroke

F. De Vico Fallani[1,3], F. Pichiorri[1,2], C. Di Lanzo[1], F. Ceccarelli[1], I. Pisotta[1], F. Cincotti[1], M. Molinari[1], F. Babiloni[3], D. Mattia[1]

[1]Fondazione Santa Lucia, IRCCS, Rome, Italy
[2]bDipartimento Di Scienze Neurologiche, Università "Sapienza", Rome, Italy
[3]Dipartimento di Fisiologia Umana, Università "Sapienza", Rome, Italy

fabrizio.devicofallani@uniroma1.it

Abstract

In the present study, we propose a methodological approach to assess the functional brain network organization underlying the motor imagery in stroke patients. Functional brain connectivity was estimated from high-density EEG signals in a group of patients ($N = 21$) suffering from unilateral cortical infarctions during the grasping imagination with both the affected and unaffected hand. The use of a graph theoretical approach allowed the characterization of the functional EEG networks and revealed a clear altered connectivity structure in the affected hemisphere during the motor imagery of the affected hand. These findings indicate that analyses of connectivity may offer new insights into the pathophysiology underlying stroke-induced neurological symptoms. Such information may help in addressing therapies to enhance motor recovery in patients, also through BCI-assisted systems.

1 Introduction

Over the last decade, there has been a growing interest in the detection of functional organization of brain networks estimated from different electromagnetic and hemodynamic signals [1–3]. As other network systems in nature, the functional brain networks estimated from the actual brain-imaging technologies (MEG, fMRI, and EEG) can be analyzed by means of the graph theory [4, 5] where graphs are mathematical representations of networks which are constituted essentially by nodes and connections between nodes. The usefulness of the graph theoretical approach in neuroscience was firstly demonstrated on a set of anatomical brain networks [6] and it has now become a well established framework [7]. Here, we applied this methodology to assess the consequences of a unilateral stroke lesion on the brain functional network organization as estimated from the EEG signals collected during covert motor actions. As such, this assessment is promising in identify possible lesion-induced disruption of functional neural networks in relation to motor performance after stroke.

2 Methods

High-density EEG data were collected from 21 patients with first ever monolateral stroke, consecutively enrolled at the Fondazione Santa Lucia where they were admitted for neurorehabilitation (mean age 58 ± 16 y, chronic/subacute 6/15, left/right lesion 7/14, subcortical/cortical 5/16, European Stroke Scale ESS 71 ± 13). All patients gave a written informed consent and the study was approved by the local ethical committee. Scalp potentials were collected from 61 electrodes and amplified by a commercial EEG system (band-pass 0.1–70 Hz; frequency sample 200 Hz). The EEG signals were referenced to the average between electrodes on the two ears. Patients were comfortably seated in a dimly lit room with upper limbs resting on a cushion and were instructed by a visual cue to image (MI) hand grasping with unaffected (Uhand) and affected hand (Ahand). Each recording session comprised 2 runs consisting of 30 trials (15 ± 1 rest trials, 15 ± 1 motor trials, randomized). A temporal epoch of 4 seconds was considered in each task (Ahand, Uhand), i.e. the "GO" period (4 s) of the motor trials and the last 4 seconds of the rest trials. All those

EEG epochs contaminated by artifacts were excluded from further analysis. hen functional network was explored during the motor and rest trials by computing the imaginary coherence [8], i.e. a measure of synchronization in the frequency domain, between the time series of all the available EEG channels. Imaginary coherence values were estimated within four frequency bands: Theta (4–7 Hz), Alpha (8–13 Hz), Beta (14–29 Hz), and Gamma (30–40 Hz). A statistical contrast between the link distributions of the motor and rest trials was performed with a multiple Student's t-test (significance level = 0.05, False Discovery Rate corrected for multiple comparisons) for each task and frequency band [5]. Thus, we obtained a scalp network containing only the significan functional connections (i.e. 0 = no link, 1 = presence of significant link) between the EEG sensors characterizing the motor trials during the two tasks (Ahand, Uhand).

2.1 Network Indexes

The estimated functional connectivity patterns were characterized by means of network metrics derived from graph theory [5]. A graph consists of a set of vertices (or nodes) and a set of edges (or connections) indicating the presence of interactions between the vertices. The adjacency matrix A contains the information about the connectivity structure of the graph. When a link connects two nodes i and j, the corresponding entry of the adjacency matrix is $aij = 1$; otherwise $aij = 0$. In our graphs, nodes (or vertices) represent scalp electrodes and links (connecting edges) represent the statistically significant synchronization between EEG signals recorded from electrode pairs. The simplest attribute of a node is its connectivity degree, which is the total number of connections with other vertices. A high degree indicates that a specific scalp electrode has a central role for the communication spreading within the network, as its removal would reduce sensibly the overall connectivity. Two measures are frequently used to characterize the global structure of brain networks. They are the global and local efficiency [9]. The former measures the efficiency of the passage of information among the nodes, the latter indicates the tendency of the network to form highly connected clusters of vertices. The global-efficiency Eg of a graph is defined as the average of the reciprocal of the shortest distances dij between all node pairs:

$$Eg(A) = \frac{1}{N(N-1)} \sum_{i \neq j}^{N} 1/d(i,j) \tag{1}$$

where N is the number of vertices composing the graph. The local-efficiency El is the average of all the sub-graphs global-efficiencies. Global (Eg) and local efficiency (El) range from 0 to 1. A high Eg value indicates a high integration of the information between all the nodes. A high El value indicates a high tendency of the network to be segregated in separate clusters of highly connected nodes.

2.2 Network Hemisphere Separation

In the present study, we considered the specific role of the functional connectivity in the two separate hemispheres (affected AH and unaffected UH) of each patient's brain. In particular, we computed the above network indexes for the two sub-networks consisting in the nodes (electrodes) and links (synchronization) within each hemisphere (Figure 1). Notably, for the node degree, which gives a value for each single node, we rather considered the mean of the degrees within each hemisphere to obtain a global measure of the hemispheric degree. Finally all the values obtained from the 20 patients were used in a T-test to assess the topological difference between the two hemispheres (AH, UH) during the two tasks (Ahand, Uhand) in each frequency band.

3 Results

The statistical contrast between the network indexes from all the patients ($N = 20$) in the two hemispheres (AH, UH) revealed a different outcome with respect to the two tasks (Ahand, Uhand)

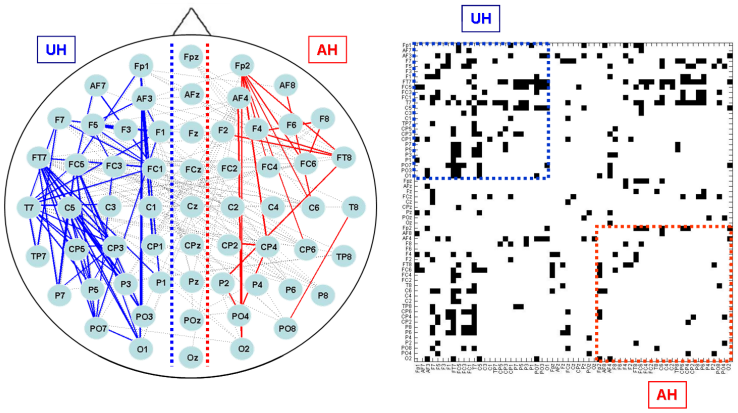


Figure 1: Scalp EEG network for a representative patient during the Ahand task in the Beta frequency band. On the left, the functional connectivity is illustrated over the scalp electrode montage. The links within the affected hemisphere (AH) are in red color, while those within the unaffected hemisphere (UH) are in blue color. On the right side, the adjacency matrix of the network is illustrated. The significant links are represented by black dots.

only in the Beta (14–29 Hz) and Gamma (30–40 Hz) band. In particular, we observed in the Beta band a significant ($p < 0.05$) inter-hemispheric difference for all the computed indexes in the Uhand task, while no significant differences were reported in the Ahand task (Figure 2). Notably, the motor imagery of the unaffected hand (Uhand) was characterized by a significantly higher connectivity in the contra-lateral unaffected hemisphere with respect to the affected one in terms of mean degree, global and local efficiency. In the Gamma band, we reported a significant ($p < 0.05$) inter-hemispheric difference for all the computed indexes during the Ahand task, while no significant differences were reported during the Uhand task. The motor imagery of the affected hand (Ahand) was characterized by a significantly lower connectivity in the contra-lateral affected hemisphere with respect to the unaffected one in terms of mean degree, global and local efficiency. No significant difference between the AH and UH hemiphere during the Ahand and Uhand task were found in the Theta and Alpha bands.

4 Discussion

The aim of the present work was to define a methodology able to estimate the functional organization of the brain in stroke patients. The strength and the usefulness of similar approaches based on functional connectivity analyses has also been recently stressed in stroke patients. To date, much of the neurobiological mechanisms leading to changes in cortical connectivity after stroke remain to be elucidated [10]. The obtained findings revealed a different pattern of connectivity in the two hemispheres (AH, UH) between the motor imagery of the affected (Ahand) and unaffected (Uhand) hand. This evidence would represent the basis for a better understanding of the functional neural reorganization in patients suffering from monolateral stroke and could help to design hypothesis-driven treatment strategies to promote recovery of motor function in patients through BCI-based assisted rehabilitations.

5 Acknowledgments

This work is supported by the European ICT Programme Project FP7-224631(TOBI). 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.

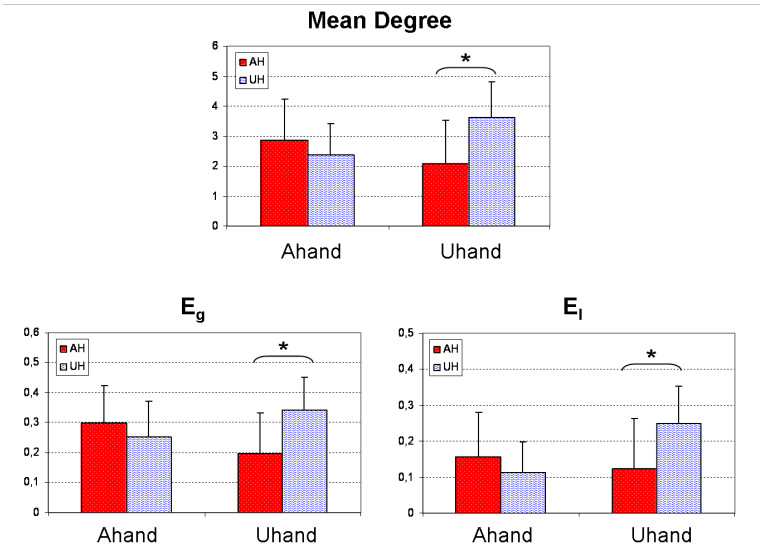


Figure 2: Grand-average values of the network indexes in the Beta frequency band. Red bars code the values for the affected hemisphere (AH), blue bars for the unaffected hemisphere (UH).Vertical bars denote standard deviations. The asterisks indicate a significant ($p < 0.05$) difference between the values of the two hemispheres.

References

[1] B. Horwitz. The elusive concept of brain connectivity. *Neuroimage*, 19(2 Pt 1):466–470, Jun 2003.

[2] O. David, D. Cosmelli, and K. J. Friston. Evaluation of different measures of functional connectivity using a neural mass model. *Neuroimage*, 21(2):659–673, Feb 2004.

[3] L. Lee, L. M. Harrison, and A. Mechelli. The functional brain connectivity workshop: report and commentary. *Network*, 14(2):R1–15, May 2003.

[4] C. J. Stam and J. C. Reijneveld. Graph theoretical analysis of complex networks in the brain. *Nonlinear Biomed Phys*, 1(1):3, 2007.

[5] F. De Vico Fallani, L. Astolfi, F. Cincotti, D. Mattia, A. Tocci, S. Salinari, M. G. Marciani, H. Witte, A. Colosimo, and F. Babiloni. Brain network analysis from high-resolution EEG recordings by the application of theoretical graph indexes. *IEEE Trans Neural Syst Rehabil Eng*, 16(5):442–452, Oct 2008.

[6] S. H. Strogatz. Exploring complex networks. *Nature*, 410(6825):268–276, Mar 2001.

[7] E. Bullmore and O. Sporns. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci*, 10(3):186–198, Mar 2009.

[8] G. Nolte, O. Bai, L. Wheaton, Z. Mari, S. Vorbach, and M. Hallett. Identifying true brain interaction from EEG data using the imaginary part of coherency. *Clin Neurophysiol*, 115(10):2292–2307, Oct 2004.

[9] V. Latora and M. Marchiori. Efficient behavior of small-world networks. *Phys Rev Lett*, 87(19):198701, Nov 2001.

[10] C. Grefkes and G. R. Fink. Reorganization of cerebral networks after stroke: new insights from neuroimaging with connectivity approaches. *Brain*, Mar 2011.

Initial Tests with Auditory BCI Paradigms in Patients with Disorders of Consciousness

C. A. Ruf[1], S. Halder[1], A. Furdea[1], T. Matuz[1], N. Birbaumer[1,2], B. Kotchoubey[1]

[1]Institute of Medical Psychology and Behavioral Neurobiology, University of Tübingen, Germany
[2]IRCCS, Ospedale San Camilo, Venezia-Lido, Italy

carolin.ruf@medizin.uni-tuebingen.de

Abstract

The diagnosis and assessment of patients with severe disorders of consciousness still present challenges for clinical neuroscience. Partially preserved cognitive abilities in some of these patients were found in several studies. Brain-computer interfaces (BCI) may help to overcome both the problems of cognitive assessment and communication in these patients. An experiment on a sample of seven vegetative state and minimally conscious state patients with two auditory paradigms (three-stimulus oddball and four-choice) provided promising offline classification accuracy. This was as high as 87.5 % in one patient in the four-choice paradigm. These results support the idea of developing BCI paradigms for patients with disorders of consciousness to avoid misdiagnosis and establish communication channels for patients with residual conscious awareness.

1 Introduction

Patients diagnosed with disorders of consciousness such as persistent vegetative state (PVS) and minimally conscious state (MCS) are in a non-responsive state, but several studies found residual awareness and partial intact cognitive processing [1–3]. Kübler and collegues suggested that a hierarchical approach based on brain-computer interface (BCI) could be used for investigation of cognitive processing in these patients [4]. Thereby, first, passive stimulation tasks with and without instruction were applied. Second, the ability of willful attentional selection is tested, e.g. by means of event-related potential (ERP)-based BCIs. These paradigms serve both the function to improve diagnosis and to detect conscious awareness in patients with VS and MCS. In case of a positive response indicated by electroencephalogram (EEG), they may provide a form of basic BCI communication.

2 Methods

2.1 Patients

The residual cognitive functions of a sample of 14 patients (4 female) diagnosed with PVS and MCS were assessed with an EEG-based test battery consisting in paradigms such as habituation, mismatch negativity and oddball with and without instructions [5]. A subsample of 7 patients (details are listed in Table 1) showed promising results (significant reactions in at least two of the experimental paradigms) and participated in the BCI experiment which includes two EEG sessions on two consecutive days. Two of the patients (MCS1, PVS5) were not able to participate in the second session.

Code	Sex	Age	Diagnosis	Years since ictus	Etiology
VS1	w	70	VS	0.5	Subarachnoidal hemorrhage
PVS2	w	55	PVS	4	Vascular hemorrhage
PVS3	m	48	PVS	4.5	Hypoxia
PVS4	m	48	PVS	1.5	TBI
PVS5	m	59	PVS	8	Hypoxia
MCS1	m	27	MCS	2	TBI
MCS2	m	63	MCS	1	TBI

Table 1: Characteristics of patient sample; (P)VS = (Persistent) Vegetative State, MCS = Minimally Conscious State, TBI = Traumatic Brain Injury.

2.2 Procedure

The patients participated in two auditory BCI experiments and the measurements were repeated on the following day. In the first part of the experiment, a three-stimulus oddball paradigm was used [6]. In each run of this paradigm, three different 75 dB tones were presented: 125 standard tones (pink noise), 25 "target one" (1000 Hz), and 25 "target two" (100 Hz) tones. The length of all tones was 80 ms with a randomized inter-trial interval (ITI) of 480–600 ms. In both sessions the patients completed 6 runs after being instructed to attend alternately to one of the targets. The patients VS1 and MCS2 did not participate in the three-stimulus oddball. In the second part, the patients participated in the auditory four-choice paradigm [7]. Four auditory stimuli (a female voice saying "yes", "no", "pass", "stop") were presented with a probability of 25 % each. The stimulus duration was 400 ms with an ITI (onset-to-onset) of 875 ms. The patients completed 8 runs with 25 sequences per run. EEG was measured using a 16-channel electrode cap (EASYCAP GmbH, Ag/AgCl electrodes) arranged according to the 10-20 system on electrode positions F3, Fz, F4, T7, C3, Cz, C4, T8, CP3, CP4, P3, Pz, P4, PO7, PO8, and Oz, referenced to the right and grounded to the left mastoid. Signals were recorded with a sampling rate of 256 Hz using a gUSBamp (g.tec) EEG amplifier and filtered between 0.1–30 Hz with an additional notch filter (48–52 Hz). Impedances were kept below 5 kΩ. The BCI2000 software controlled all aspects of stimulus presentation and data collection. Data was classified offline using stepwise linear discriminant analysis (SWLDA) for both paradigms. Segments of 800 ms length at the end of stimulus presentation were selected for classification of both targets and non-targets. The classifier was trained with a set of 5 runs (three-stimulus oddball) respectively 7 runs (four-choice paradigm) and tested on the residual run and this was repeated for every run as test set (leave-one-run-out-cross validation). The obtained classification accuracies were averaged for each subject and session.

3 Results

For the three-stimulus oddball mean accuracies between 0 and 66.67 % were achieved. With a chance level of 50 %, only one patient (PVS3) achieved a better-than-chance accuracy in one session (Figure 1 left).

In the four-choice BCI, the patients obtained offline accuracies between 0 and 87.5 % (Figure 1 right). Here, the accuracy of the best subject (MCS2) varied considerably between sessions. Besides MCS2 also PVS3 and PVS4 achieved an accuracy above chance level (25 %) in one session. Figure 2 shows averaged EEG responses of all patients in the four-choice paradigm for electrode Cz.

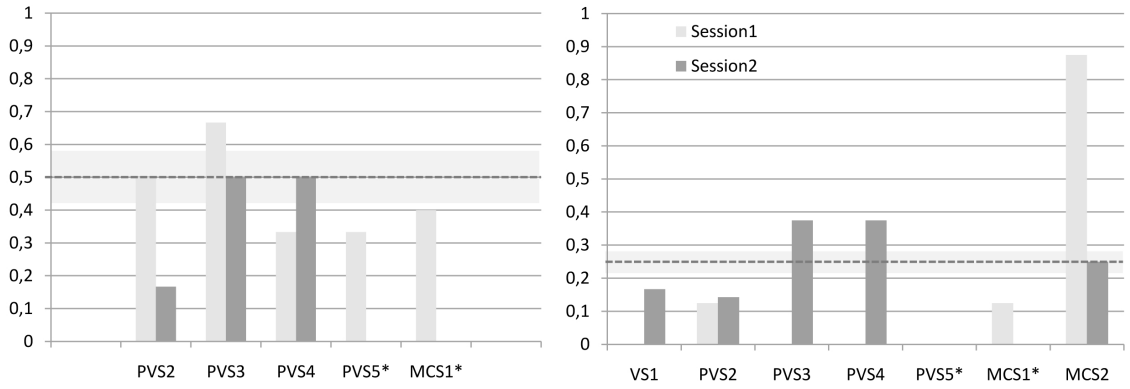


Figure 1: Offline classification accuracy for the three-stimulus oddball (left) and four-choice (right) paradigms. Dashed lines show chance levels, light grey bars the confidence intervals. Asterisks indicate the two patients who did not participate in the second session.

4 Discussion

Three out of the seven tested patients were able to generate classifiable responses (better-than-chance). Subject MCS2 provided the highest accuracy of 87.5 %. The ERP difference between target and non-target in this patient was larger in the second session than in the first one, but it was not classifiable above chance level. This indicates that the inter-trial variability obviously was very high, i.e., that the subject probably produced very atypical responses on different runs of the same session.

The classification accuracy of over 80 % in the four-choice task suggests that MCS2 was able to understand the instruction and to focus attention on the target word in the first session. [7] found an online classification accuracy of about 61.2 % in two sessions with healthy subjects and ALS patients in the same paradigm. A mean offline accuracy of about 92.5 % in 20 healthy subjects in the three-stimulus oddball paradigm showed the feasibility of this auditory paradigm in healthy subjects [6], but this could not be proven for our patient sample. We assume a higher workload having two rare target tones and a frequent standard tone. The differences in performance between the two sessions could be an indicator of large fluctuations of attention and consciousness in the tested patients. Neurophysiological measurements usually underestimate the existence of cognitive functions, and a lack of ERP components does not necessary imply the lack of cognitive functions [5], therefore it is not possible to make a negative judgement about the patients with low classification accuracy. Based on these results and as it was shown that patients can use residual awareness to express thoughts [1], the four-choice seems to be the most promising paradigm for BCI communication in patients referred to as non-responsive .

5 Conclusion

The study provides a justification for the implementation of BCI systems for improved diagnosis of conscious awareness and for establishing communication in non-responsive patients.

6 Acknowledgments

The authors would like to thank Colleen Dockery, Slavica von Hartlieb and Dimitra Pagania for their help in data collection. This work is supported by the Bundesministerium für Bildung und Forschung (BMBF) as part of Bernstein Focus Neurotechnology and CERES Tübingen. 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.

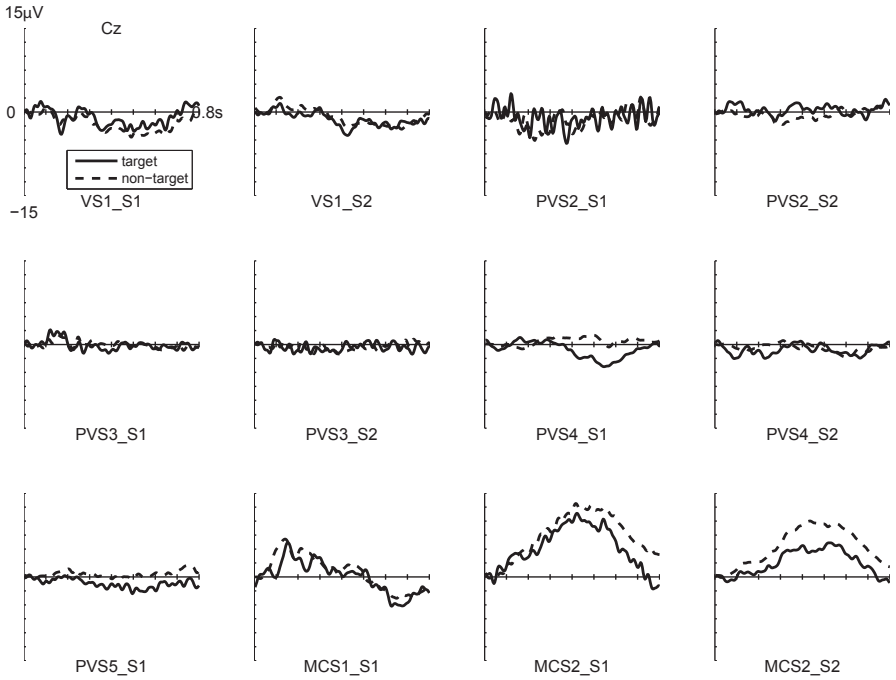


Figure 2: Averaged ERP responses of all patients for both sessions (four choice) for electrode Cz. S1 = first session, S2 = second session. Reference was not adapted.

References

[1] M. M. Monti, A. Vanhaudenhuyse, M. R. Coleman, M. Boly, J. D. Pickard, L. Tshibanda, A. M. Owen, and S. Laureys. Willful modulation of brain activity in disorders of consciousness. *N Engl J Med*, 362(7):579–589, Feb 2010.

[2] T. A. Bekinschtein, D. E. Shalom, C. Forcato, M. Herrera, M. R. Coleman, F. F. Manes, and M. Sigman. Classical conditioning in the vegetative and minimally conscious state. *Nat Neurosci*, 12(10):1343–1349, Oct 2009.

[3] B. Kotchoubey, S. Lang, G. Mezger, D. Schmalohr, M. Schneck, A. Semmler, V. Bostanov, and N. Birbaumer. Information processing in severe disorders of consciousness: Vegetative state and minimally conscious state. *Clinical Neurophysiology*, 116(10):2441 – 2453, 2005.

[4] A. Kübler and B. Kotchoubey. Brain-computer interfaces in the continuum of consciousness. *Curr Opin Neurol*, 20(6):643–649, Dec 2007.

[5] N. Neumann and B. Kotchoubey. Assessment of cognitive functions in severely paralysed and severely brain-damaged patients: neuropsychological and electrophysiological methods. *Brain Res Brain Res Protoc*, 14(1):25–36, Nov 2004.

[6] 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):516–523, Apr 2010.

[7] E. W. Sellers and E. Donchin. A P300-based brain-computer interface: initial tests by ALS patients. *Clin Neurophysiol*, 117(3):538–548, Mar 2006.

Fast and Reliable P300-Based BCI with Facial Images

A. Onishi[1,3], Y. Zhang[2,3], Q. Zhao[3], A. Cichocki[3]

[1]Department of Brain Science and Engineering, Kyushu Institute of Technology, Fukuoka, Japan
[2]School of Information Science and Engineering, East China University of Science and Technology, Shanghai, China
[3]Lab. for Advanced Brain Signal Processing, Brain Science Institute, Riken, Saitama, Japan

onishi@brain.riken.jp

Abstract

A P300-based brain-computer interface (BCI) often called "P300 speller" is a promising approach to help disabled people communicate with external world. However, the spellers using letters or symbols as stimuli usually require more than 5 repetitions to achieve high classification accuracy due to weak P300 evoked potential. We propose a suitable platform which has 8 command intuitive interface and uses human's facial images as flashes. The experimental results from on-line tests demonstrated that our P300-based BCI with facial images achieved over 90% accuracies by using only two repetitions. We also analyzed off-line EEG data using two or three channels and 2 paradigms achieved over 80 % accuracies. In addition, we found that the latencies of event-related potentials evoked by a clockwise flash order were shorter than that evoked by a random flash order.

1 Introduction

An electroencephalographic (EEG) brain-computer interface (BCI) can provide a non-invasive, low cost, non-muscular means of direct communication between a human brain and a computer or a robot [1, 2]. One of the most reliable and promising multi-command BCI system is based on P300 evoked potential paradigm. Most of P300-based BCI platforms exploit so called P300 speller with small abstract elements like letters and symbols arranged in the form of a matrix where each row and column is intensified in a random sequence [1]. Recently, big progresses have been made in optimization of the P300 speller, for example, on color and intensity of stimulation [3], size of the matrix [4], EEG electrode locations [5] and classification algorithms [6]. Most of the spellers using letters or symbols as stimuli require more than 5 repetitions to reliably extract P300 evoked potential since the P300 evoked potential is relatively weak and occurs amid other ongoing EEG activities. Our objective in this paper is to demonstrate how to enhance or increase P300 evoked potentials by suitably designed stimuli.

We present a novel P300-based BCI whose commands are intensified by natural images of human faces – the affective face driven paradigm (AFDP). In this paper we use P300-based BCI that has 8 independent commands to navigate, for example, a robot arm or a wheel chair moving in 8 different directions. Our experiments demonstrated that P300-based BCI using AFDP showed higher accuracies than that a gray/white flash of abstract symbols.

2 Methods

We tested 6 paradigms illustrated in Figure 1. The central colored arrow image shows a target or an output as a feedback. The other arrow images indicate control commands and the part is intensified sequentially during each trial. We prepared two different flash orders: random (R)

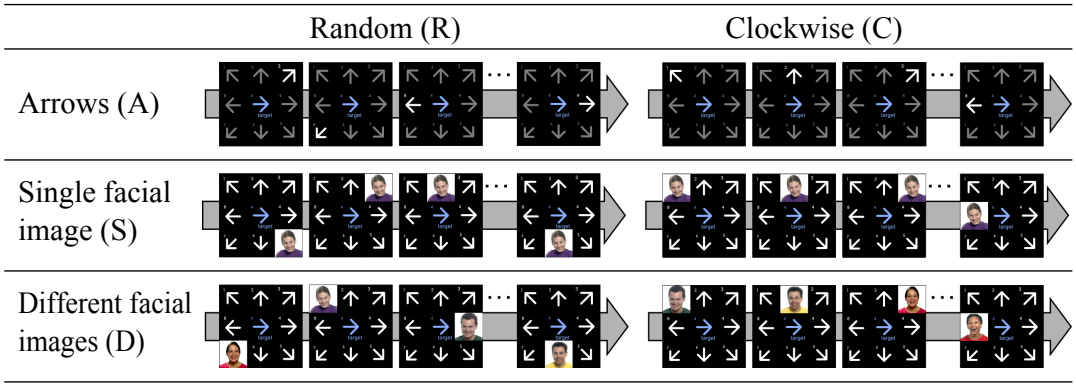


Figure 1: Illustration of visual stimuli for novel BCIs. The background black screen with 8 arrows was displayed continuously. The facial images or white arrows were flashed in random or clockwise order during 100 milliseconds (ms). The pause between two consecutive flashes was 20 ms. The top left panel shows a paradigm referred to as RA in which arrows are flashed randomly. In top right paradigm CA, arrows are also used as flashes but they flashed in clockwise order. RS paradigm means that a single face is displayed randomly. In CS paradigm the single face is flashed clockwise. In RD paradigm different facial images chosen randomly are used and finally in CD paradigm different facial images are flashed clockwise.

and clockwise (C) order. We also exploited three types of flash representations: arrows (A), a single facial image (S), and an image randomly selected from different facial images (D). The representation (A) is a gray/white symbol flicker but the other flash representations employ an natural image of human faces as stimuli. We randomized the experimental order to avoid order effects.

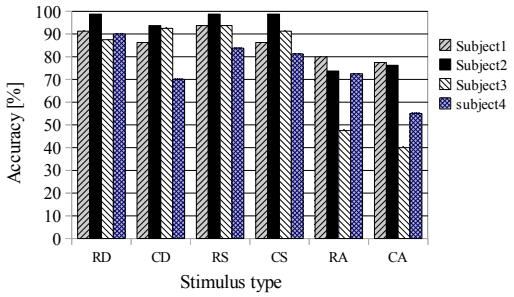
In our experiments, 4 healthy subjects were instructed to focus on a target arrow image cued by a center arrow image and silently count how many times the target was flashed. In a training stage, 16 targets were shown and each target was flashed 5 times. In a run of on-line test stages, 16 targets were shown but the number of target flashes to output a result was only 2 times. We repeated 5 runs for each stimuli. Subjects were seated in a comfortable chair placed 60 cm from a 17 inch LCD screen. Each stimulus cell size was 6×6 cm on the screen. We recorded EEG data and computed on-line classification accuracies. EEG signals were recorded at Fz, Cz, P3, Pz, P4, PO7, Oz and PO8 site following the 10-20 international system and each channel data was amplified by g.USBamp (Guger Technologies, Austria). The average of two mastoid electrodes was used as reference and the ground electrode was placed at AFz.

The 8 channel EEG data with 256 Hz sampling rate were filtered by a 50 Hz notch filter and a bandpass filter between 0.5 and 30 Hz and moving average was applied. Then the signals were down-sampled to 64 Hz. 700 ms buffers were made from each start time of a stimulus, subtracting baseline correlation by using 100 ms pre-stimulus data. After all arrows were intensified twice, each buffer was reconstructed into a feature matrix in which rows were feature vectors of all channels and columns indicated the stimuli. After that linear discriminant analysis (LDA) classifier was used to calculate the posterior probability of each column of the feature matrix, which was provided as Discriminant Analysis Toolbox by Dr. Michael Kiefte [http://www.mathworks.com/matlabcentral/fileexchange/189-discrim]. Finally, each probability was averaged to find the maximum score of all 8 arrows, taking it as an output.

3 Results

On-line classification performances of 6 paradigms are shown in Figure 2 (a). A classification accuracy was calculated as (# correct outputs) / (# total outputs). The number of flash repetitions

(a)



(b)

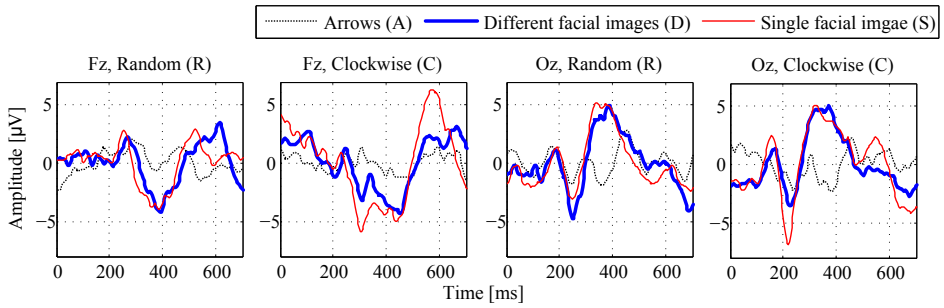


Figure 2: (a) On-line performance of our (8-channel) BCI system for four subjects. (b) Examples of averaged event-related potential waveforms obtained from training data of second subject.

for an output was only 2 times. Classification accuracies of RD, RS and CS, which were related to facial images, were over 80 % for all subjects, while that of RA and CA were less than 80 %. Every subjects achieved over 90 % individually in one of paradigms related to facial images. Clockwise order stimulus CS which contained no randomness also worked with more than 80 % accuracy.

Figure 2 (b) shows an example of averaged EEG waveforms of training sessions. Peak-to-peak voltages of the waveforms around 300 ms were over 5 μV when the arrows were intensified by different facial images or the single facial image, though the flashes by arrows evoked peaks below 5 μV. Also the latencies of event-related potentials evoked by the clockwise flash order were shorter than that evoked by the random flash order.

To be practical, the number of EEG electrodes should be minimum. We analyzed off-line EEG data using one, two and three channel combinations. Classifiers were built from the recorded training data. The number of repetitions was also 2 times.

Figure 3 depicts the off-line accuracies calculated by combinations of EEG channels for all 6 paradigms, where the four subject accuracies were averaged. Among one channel off-line classification results, Oz site accuracies of RD, CD, RS and CS paradigm showed over 50 % though Fz, Cz and Pz site accuracies were less than 50 %. Among two channel combinations, RD and RS paradigm accuracies of Cz & Oz and Pz & Oz showed above 80 %. Also we tested three channel combinations. The RD, RS, and CS paradigm accuracies with Cz & Pz & Oz, P3 & P4 & Oz and PO7 & Oz & PO8 also achieved above 80 %. But RA and CA paradigm accuracies of those channel combinations were less than 60 %.

4 Discussion

In the AFDP, features at Oz site enhanced its classification performance. Among the one channel off-line accuracies shown in Figure 3, only RD, CD, RS, CS at Oz site, in which facial images were used, showed over 50 % classification accuracies. In Figure 2 (b), Oz site waveforms of the single and different facial images showed over 5 μV peak-to-peak voltages and this feature might have

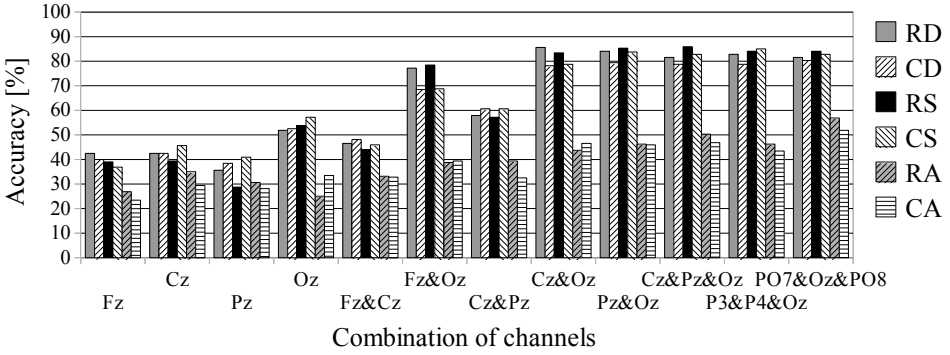


Figure 3: Off-line performance of our BCI systems using one, two and three channel combinations.

contributed to the classification accuracy. Taking those results into consideration, the features in Oz site were influenced by facial perception or were the mixture of visual evoked potentials from the occipital cortex. Thus the facial images increased the classification accuracies.

5 Conclusion

We have designed, implemented and preliminary tested new visual stimuli for P300-based BCI. Through the on-line experiment using only two repetitions, flashes with a single facial image and different facial images showed over 90 % accuracies. According to averaged waveforms we found that the waveform of flashes related to faces evoked higher peak-to-peak voltage compared to flashes with arrows. Also in the off-line tests, even two or three channels data of EEG we have obtained over 80 % accuracies. To our surprise, a paradigm which does not contain randomness also showed good performance. In our future work, we would like to extend these concepts to BCI system to control a robot arm, optimizing the stimulus, algorithms and functions.

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(6):510 – 523, 1988.

[2] R. Scherer, G. R. Müller, C. Neuper, B. Graimann, and G. Pfurtscheller. An asynchronously controlled EEG-based virtual keyboard: improvement of the spelling rate. *IEEE Transactions on Biomedical Engineering*, 51(6):979 –984, 2004.

[3] K. Takano, T. Komatsu, N. Hata, Y. Nakajima, and K. Kansaku. Visual stimuli for the P300 brain-computer interface: A comparison of white/gray and green/blue flicker matrices. *Clinical Neurophysiology*, 120(8):1562 – 1566, 2009.

[4] B. Z. Allison and J. A. Pineda. ERPs evoked by different matrix sizes: Implications for a brain computer interface (BCI) system. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11(2):110 –113, 2003.

[5] 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.

[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. *Journal of Neural Engineering*, 3(4):299–305, 2006.

Accuracy of a P300 Speller for Different Conditions: A Comparison

R. Ortner[1,2], R. Prückl[1,2], V. Putz[1], J. Scharinger[2], M. Bruckner[1], A. Schnürer[1], C. Guger[1]

[1]Guger Technologies OG, Sierningstrasse 14, Schiedlberg, Austria
[2]Department of Computational Perception, Johannes Kepler University, Linz, Austria

www.gtec.at

Abstract

When using a Brain-Computer Interface (BCI) as a spelling device, the P300-based BCI is mostly common, because of its high speed and accuracy, compared to other BCI paradigms. During the last years many studies have been performed to optimize the single spelling parameters, the stimulation frequency and stimulation intensity. Also different options for the electrode position and classification algorithm were analyzed intensively. But most of these studies were tested on healthy, young participants, subjects who may achieve better results than the people BCIs are mainly thought for: people with motor impairments. The tests were performed in a quiet laboratory environment were all possible outer distractions were filtered out. The aim of this publication is to take the results of a previous study, performed on healthy people within a laboratory (Group A) and to compare them to measurements on two other groups. Group B were healthy people that tested the device in a noisy environment, group C in contrast consisted of people with motor impairments. The conclusions deliver valuable results towards bringing BCIs to the end-user as they point out differences in performance for different groups of users.

1 Introduction

A Brain-Computer Interface (BCI) provides a new output pathway and so an additional technology to control external devices. According to Wolpaw et al. [1] it can be described as system for controlling a device e.g. computer, wheelchair or a neuroprothesis by human intention which does not depend on the brain's normal output pathways of peripheral nerves and muscles. For spelling devices a P300 based BCI is most common, due to its higher number of characters it is leading to a higher communication rate [2]. The P300 elicits when an unlikely event occurs randomly between events with high possibility (a so called oddball paradigm [3] is used for that) and was the first time described by Farwell and Donchin [4] in 1988. To elicit the P300 the user has to be mentally engaged to the task and wait for the unlikely event to occur. During the last years studies using a P300-based BCI were performed for people with motor impairments, e.g. Nijboer et al. [5] and Sellers and Donchin [6] tested a P300-based speller with people suffering amyotrophic lateral sclerosis (ALS). Nijboer et al. [7] also investigated the influence of psychological state and motivation in patients with ALS and Piccione et al. [8] examined the reliability and performance in healthy and paralysed participants. This study aims to compare the accuracy of different groups of participants when using the same paradigm and settings. In a previous study [9] we examined the overall accuracy of a P300 speller for healthy subjects. After five minutes of training the subjects were asked to spell 5 characters. It was up to the subjects to choose between a row/column (RC) speller and a single character (SC) speller. 72.8 % ($N = 81$) were able to spell with 100 % accuracy in the RC paradigm and 55.3 % ($N = 38$) spelled with 100 % accuracy in the SC paradigm. Less than 3 % of the subjects did not spell any character correctly. Following this first study, we tested the RC speller with exactly the same settings but different conditions.

2 Participants, Paradigm and Data Recording

Members of group A are the 81 subjects that tested the RC speller. For group B ($N = 50$) the environment was changed into a setting that resembled a possible "everyday situation". For that we asked visitors of the g.tec exhibition booth during the CeBIT 2011 in Hannover to instantly test the speller. During the session many other people were walking around, talking to the subject and also to each other. No exclusion criteria were given, apart of the anonymized EEG data no further information about the subjects (e.g. age or sex) was noted.

The measurements of group C were done at the Fundació Privada Institut de Neurorehabilitació Guttmann, Barcelona, Spain. Here, 10 people with motor impairments (two of them suffering locked in syndrome (LIS), and 8 person suffering SCI), used the device. Inclusion criteria were: Cervical Spinal Cord Injury (SCI) (between C2 and C6) and massive subcortical stroke patients with preserved cognitive function. The scale of the American Spinal Injury Association (ASIA) is evaluated for spinal cord injury patients to show the level of completeness of the injury. Six persons were classified A according to this scale, two of them were classified B. One person suffered diplopia.

The EEG data were acquired using eight active electrodes (positions: Fz, Cz, P3, Pz, P4, PO7, Oz, PO8; according to the international 10/20 system) made of sintered Ag/AgCl material. The ground electrode was located on the forehead; the reference was mounted on the right earlobe.

The speller showed a matrix (size 5×10), consisting of 26 characters (A, B,... Z), 10 numerical characters (0... 9), 5 punctuation marks (,.:!?) and 9 extra token (e.g. for ringing an alarm) on the computer screen. The RC speller highlighted a whole column or row for 100 ms (flashing time). The pause between the flashing was 75 ms each time (dark time). Each row and column was flashed fifteen times in a random order before a classification was done. After that the signal processing unit calculated the evoked potential for each character and performed a linear discriminant analysis (LDA) classification to determine which matrix item the subject was attending to. Then the highlighting sequence started again and the subject was prompted to attend to the next character. The subject's task was to attend to (or look at) the character she/he was prompted to spell and count how many times the character was highlighted. The BCI system was trained first on individual EEG data and therefore the subject was asked to "select" (or attend to) the word WATER, one letter at a time. This process took about 5 minutes. After training the BCI using this calibration data, the subject was asked to write the word LUKAS, one character at a time, taking about 5 more minutes. The spelling accuracy of each person was calculated by looking at the number of correctly spelled characters of the word LUKAS. When one person misspelled one character (e.g. LUFAS instead of LUKAS) then 4 out of 5 characters were correct and the accuracy was 80 %.

In a second step the accuracy versus the number of flashes were calculated offline for each classifier, using the validation run of the copy spelling task, where the subjects were asked to spell the word LUKAS. These plots show the reached accuracy for every single number of used flashes. This means that for the accuracy level at flash number one only the data of the first flashing event (each row and each column flashes one time) for each of the single characters was used for classification. For the level at flash number two the data of both, the first and second flashing event was taken, and so on. These plots help to optimize the number of used flashes for additional measurements.

3 Results

Table 1 shows the accuracy levels of the single groups. The average accuracy of group A was 91 %. In group B were 31 persons able to control the speller with 100 % accuracy, 9 persons reached 80 %, 6 persons 60 %, and 3 persons 40 %. Only one participant of this group was not able to spell one single character correctly (accuracy: 0 %).

Three out of ten people in group C controlled the speller with 100 % accuracy, another three subjects spelled one of five characters wrong (80 %). One person achieved an accuracy of 60 %.

Accuracy level (%)	Group A (%)	Group B (%)	Group C (%)
100	72.8	62.0	30.0
80	16.1	18.0	30.0
60	6.2	12.0	10.0
40	3.7	6.0	0.0
20	0.0	0.0	10.0
0	1.2	2.0	20.0
Average accuracy of all subjects (%)	91.0	86.0	62.0

Table 1: Accuracy levels of the single groups.

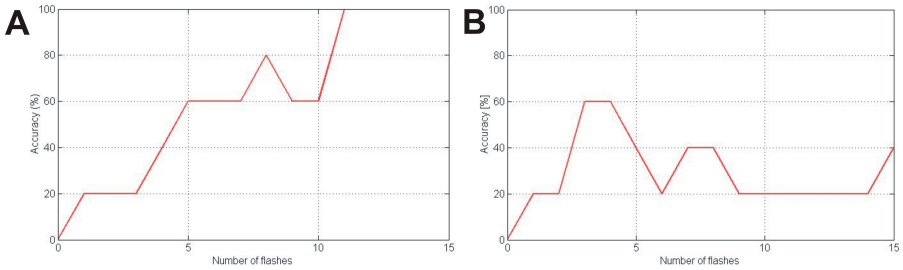


Figure 1: Plot of accuracy versus number of used flashes. A: The patient reached an accuracy of 100 % when using 11 or more flashes. Using less flashing events would lead to a faster spelling rate, but produces errors. B: This plot was processed out of the data of one of the patients suffering LIS with changed settings. Here the accuracy is best with only three or four flashes, using more of them even decreases the accuracy.

One subject had to stay in bed during the measurement. However, the monitor was placed on the desk as we did for the other patients, so the configuration did not fit the special needs of this patient. Due to this, the accuracy obtained by him was poor (20 %). Two subjects suffered LIS and were not able to control the speller with the given settings. These results were filled in the table for comparison of the accuracy. Nevertheless afterwards another measurement was performed, applying different settings (flashing time: 150 ms, dark time: 100 ms). Here the subjects reached an accuracy of 40 % and 20 %, respectively.

Figure 1 shows two plots of accuracy versus the number of used flashes for two subjects out of group C. In Figure 1 A the patient reached an accuracy of 100 % when using 11 or more flashes. Using less flashing events would lead to a faster spelling rate, but produces errors. Of special interest is the plot in Figure 1 B, because it was processed out of the data of one of the patients suffering LIS.

4 Discussion and Conclusion

In this study we compare data from 141 people, divided into three groups. As expected, the overall accuracy of group A, were healthy subjects sat in a laboratory environment, was best. The overall accuracy of group B (healthy people, noisy environment) is only five percentage points lower than in group A. This shows clearly that the device is also practicable within "everyday life". The accuracy of group C, is very heterogeneously. Two people suffering LIS were not able to control the speller with the predefined setting. However, changing the settings to a slower spelling rate (flashing time: 150 ms, dark time: 100 ms) resulted in at least one correct spelled character for each patient. Therefore, tuning these parameters seems to be necessary. Giving into account that not more time than two hours was given with each subject, we think that with more training time and different settings it might still be possible to use the device for LIS patients. One

important hint for this assumption is the fact that for one of them the offline accuracy reached 60 % (see Figure 1 B) when using only three or four flashes (and setting the flashing time and dark time to higher values). This offline analysis means that the person would have done best when the paradigm would have stimulated with three or four flashes. We believe that the decrease of the accuracy for an increasing number of flashes is due to the fact that this patient suffers from diplopia so it was difficult for him to concentrate on one character for a longer period of time.

5 Acknowledgments

The research leading to these results has received funding from the European Community's, Seventh Framework Programme FP7/2007-2013 under BrainAble project, grant agreement n° 247447. The authors further gratefully acknowledge the funding by the European Commission under contract FP7-ICT-2009-247935 (Brain-Neural Computer Interaction for Evaluation and Testing of Physical Therapies in Stroke Rehabilitation of Gait Disorders).

References

[1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan. Brain-computer interfaces for communication and control. *Clinical Neurophysiology*, 113:767–791, 2002.

[2] D. Krusienski, E. Sellers, F. Cabestaing, S. Bayoudh, D. McFarland, T. Vaughan, and J. Wolpaw. A comparison of classification techniques for the P300 speller. *Journal of Neural Engineering*, 6:299–305, 2006.

[3] E. Donchin, W. Ritter, and C. McCallum. *Brain-event related potentials in man*, chapter Cognitive psychphysiology: the endogenous components of the ERP, pages 349–411. Academic, New York, 1978.

[4] L. A. Farwell and E. Donchin. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalogr. Clin. Neurophysiol*, 70(6):510–523, 1988.

[5] F. Nijboer, E. W. Sellers, J. Mellinger, M. A. Jordan, T. Matuz, A. Furdea, S. Halder, U. Mochty, D. J. Krusienski, T. M. Vaughan, J. R. Wolpaw, N. Birbaumer, and A. Kübler. A P300-based Brain-Computer interface for people with amyotrophic lateral sclerosis. *Clin Neurophysiol*, 119(8):1909–1916, 2008.

[6] E. Sellers and E. Donchin. A P300-based brain-computer interface: initial tests by ALS patient. *Clin Neurophysiol*, 2006:538–548, 2006.

[7] F. Nijboer, N. Birbaumer, and A. Kübler. The influence of psychological state and motivation on brain-computer interface performance in patients with amyotrophic lateral sclerosis - a longitudinal study. *Front Neurosci.*, 4:55, 2010.

[8] 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. *Clin Neurophysiol*, 117:531–537, 2006.

[9] C. Guger, S. Daban, E. Sellers, C. Holzner, G. Krausz, R. Carabalona, F. Gramatica, and G. Edlinger. How many people are able to control a P300-based brain-computer interface (BCI)? *Neuroscience Letters*, 462(1):94–8, 2009.

Towards a Single-Switch BCI Based on Steady-State Somatosensory Evoked Potentials

C. Pokorny[1], C. Breitwieser[1], C. Neuper[1,2], G. R. Müller-Putz[1]

[1]Institute for Knowledge Discovery, BCI Lab, Graz University of Technology, Graz, Austria
[2]Department of Psychology, University of Graz, Graz, Austria

christoph.pokorny@tugraz.at

Abstract

A single-switch Brain-Computer Interface (BCI) based on steady-state somatosensory evoked potentials (SSSEPs) was designed with the aim to provide non-responsive patients with a means of communication. The main focus of this study was to investigate, whether a single-switch BCI can be realized at all based on SSSEP. First, two different stimulation frequencies were selected using a screening procedure. Then, tactile stimuli with the selected frequencies were applied to the thumb and the middle finger of the subjects' right hand. The subjects were instructed to randomly focus attention on one of the fingers. Both classes were classified against the reference period using lock-in analyzer system (LAS) features and two linear discriminant analysis (LDA) classifiers. The offline classification results were compared and the class with the higher classification accuracy was selected to be used to activate a brain switch. Thirteen out of fourteen healthy subjects performed above chance level for at least one class. This study shows that a single-switch BCI can be realized based on SSSEP in healthy subjects.

1 Introduction

A Brain-Computer Interface (BCI) can provide a means of communication for non-responsive patients. Non-responsive patients are patients who have lost all motor functions due to a severe neurological disease, such as amyotrophic lateral sclerosis (ALS). Several studies showed that patients suffering from ALS are able to operate different kinds of BCIs [1, 2]. Patients with severe neurological diseases may also lose volitional gaze control, making them unable to use vision-based BCIs [3]. To overcome this problem, BCIs based on auditory or tactile stimuli can be used as alternatives. In healthy subjects, a tactile two-class BCI based on steady-state somatosensory evoked potentials (SSSEPs) was successfully applied [3]. In that study, the index fingers of the left and right hand were simultaneously stimulated with different frequencies found in a screening procedure. Two out of four subjects learned to modulate the elicited SSSEPs by focusing attention on one of the fingers.

For patients with severe neurological diseases operating a two-class BCI might be too demanding. However, using a single-switch BCI, only one single brain response that can reliably be detected is sufficient to activate a switch. Eventually, the single-switch BCI can be connected to assistive technology devices that can be controlled by a conventional single switch in order to provide patients with means of communication and control.

The focus of the present study was to investigate, if it is possible to realize a single-switch BCI based on SSSEP in healthy subjects. Starting point of this investigation was a two-class BCI based on SSSEP with two fingers of one hand simultaneously stimulated with different frequencies, individually selected for each subject. The two classes were compared in an offline analysis to find the more responsive class which can be used to activate a brain switch.

2 Methods

Fourteen healthy subjects participated in this study. The subjects were informed about the purpose of the study and were paid for participation. All subjects participated voluntarily and gave informed consent.

Tactile stimuli were applied to the thumb and the middle finger of the subjects' right hand. The fingers were stimulated using C2 tactors [Engineering Acoustics, Inc., Casselberry, Florida, USA]. The tactors were attached to the subjects' fingers using finger clips. The stimulation patterns were created using a self-made stimulation device and consisted of a 200 Hz sinusoidal signal modulated with a rectangular signal (duty cycle 50 %) of the respective stimulation frequency [3].

The EEG was recorded using three g.USBAmps [Guger Technologies OG, Graz, Austria] with 48 Ag/AgCl electrodes. The ground electrode was attached to the tip of the nose. Linked mastoids were used as reference. All impedances were kept below 5 kΩ. A sampling rate of 2.4 kHz was used and all EEG measurements were done in a shielded room. For further analyses, the exact electrode positions were measured using the electrode positioning system ELPOS [zebris Medical GmbH, Isny, Germany].

2.1 Experimental Paradigms

The first paradigm of this study was a screening procedure (similar as proposed in [4]) to identify the subjects' "resonance-like" frequencies together with a tuning curve. After the screening procedure, two stimulation frequencies were individually selected for each subject to be used in the single-switch paradigm.

2.1.1 Screening Paradigm

The thumb of each subject was randomly stimulated with twelve frequencies ranging from 13 Hz to 35 Hz in steps of 2 Hz, with 60 trials per frequency. Only the thumb was stimulated assuming that the tuning curves of all fingers are similar [5]. Each trial started with a reference period without stimulation followed by twelve stimulation periods with random stimulation frequencies, as shown in Figure 1. To avoid the subjects to focus attention on the stimuli during the screening procedure, the subjects had to perform a distracting mental arithmetic task.

After the screening procedure, the fast Fourier transform (FFT) was computed and two stimulation frequencies were manually selected for each subject to be used in the single-switch paradigm. The two stimulation frequencies with the highest peaks in the spectrum (compared to the reference period) which were similarly high and separated by at least one other stimulation frequency were selected.

2.1.2 Single-Switch Paradigm

The subjects' thumb and middle finger were simultaneously stimulated using the two frequencies selected after the screening procedure. The subjects were instructed to randomly focus attention on one of the fingers. In order to make it easier to focus attention on one of the fingers, short twitches (i.e., amplitude attenuations similar as proposed in [3]) were inserted in the stimuli of both fingers at random time points. As shown in Figure 2, each trial started with a reference period during which the subject was instructed to just look at the center of a blank screen. After that, a text faded in on the screen instructing the subject to focus attention on one of the fingers. Tactile stimulation was applied on both fingers during the whole trial. In total, the paradigm consisted of 80 trials per class.

2.2 Analysis

EMG (electromyogram) artifacts were manually selected in all data sets and trials containing artifacts were excluded from data analysis. As a first attempt, only the bipolar channel FC3-CP3 that had the highest FFT magnitude across most subjects was selected for classification.

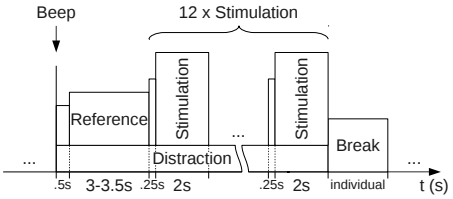


Figure 1: Screening paradigm. After a beep tone, each trial started with a reference period without stimulation (random length 3 s to 3.5 s) followed by twelve stimulation periods (length 2 s each) with random stimulation frequencies. A distracting mental arithmetic task was presented during the whole trial.

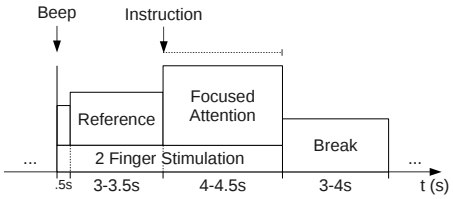


Figure 2: Single-switch paradigm. After a beep tone, each trial started with a reference period (random length 3 s to 3.5 s) followed by a period of focused attention on one of the fingers (random length 4 s to 4.5 s). Tactile stimulation was applied on both fingers during the whole trial.

Magnitude features were extracted at the two stimulation frequencies using a lock-in analyzer system (LAS) [3]. The resulting two time series were smoothed with a moving average filter (length 1 s) and logarithmized. Two linear discriminant analysis (LDA) classifiers (Fisher's LDA) were trained using 10×10 cross-validation. Both classes (focused attention on the thumb and on the middle finger) were separately classified against the reference period (no focused attention). The classification results were compared and the class with the higher classification accuracy was selected and can in principle (i.e., when being above chance level) be used to activate a switch.

3 Results

As an example, Figure 3 shows the offline classification accuracies of both classes (focused attention on the thumb and on the middle finger) against the reference period (no focused attention) for subject s7 (randomly chosen). While the classification accuracy for the thumb remains at chance level, the classification accuracy for the middle finger starts increasing after around 1 s and reaches a maximum value of 65 % after around 2 s. The maximum accuracy for the middle finger is above chance level (significance level $\alpha = 5\,\%$) [6]. Because of removed trials due to EMG artifacts, the number of used trials was varying. Therefore, also the confidence limit of the chance level is different across classes and subjects. Table 1 shows the classification accuracies of the single-switch paradigm for all subjects. The time points with the highest classification accuracy were chosen.

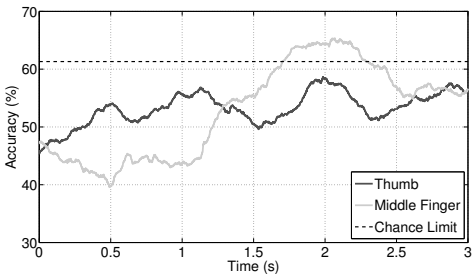


Figure 3: Offline classification accuracies for focused attention on the thumb (dark gray curve) and on the middle finger (light gray curve) against reference for subject s7. The dashed horizontal line at 61 % indicates the confidence limit of the chance level ($\alpha = 5\,\%$).

Subject	s1	s2	s3	s4	s5	s6	s7
Cl1 (%)	**74**	**72**	64*	76	**69**	**70**	58*
Cl2 (%)	73	64	**65**	**78**	64	68	**65**
Subject	s8	s9	s10	s11	s12	s13	s14
Cl1 (%)	**74**	65	**67**	**63**	60*	**64**	57*
Cl2 (%)	72	**66**	66	58*	**68**	**64**	59*

* Below chance limit ($\alpha = 5\,\%$).

Table 1: Maximum offline classification accuracies for focused attention on the thumb (Cl1) and on the middle finger (Cl2) against reference. The class with the higher accuracy (marked bold) was selected and can be used to activate a switch.

4 Discussion

This study shows that a single-switch BCI can be realized based on SSSEP in healthy subjects. Thirteen out of fourteen healthy subjects performed above chance level for at least one class, nine out of fourteen even for both classes. The class (i.e., the finger) with the higher accuracy can in principle be used to activate a brain switch. Within this study, the transition from a two-class BCI to a single-switch BCI based on SSSEP was demonstrated when two fingers of one hand are simultaneously stimulated with two different frequencies. The frequencies used for stimulation were individually selected by a screening procedure where similar effects as described by Müller-Putz et al. [4] could be observed.

So far, only offline measurements were carried out. To activate a switch in an online paradigm, the output of the classifier with the higher classification accuracy together with a threshold can be used. To increase classification accuracy, investigations about the optimal electrode positions and channel combinations (among all recorded 48 channels) are necessary. According to Burton et al. [7], also regions other than the primary somatosensory cortex are involved in vibrotactile attention. Moreover, higher harmonics of the stimulation frequencies have not yet been taken into account for classification. If a single-switch BCI can be realized with stimuli applied to just one single finger remains an open question and needs to be further investigated, too.

In summary, this study shows that a single-switch BCI can be realized based on SSSEP in healthy subjects. This study is the first step towards a single-switch BCI that eventually may provide non-responsive patients with an alternative means of communication and control.

5 Acknowledgments

The authors would like to thank Fondazione Santa Lucia for providing their C2 tactors. 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] A. Kübler, F. Nijboer, J. Mellinger, T. M. Vaughan, H. Pawelzik, G. Schalk, D. J. McFarland, N. Birbaumer, and J. R. Wolpaw. Patients with ALS can use sensorimotor rhythms to operate a brain-computer interface. *Neurology*, 64:1775–1777, 2005.

[2] F. Nijboer, E. W. Sellers, J. Mellinger, M. A. Jordan, T. Matuz, A. Furdea, S. Halder, U. Mochty, D. J. Krusienski, T. M. Vaughan, J. R. Wolpaw, N. Birbaumer, and A. Kübler. A P300-based brain-computer interface for people with amyotrophic lateral sclerosis. *Clinical Neurophysiology*, 119:1909–1916, 2008.

[3] G. R. Müller-Putz, R. Scherer, C. Neuper, and G. Pfurtscheller. Steady-state somatosensory evoked potentials: suitable brain signals for brain-computer interfaces? *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 14:30–37, 2006.

[4] G. R. Müller-Putz, C. Neuper, and G. Pfurtscheller. Resonance-like frequencies of sensorimotor areas evoked by repetitive tactile stimulation. *Biomedizinische Technik*, 46:186–190, 2001.

[5] C. Breitwieser. Are steady-state somatosensory evoked potentials suitable for biometric use? Master's thesis, Graz University of Technology, 2009.

[6] G. R. Müller-Putz, R. Scherer, C. Brunner, R. Leeb, and G. Pfurtscheller. Better than random? A closer look on BCI results. *International Journal of Bioelectromagnetism*, 10:52–55, 2008.

[7] H. Burton, R. J. Sinclair, and D. G. McLaren. Cortical network for vibrotactile attention: a fMRI study. *Human Brain Mapping*, 29(2):207–221, 2008.

Detection of Error Potentials During a Car-Game with Combined Continuous and Discrete Feedback

A. Kreilinger[1], C. Neuper[1,2], G. R. Müller-Putz[1]

[1]Institute for Knowledge Discovery, Graz University of Technology, Graz, Austria
[2]Department of Psychology, University of Graz, Graz, Austria

alex.kreilinger@tugraz.at

Abstract

This work describes an experiment designed to use continuous feedback in terms of a car game with additional discrete feedback to record error potentials (ErrPs). The game feedback allowed free movement of a car from the left to the right side of a street while moving forward with constant speed. Randomly appearing coins and barriers were required to be picked up or avoided. In case of successful collections or unwanted collisions visual and acoustic signals were presented as discrete feedback. An offline analysis was conducted to evaluate time periods after these discrete feedback events to investigate ErrPs after collisions with barriers. The found detection rates were above chance level for most of the subjects.

1 Introduction

Error potentials (ErrPs) [1], specifically interaction ErrPs [2], provide a promising possibility to increase the performance of brain-computer interfaces (BCIs). By detecting specific reactions to errors that differ from reactions to correct events, false actions can be inhibited and therefore the accuracy of BCI-driven systems can be increased. If an ErrP classifier detects a possibly false reaction it allows users to redo commands instead of picking another command automatically. Several studies already mentioned the technical capabilities of error correction for various paradigms [2, 3]. The paradigms used in these experiments have in common that they are designed to work well for ErrP processing, i.e., due to a discrete setup of the feedback ErrPs can be detected easily by evaluating time periods following discrete events. This study aimed to show that ErrPs can also be recorded during continuous feedback which might be useful in real-life applications. The difficulty here is, however, that ErrPs are time- and phase-locked, i.e., they need specific discrete events to facilitate detection. Also, processing of errors needs definite triggers to know when to look for eventual errors. The study introduced here is a follow-up to [4] where a continuous feedback was already coupled with additional discrete events and ErrPs were successfully found in offline analysis. However, the accuracy for single trial detection of these ErrPs was not enough to be feasible in online applications.

A new paradigm was designed to find a better solution for the integration of ErrPs into online, continuous feedback. A game-like feedback, controlled with motor imagery (MI) [5], was set up to deliver information about the current state in two ways: (i) continuously by a movable car; (ii) discretely by applying distinct flashes/sounds in case of encounters with objects. The main goal of the study was to show that ErrPs could be evoked after presentation of negative discrete feedback after collision with a barrier on top of the constantly active continuous feedback. Another interesting point was whether the modality of the feedback type for discrete events has a positive or negative effect on the evoked ErrPs. Therefore, half of the conducted runs were recorded with visual and acoustic discrete feedback whereas the other half was only using the visual kind.

2 Methods

2.1 Subjects, Hardware, and Recording

Five female and five male subjects (24.9±2.3 years) participated in the study. Data was recorded with two g.USBamps (Guger Technologies OEG, Graz, Austria). The 32 Ag/AgCl-electrodes were placed on the scalp of the subjects following the international 10-20 system. Therefore, all of the important regions for MI (C3, Cz, C4) and ErrP detection (the area over the anterior cingulate cortex (ACC) [6] at channels Fz and Cz) were covered. All channels were recorded monopolarly with a reference electrode at the left mastoid. The ground electrode was mounted on the right mastoid. The sample rate was set to 512 Hz with a high-pass filter at 0.5 Hz, a low-pass filter at 100 Hz, and a notch filter at 50 Hz.

2.2 Experiment Setup

The first part of the experiment consisted of two training runs with the standard Graz-BCI paradigm [5] where subjects were asked to perform MI of the two required classes (right hand vs. both feet), in total 40 trials per class. The low number of trials was possible because all subjects, minus one exception, had previous experience in MI BCI. The person without experience was asked to perform 80 trials per class. ERD/S maps [7] were calculated after common average reference (CAR) was applied on channels C3, Cz, and C4. According to these maps, features (band powers in frequency bands of the three channels) were selected manually to generate a classifier based on linear discriminant analysis (LDA).
The second part consisted of six runs with 20 trials left versus 20 trials right. The LDA classifier was constantly active, letting the subjects move the car all the time. However, actual control of the car was only required when coins were able to be collected on either one side of the street. During every trial four coins could be collected, accompanied by exactly the same number of barriers on the opposite side. Both objects appeared in intervals of one second. In total, a trial lasted 10 s from the appearance of a starting line to the collection of the last object. This setup allowed a maximum collection of 960 coins after all six runs. Half of the runs were recorded with sound (short beeps with different frequencies) and visual (increased size and change of color of the car) feedback combined, the other half used only visual feedback. The discrete feedback events were chosen to be neutral for coin and barrier collisions: instead of linking a bright/positive flashing color and an encouraging sound to coins, the sounds and flashes were merely depending on the side of the street where the current collision occurred.

2.3 Analysis

For offline analysis all channels were spatially filtered with CAR and a bandpass between 0.5 and 8 Hz was applied. Only data within a one-second window following discrete events (collisions with coins and barriers) was evaluated. Here, the data was split into two different classes, the time windows containing either reactions to correct events or to errors. A 10×10 crossvalidation was applied to test the classification accuracy. For each crossvalidation cycle, features were selected with a discriminant power algorithm.

3 Results

3.1 Motor Imagery BCI

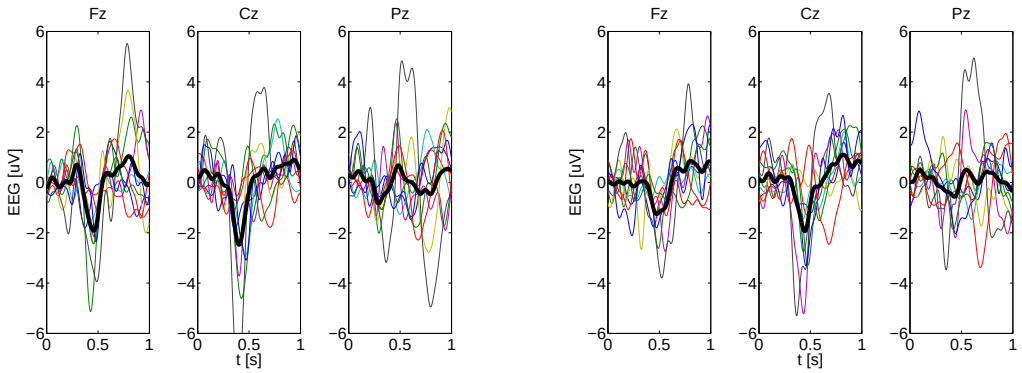
Online performance was measured in terms of scoring points. The score was increased by collecting a coin and reduced after collisions with barriers. Subjects could also miss coins within trials; in this case the score was not altered and neither a positive nor a positive feedback occurred. Further, a negative score was not possible. Table 1 shows the performance for each participant.

Subject	Total Score	Coins:Barriers	Subject	Total Score	Coins:Barriers
S1	409	599:192	S6	202	489:310
S2	718	782:65	S7	703	774:72
S3	588	709:127	S8	751	823:72
S4	189	367:185	S9	100	404:349
S5	722	797:75	S10	265	541:302

Table 1: Online performance in terms of total score out of 960 maximum points and coin:barrier collection rate for each participant which is also the correct:erroneous trial rate for ErrP analysis.

3.2 Error Potential Detection

The performance depending occurrences of errors evoked differently strong ErrPs for the 10 subjects. The error-minus-correct waveform for all the participants can be observed in Figure 1. Three channels, Fz, Cz, and Pz, are shown for the two different modalities 'Sound' and 'No Sound'. The greatest effect is visible over Fz and Cz, the channels directly over the ACC. On average there is a measurable negativity about 400 ms after the moment of a collision with a barrier (event of an error).



(a) Sound (b) No Sound

Figure 1: Recorded ErrPs for all individual subjects and averaged waveforms, shown specifically for 'Sound' and 'No Sound' modality. Here, point in time zero is the time of the collision with an object on the street. Three subplots demonstrate the measured ErrPs on three different channel locations: Fz, Cz, and Pz.

An offline analysis of the recorded data yielded detection rates for correct and erroneous trials above random for all subjects except S10 in at least one modality ('Sound','No Sound', or all trials combined). The chance level was on average 53 % (all trials combined), 54.4 % ('Sound'), and 54.3 % ('No Sound') according to [8]. The particular values can be seen in Table 2. Feature extraction and cross validation was performed individually for each modality. The comparison of detection rates for feedback with and without sound brought forth following results: the average accuracy for correct detections of correct and erroneous trials was 61.4 % for the two modalities combined, 61.9 % for only runs with sound, and 59.2 % for runs without sound feedback. A performed t-test showed no significant differences in terms of timing and shape of the waveforms.

4 Discussion & Conclusion

The measured ErrPs were different to the interaction ErrPs described in [2] (with a negative peak followed by a positive and another negative peak) but showed an error-related negativity (ERN) about 400 ms after objects were picked up. Addition of sound feedback had a beneficial effect for 7 out of 10 subjects when compared to trials with only visual feedback but did not alter the resulting

Subject	Accuracy [%]			Subject	Accuracy [%]		
	Sound	No Sound	Combined		Sound	No Sound	Combined
S1	64.3	61.1	60.2	S6	59.4	56.0	52.8
S2	75.8	66.5	71.1	S7	79.2	68.7	76.0
S3	56.0	51.1	54.0	S8	58.1	61.0	64.6
S4	64.5	55.6	59.4	S9	53.1	59.8	57.7
S5	56.3	60.7	63.7	S10	52.3	51.6	54.0

Table 2: Offline detection rates for correct and erroneous trials. The results for 'Sound', 'No Sound', and 'Combined' are shown to demonstrate the difference between the feedback modalities.

waveform of the ErrPs significantly. Considering that the paradigm was not optimized to trigger ErrPs, e.g., by giving only discrete feedback and artificially setting a fixed error rate of 20 %, it was still possible to find the ERN after collisions with barriers. Still, for future experiments it will be necessary to find a practical compromise of optimization and applicable feedback systems. A feasible daily life application for the type of feedback used in the study would be to present discrete feedback during continuous states. As an example, users could control a wheelchair continuously but during particular states (e.g., moving forward, turning left/right) visual and/or acoustic events could be added to inform about the current state. Events during unintended states could then be used to evoke ErrPs and alter decisions that can support moving to another state which should be the initially intended one.

Acknowledgments

We thank University of Glasgow for providing the car-game. This work is supported by the European ICT Programme Project FP7-224631. 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] M. Falkenstein, J. Hoormann, S. Christ, and J. Hohnsbein. ERP components on reaction errors and their functional significance: a tutorial. *Biological Psychology*, 51(2–3):87–107, 2000.

[2] P. W. Ferrez and J. del R. Millán. Error-related EEG potentials generated during simulated brain-computer interaction. *IEEE Transactions on Biomedical Engineering*, 55(3):923–929, 2008.

[3] B. Dal Seno, M. Matteucci, and L. Mainardi. Online detection of P300 and error potentials in a BCI speller. *Computational Intelligence and Neuroscience*, (307254):5 pages, 2010.

[4] A. Kreilinger, G. R. Müller-Putz, and C. Neuper. Time coded motor imagery BCI to control an artificial limb with additional discrete feedback to detect error potentials. In *Proceedings of the 1st TOBI Workshop 2010*, 2010.

[5] G. Pfurtscheller and C. Neuper. Motor imagery and direct brain-computer communication. *Proceedings of the IEEE*, 89:1123–1134, 2001.

[6] D. H. Mathalon, S. L. Whitfield, and J. M. Ford. Anatomy of an error: ERP and fMRI. *Biological Psychology*, 64(1–2):119–141, 2003.

[7] G. Pfurtscheller and F. H. Lopes da Silva. Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical Neurophysiology*, 110:1842–1857, 1999.

[8] G. R. Müller-Putz, R. Scherer, C. Brunner, R. Leeb, and G. Pfurtscheller. Better than random? A closer look on BCI results. *International Journal of Bioelectromagnetism*, 10:52–55, 2008.

Online Detection of Error-Related Potentials Boosts the Communication Speed of Visual Spellers

N. M. Schmidt[1,2], B. Blankertz[1], M. S. Treder[1]

[1]Machine Learning Laboratory, Berlin Institute of Technology, Berlin, Germany
[2]Artificial Intelligence Laboratory, University of Zurich, Zurich, Switzerland

nico.schmidt@uzh.ch

Abstract

Increasing the communication speed of brain-computer interfaces (BCIs) is a major aim of current BCI-research. One so far little explored approach is to automatically detect error-related potentials (ErrPs) in order to veto erroneous decisions of a BCI. In a study with eleven participants, an ErrP detection mechanism was implemented in an electroencephalography (EEG) based visual speller. Single-trial ErrPs were detected online with a mean accuracy of 79 %. The spelling speed was increased on average by 29.4 % by using ErrP detection.

1 Introduction

A brain-computer interface (BCI) based on event-related potentials (ERPs) exploits the fact that the neural processing of a stimulus is modulated by attention, like in the classic *Matrix Speller* [6], or the recently presented gaze independent *Center Speller* [12, 13]. For the Center Speller, a previous study showed an average spelling speed of about 1.5 characters/minute. One so far little explored approach to increase the spelling speed is the detection of error-related potentials (ErrPs). ErrPs are a certain type of ERPs that are present in the EEG signals when the user is aware of erroneous behavior. The characteristics of the ErrPs vary, depending on the situation in which the erroneous behavior was perceived. [7] distinguish between "response ErrPs" that were observed during errors made in choice reaction tasks [1], "feedback ErrPs" in reinforcement learning tasks [8], "observation ErrPs" when observing erroneous behavior [14], and "interaction ErrPs" which can be observed when the BCI performs wrongly [4, 7, 10].

In BCI field, few studies have been conducted on the detection of interaction ErrPs. ErrP detection has been used to detect error trials offline in EEG-data of motor imagery experiments [9] and in EEG-data of button press experiments with artificially induced errors [7]. [4] used online ErrP detection in an online BCI Matrix Speller experiment with three participants. The only comprehensive online study so far is [10] which demonstrates successful online ErrP detection with the Matrix Speller in 12 healthy participants and 4 patients with motor disorders.

The aim of the present study was to investigate, whether the communication rate of the gaze-independent Center Speller can be increased using online detection of ErrPs.

2 Methods

Center Speller The selection process of one symobl in the Center Speller is broken down into two levels, see Figure 1.

ErrP Calibration Speller To train an ErrP classifier, one has to collect a sufficient amount of trials. However, because spelling with the Center Speller is slow, participants need to spell a lot of characters in order to obtain a moderate training set for the ErrP detection. Depending on the desired size of the training set, this could exceed the reasonable duration of an experiment.

As an alternative, a calibration experiment was designed, wherein the participants spell via key press in a much faster way. The spelling process and the appearance of this experiment was

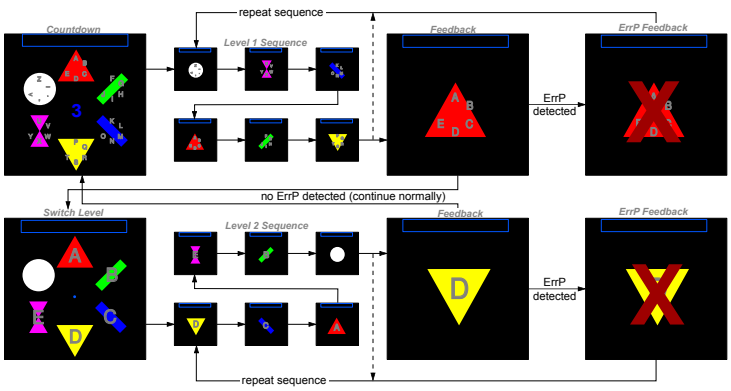


Figure 1 – Program flow of the Center Speller with ErrP detection: The symbol set is divided into six groups of five characters each, and each group is associated with a polygon of unique geometric shape and color. After a countdown (top left), stimuli are presented several times sequentially in a random order. Then the feedback (selected symbol group) is presented, and the user's response to that is evaluated. If an ErrP is detected, the stimulus sequence repeats, otherwise the second level starts, where the selection process is repeated on the single character level. Finally, the selected character is appended to the phrase.

designed to be as similar as possible to the Center Speller, in order to assure that the classifier trained on the data of this transfers to the Center Speller. Instead of presenting the elements sequentially in the center of the screen, they are shown in a hexagonal arrangement throughout the selection process, while a small arrow rotating clockwise in the center of the screen is used to select the target. The participant has to press a key when the arrow points on the target symbol. The feedback presentation is identical to the Center Speller. In 15 % of the trials, errors were induced artificially, in order to ensure a fast collection of error trials.

Procedure Three conditions were compared in nine Center Speller blocks: spelling without ErrP detection, spelling with ErrP detection using a classifier that was trained on the data of the ErrP Calibration Speller (classifier A) and spelling with ErrP detection using a classifier that was trained on the Center Speller data (classifier B). The training data of classifier A was captured in a preceding ErrP Calibration Speller block, classifier B was trained on the data of the first four Center Speller blocks.

Participants, Apparatus and Data Analysis Eleven participants (6 m, 5 f), aged 23–31 years ($\mu = 26$), took part in the study. All gave written consent and the study was performed in accordance with the Declaration of Helsinki. EEG was acquired from 64 electrodes placed according to the international 10-10 system with a nose reference and sampled at 100 Hz. Both the spelling classifier and the two ErrP classifiers were based on linear discriminant analysis (LDA) with shrinkage of the covariance matrix (see *e.g.* [2]). As features for classification, the averages of the raw signals within 3 to 6 semi-automatically chosen time intervals from 57 channels were used. Both spellers were implemented in Python using the open-source-framework Pyff [15] with VisionEgg [11].

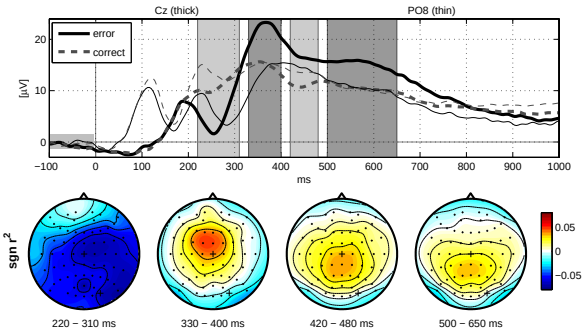


Figure 2 – Grand-average of the event-related potentials for error and non-error conditions in the ErrP Calibration Speller and in the Center Speller experiment. The top panel shows the raw voltage trace at electrode Cz (thick lines) and PO8 (thin lines) for the error (solid lines) and non-error (dashed lines) condition. The lower panel shows the scalp topographies of the point biserial correlation coefficient ($sgn\ r^2$ values), averaged in the gray intervals. The channels Cz and PO8 are indicated as black crosses.

3 Results

Error-Related Potentials (ErrPs) The error condition (solid lines) in the grand-average of the Center Speller (Figure 2) clearly deviates from the non-error condition (dashed lines). In the $sgn\, r^2$ values (scalp maps) this is reflected by a negativity 200–350 ms after onset of the feedback, referred to as error negativity (N_E). In contrast to the fronto-central distribution of N_E reported in previous studies [1, 5], we observed it over the whole scalp, with peaks over right occipital regions. The N_E was followed by a large error positivity (P_E) 400–800 ms after feedback onset. It appeared first over fronto-central regions, later over centro-parietal regions, which is in accordance with previous studies.

ErrP Detection Performance of the ErrP classification was measured in terms of the area under the ROC curve (AUC). For all participants, AUC was above 0.6 for classifier A and above 0.75 for classifier B. The performance of classifier B with a mean of 0.89 AUC was significantly higher than that of classifier A with a mean of 0.78 AUC ($t = 3.7$, $p < 0.01$).

Spelling Speed Improvement The left panel of Figure 3 compares the spelling speed of the three conditions. The average spelling speed is the highest for classifier B with 2.2 char/min. This is significantly higher than for classifier A (1.9 char/min, $t = 2.8$, $p < 0.05$) and higher than without ErrP detection (1.7 char/min, $t = 2.7$, $p < 0.05$). For classifier B ErrP-detection yields a 29.4 % mean increase in the spelling speed. Using classifier A yields no significantly higher spelling speed compared to the condition without ErrP detection ($t = 1.3$, $p = 0.23$).

The increase in spelling speed due to ErrP detection varies with the spelling accuracy (right panel of Figure 3). If many errors occur during spelling, more errors can be detected. The chance of increasing the spelling speed is thus larger than in cases where only few errors occur. Furthermore, false alarms reduce the spelling speed and in cases with high spelling accuracy, the loss in speed due to false alarms can exceed the improvement in speed due to hits. Not surprisingly, spelling speed is low when spelling accuracy is low, and vice versa. Using ErrP detection, however, the effect of spelling accuracy on speed is attenuated, although at high accuracies, false alarms outweigh the hits, so that fastest spelling is obtained without ErrP detection.

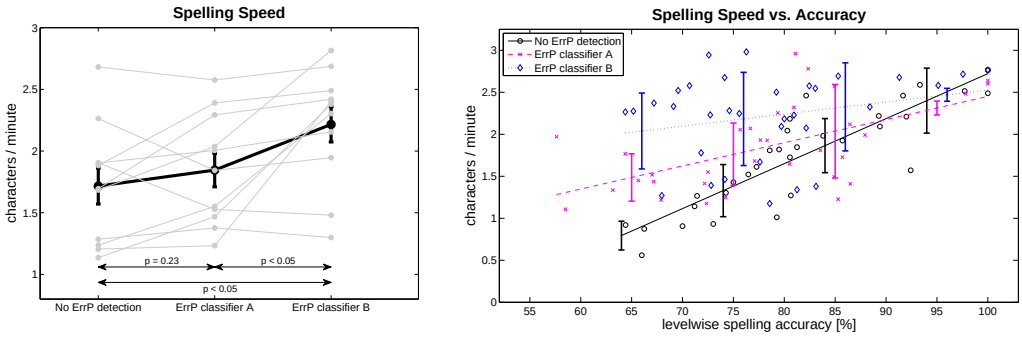


Figure 3 – Left: Spelling Speed of each participant in the three conditions: without ErrP detection, with ErrP classifier A and with ErrP classifier B. The gray error bars give the mean speed and standard error for each condition. The double arrows indicate the three paired-samples t-tests with the respective p values. Right: Scatter plot of the spelling speed as a function of levelwise selection accuracy. Each data point refers to one experimental block of one participant.

4 Discussion

Single-trial ErrPs were detected online with a mean accuracy of 79 %. The online detection rate is similar to the cross-validation results in offline studies, where 82 % [7] and 80 % [3] have been reported. ErrP detection using a classifier trained on the online data increased the mean spelling

speed by 29.4 % compared to the case without ErrP detection. A comparable improvement was obtained by [10] with ErrP detection in Matrix Speller experiments (29.5 % increase of the bit rate). This illustrates that ERP spellers can be enhanced significantly by detecting and vetoing erroneous decisions of the BCI based on error potentials.

Furthermore, the gain in communication speed is higher for participants with a medium or low BCI performance (say, > 10 % errors) than for participants with a high BCI performance. This is due to false alarms, which prolong the spelling and outweigh the improvement due to error detection in participants that have only few errors. This shows that the balance between hits and false alarms should be adjusted depending on BCI performance. The trade-off that maximizes communication speed is a function of spelling accuracy.

A drawback of using ErrPs is the fact that one has to collect a substantial number of error trials in order to train the ErrP classifier. Two possibilities were shown in the present study. One is to perform spelling first without error detection and then label the collected data to train a classifier (just as we did for classifier B). Alternatively, one might design an experiment with similar design, where many error trials are artificially generated in short time (as classifier A in the present study). Our data show that an ErrP classifier trained in one paradigm can transfer to a similar paradigm, albeit with a reduced performance.

Concluding, we demonstrated a statistically significant increase of communication speed of gaze-independent ERP spellers using error potentials. Furthermore, we found that this increase is higher for participants with low spelling performance than for participants with high spelling performance. Consequently, error detection might prove beneficial in clinical application, because patients' performance is usually lower than the performance of healthy participants.

References

[1] B. Blankertz, G. Dornhege, C. Schäfer, R. Krepki, J. Kohlmorgen, K. R. Müller, V. Kunzmann, F. Losch, and G. Curio. Boosting bit rates and error detection for the classification of fast-paced motor commands based on single-trial EEG analysis. *IEEE Trans Neural Syst Rehabil Eng*, 11(2):127–131, 2003.

[2] B. Blankertz, S. Lemm, M. S. Treder, S. Haufe, and K. R. Müller. Single-trial analysis and classification of ERP components – a tutorial. *Neuroimage*, 56:814–825, 2011. URL http://dx.doi.org/10.1016/j.neuroimage.2010.06.048.

[3] B. Dal Seno, M. Matteucci, and L. Mainardi. A genetic algorithm for automatic feature extraction in P300 detection. In *Proceedings of the International Joint Conference on Neural Networks (IJCNN '08)*, pages 3145–3152, 2008.

[4] B. Dal Seno, M. Matteucci, and L. Mainardi. Online detection of P300 and error potentials in a BCI speller. *Intell. Neuroscience*, 2010:11:1–11:1, January 2010. ISSN 1687-5265.

[5] M. Falkenstein, J. Hoormann, S. Christ, and J. Hohnsbein. ERP components on reaction errors and their functional significance: a tutorial. *Biol Psychol*, 51(2-3):87–107, 2000.

[6] L. Farwell and E. Donchin. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalogr Clin Neurophysiol*, 70:510–523, 1988.

[7] P. Ferrez and J. Millán. Error-related EEG potentials generated during simulated brain-computer interaction. *IEEE Trans Biomed Eng*, 55:923–929, 2008.

[8] C. Holroyd and M. Coles. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol Rev*, 109:679–709, Oct 2002.

[9] G. Schalk, J. R. Wolpaw, D. J. McFarland, and G. Pfurtscheller. EEG-based communication: presence of an error potential. *Clin Neurophysiol*, 111:2138–2144, 2000.

[10] M. Spüler. Online Erkennung von Fehlerpotentialen zur Fehlerkorrektur bei einem P300 Brain-Computer Interface. Eberhard Karls Universität Tübingen, Feb. 2010. Diploma thesis.

[11] A. D. Straw. Vision Egg: an open-source library for realtime visual stimulus generation. *Front Neuroinformatics*, 2: 4, 2008.

[12] M. S. Treder, N. M. Schmidt, and B. Blankertz. Towards gaze-independent visual brain-computer interfaces. In *Front Comput Neurosci*, 2010. Conference Abstract: Bernstein Conference on Computational Neuroscience 2010.

[13] M. S. Treder, N. M. Schmidt, and B. Blankertz. Gaze-independent visual brain-computer interfaces based on covert attention and feature attention. *J Neural Eng*, 2010. submitted.

[14] H. T. van Schie, R. B. Mars, M. G. Coles, and H. Bekkering. Modulation of activity in medial frontal and motor cortices during error observation. *Nat Neurosci*, 7(5):549–554, 2004.

[15] B. Venthur, S. Scholler, J. Williamson, S. Dähne, M. S. Treder, M. T. Kramarek, K. R. Müller, and B. Blankertz. Pyff – a Pythonic framework for feedback applications and stimulus presentation in neuroscience. *Front Neuroscience*, 4: 179, 2010.

Validation of an Asynchronous P300-Based BCI with Potential End Users to Control a Virtual Environment

F. Aloise[1,2], F. Schettini[1,2], P. Aricò[1,2], S. Salinari[2], C. Guger[3], J. Rinsma[4], M. Aiello[5], D. Mattia[1], F. Cincotti[1]

[1]Neuroelectrical Imaging and BCI Lab, Fondazione Santa Lucia IRCCS, Rome, Italy
[2]Dept. of Computer Science, University. of Rome "Sapienza", Rome, Italy
[3]g.tec Guger Technologies OEG, Austria
[4]Thuiszorg Het Friese Land, The Netherlands
[5]Distributed Systems Group University of Groningen, The Netherlands

f.aloise@hsantalucia.it

Abstract

Brain computer interface technology combined with accessible user interfaces for domotic appliances can return independence to individuals who (partially) lost motor abilities increasing their independence and quality of life. In this study, an asynchronous P300-based BCI, which was recently validated with young and healthy users, was tested to operate a virtual environment involving 7 potential end users. Results showed that the asynchronous system maintained its reliability with potential end users. In fact, the asynchronous system was strong in avoiding false positives when subjects were not attending to the stimulation (0.225 false positives/min). Furthermore it exhibited a low error rate ($< 2\,\%$) because was able to abstain from selection when the EEG pattern was not reliable enough.

1 Introduction

Non-invasive Brain-Computer Interface (BCI) systems aim at providing with an additional channel to restore communication and interaction with the external world people suffering from severe motor disability. Combining a BCI channel with home automation systems can be a way to augment the independency in the environmental interaction of disabled persons thus, reducing the assistance of a second person (family member or caregiver). The opportunity to control simple devices such as TV, air conditioning and/or to perform common daily actions such as to open a door, to switch on/off the light may have a remarkable impact on the quality of life. One of the EEG feature widely used to operate a BCI system is the P300 potential [1]. P300 is a positive detection (around $10\,\mu$V) recorded over the scalp central-parietal region that occurs 250–400 ms after the subject recognizes a rare or relevant sensory stimulus (Target) within a train of frequent stimuli (Non-Target) [2]. The P300-based BCIs have been already proved as suitable to control "smart" homes both under real and virtual environment [3, 4]. Typically the P300-based BCI paradigm operates in a synchronous mode, by assuming that the user is constantly controlling the BCI system and its application. Also, under this BCI paradigm a pre-defined number of stimuli is provided in order to recognize the Target stimulus. A substantial increase in system's usability and reliability would rely on their capability of dynamically adapting to the actual intent of the subject to send a command and possibly to adjust their speed to the subject's level of attention. As such, these system's features would prevent possible "misclassification" occurring when the user diverts his/her attention from the interface. The issues have been addressed in a previous work [3], in which an asynchronous P300-based-BCI system to actuate domotic appliances in a real environment was implemented and successfully tested with healthy participants. Here we

present the initial findings on the feasibility of such asynchronous system operated by potential end-users to master domotic appliances in a virtual apartment.

2 Methods

Seven elderly (4 males, 3 females; mean age = 64.85±5.81 years) individuals who are clients of the Frisian home care organization (THFL, The Netherlands) joined the recording protocol. Variable degree of motor impairment was present in four of them due to different neurological disorders such as amyotrophic lateral sclerosis (ALS; 1 patient), multiple sclerosis (MS; 2 patients) and chronic stroke (1 patient). Gross cognitive functions were preserved in all of them as indicated by the MMSE scores (between 27–30) [5].
The virtual apartment was composed of four rooms: two bedrooms, a kitchen and a living room and the devices operable by means of the BCI were the lights, doors, curtains, windows, bed, TV, air conditioning, alarm and pause applications. A beamer was used to display the apartment with a bird-eye's view and a laptop was used to run the BCI software.
The acquisition protocol was based on the Speller paradigm adapted for environmental control: a 4 by 4 matrix contained 16 icons which represented the available devices in the virtual apartment. The icons were flashing in rows and columns and a fixation cross was placed in the middle of the interface. Each stimulus was intensified for 125 ms with an inter stimulus interval (ISI) of 125 ms. When the BCI system recognized a Target (desired icon), a message was sent to an external application that parsed the information and it generated a call request to the virtual environment by using the XML based SOAP protocol. Finally, the software provided the user with a feedback through the animation of the selected device in the virtual environment. Scalp EEG potentials were recorded from 16 positions according to the 10-10 standard (Fz, FCz, Cz, CPz, Pz, Oz, F3, F4, C3, C4, CP3, CP4, P3, P4, PO7, PO8) with g.LADYbird active ring electrode using the g.USBamp amplifier (256 Hz, g.tec medical engineering GmbH, Austria). Each channel was referenced to the left earlobe and grounded to the right mastoid. Each subject underwent one recording session which was composed by a total of 9 runs: 6 runs were operated in synchronous and 3 in asynchronous modality, respectively.
The synchronous modality consisted of 4 Copy mode and 2 No-Control runs. Each run consisted of five trials; each trial was composed by 12 sequences, where a sequence consisted of one complete cycle of four rows and four columns flashes. During each Copy mode run, the subjects was asked to operate five different devices contained in the virtual environment that were cued (Target icon) at the beginning of each trial. The first two runs were used to extract control parameters through a Stepwise Linear Discriminant Analysis (SWLDA) and no feedback was provided to the subject. From the third run on, the user was provided with a feedback at the end of the trial by intensifying the selected icon and by changing the actual state of the corresponding device in the virtual environment. During the two No-Control runs, the subjects voluntarily diverted their attention from the running stimulation (flashing icons) either by concurrently gazing the fixation cross placed in the middle of the interface (No-Control Task 1) or by gazing the fixation cross and talking with the operator or writing on a text-to-speech communication aid if he/she was not able to verbally communicate (No-Control Task 2). Gazing the fixation cross surrounded by random stimuli made the asynchronous classifier stronger to misclassifications due to visual stimuli that can affect the subject while talking is a simple way to introduce muscular artifacts in the recorded EEG. Finally, both the four copy mode and the two No-Control runs were also used to extract control features and to set the thresholds for the asynchronous classifier. This latter procedure relied on a ROC curve plotting of score values as reported in Aloise et al. [3]. The asynchronous part consisted of one Control and two No-Control runs. During the first Control run, the subjects were asked to operate a total of five different devices of the environment within 10 minutes, (uncompleted run was considered as unsuccessful). In order to estimate the online performance under the asynchronous modality, the operator suggested the icon which the user

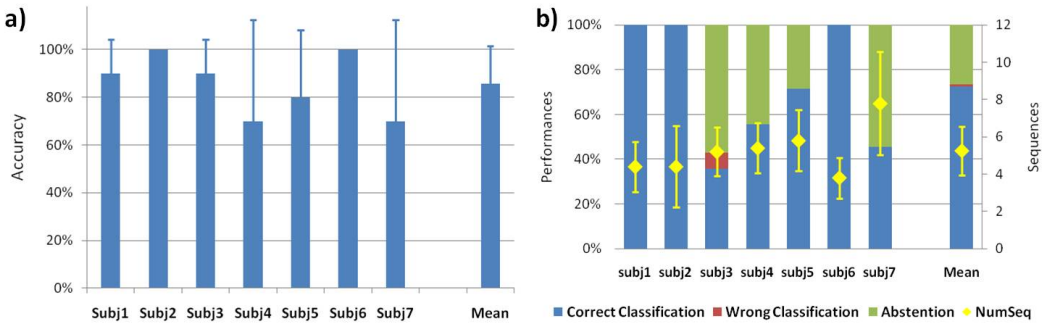


Figure 1: a) Accuracy achieved by each subject during synchronous Control runs; the error bars express the standard deviation over the two runs b) Online results obtained with the asynchronous mode; The yellow dots indicate the mean value of the number of stimulation sequences needed to obtain a classification.

had to focus on before the trial begun. In case of an unwanted abstention or a misclassification occurred the subject was invited to try again the selection of the desired (suggested) icon. The two No-Control runs (lasting 2 and half minute each), consisted of the above described two No-Control tasks and had the aim of quantifying the robustness of the asynchronous system with respect to false positives.

3 Results

The Figures 1 a) and 1 b) illustrate the online results obtained under synchronous and asynchronous modality, respectively. As for the synchronous mode, all subjects achieved at least a mean accuracy of 70 % with an overall mean of accuracy of 85 % ($n = 8$ subjects). Under the asynchronous mode, 3 subjects achieved 100 % correct classifications although the overall mean of correct classification was 73 % ($n = 8$ subjects). The observed difference between the correct classification mean value relative to the asynchronous mode (73 %) and the mean accuracy obtained under the synchronous modality (85 %) did not reach statistical significance as revealed by the paired t-test ($t = 1.49$, $p - value = 0.15$). The asynchronous system was reliable in avoiding misclassifications. Indeed, misclassifications occurred only with the subject 3 (7 % of errors). Moreover, the percentage of abstention was equal to 26 % on average but it should be noted that during the synchronous runs 15 % of wrong classifications on average occurred. Finally, during the 2 No-Control runs of the asynchronous modality, 0.225 false positives/minute were detected on average, thus indicating that the system exhibited a "strong" reliability in avoiding false positives when the subjects were challenged with the other concurrent tasks (Figure 1 b), yellow dots).

4 Discussion

One of the most important features of an asynchronous system is the capability to estimate when the user intends to exercise his control on through the system from the ongoing EEG, avoiding misclassifications when the user focuses his attention elsewhere. From this point of view the proposed asynchronous system proved very reliable also with potential end users, exhibiting on average 0.225 false positives/min. Regarding its accuracy recognition the asynchronous system was able to adapt its speed to user state. The performances far exceed the chance level [6] providing a

valuable mean for communication and control [7]. The actual findings reinforce the usability and reliability of an asynchronous BCI system for environmental control, decreasing the gap between BCI systems and classical input devices.

5 Conclusion

An asynchronous P300-based BCI system demonstrated advantages with respect to a synchronous one, in terms of usability and efficiency, also involving potential end users. In fact it is able toautomatically suspend the control when the user divert his/her attention from the stimulation, it can avoid misclassifications through the mechanism of abstentions, and finally it continuously adapts the time required for each classification to the changes of user state.

6 Acknowledgments

The work is partly supported by the EU grant FP7-224332 "SM4ALL" (Smart homes for All anembedded middleware platform for pervasive and immersive environments for all) project. 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. We also would to thank Francesco Leotta, Eirini Kaldeli and Alexander Lazovik for the developing part. Dinie Sieswerda, Willy van der Brink, Allard Naberand Ehsan Warriach for their precious support during the experimentation. Finally a particular thanks goes to all the subjects that participated in this study.

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(6):510–523, 1988.

[2] J. Polich. P300 clinical utility and control of variability. *Journal of Clinical Neurophysiology: Official Publication of the American Electroencephalographic Society*, 15(1):14–33, 1998.

[3] F. Aloise, F. Schettini, P. Aricò, F. Leotta, S. Salinari, D. Mattia, F. Babiloni, and F. Cincotti. P300-based brain-computer interface for environmental control: an asynchronous approach. *Journal of Neural Engineering*, 8(2):025025, 2011.

[4] C. Guger, C. Holzner, C. Grönegress, G. Edlinger, and M. Slater. Control of a smart home with a brain-computer interface. In *4th BCI Workshop*, Graz, 2008.

[5] M. F. Folstein, S. E. Folstein, and P. R. McHugh. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12(3):189–198, 1975.

[6] 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.

[7] A. Kübler, A. Furdea, S. Halder, E. M. Hammer, F. Nijboer, and B. Kotchoubey. A brain-computer interface controlled auditory event-related potential (P300) spelling system for locked-in patients. *Annals of the New York Academy of Sciences*, 1157:90–100, 2009.

Design of a Novel Covert SSVEP-Based BCI

D. Lesenfants[1], N. Partoune[1], A. Soddu[1], R. Lehembre[1], G. Müller-Putz[2], S. Laureys[1], Q. Noirhomme[1]

[1]Coma Science Group, Cyclotron Research Centre and Neurology department, University of Liège, Liège, Belgium

[2]Laboratory of Brain–Computer Interfaces, Institute for Computer Graphics and Vision, Graz University of Technology, Graz, Austria

Damien.Lesenfants@student.ulg.ac.be

Abstract

Brain computer interfaces (BCI) employing steady-state visually evoked potential (SSVEP) modulations have been investigated increasingly in the last years because of their high signal-to-noise ratio and information transfer rate. However, independent SSVEP BCI based on covert attention show a drop in robustness which makes it difficult to use on patients with impaired or nonexistent ocular motor control. In the present paper, offline analysis is aimed at investigating the influence of three important parameters on the performance of covert SSVEP BCI: feature extraction algorithms, window length and number of harmonics. We also proposed a new "checkerboard" pattern and compared its performance with lines pattern. We have shown that the use of this pattern and only one harmonic yielded an average accuracy of approximately 79 % across five subjects (with four subjects at more than 81 %) with 6 s-window length and feature extraction algorithm based on canonical correlation analysis or lock-in analyzer system. The short 5 or 6 s-concentration time, the absence of training due to the use of only one harmonic, the robustness make this method very well suited for detecting command following and testing communication in unresponsive post-comatose patients.

1 Introduction

Current brain-computer interfaces (BCI) [1] relying on steady-state visually evoked potentials (SSVEP), while demonstrating high information transfer rates and considerable robustness, depend on gaze control [2, 3]. This rules out applicability to those whose severe disabilities extend to impaired or nonexistent ocular motor control and prevents the use as a diagnostic tool for post-comatose patients. Independent SSVEP BCI based on covert attention have been proposed but have shown a drop in robustness. First covert SSVEP BCI were based on block pattern [4]. While it works on healthy controls who can move their gaze away of both pattern, it is not suitable for patients without gaze control. They could unintentionally move their gaze on the wrong stimulation or move their gaze away of both patterns preventing detection of their wishes. Other stimulation patterns take the advantage of the ability to covertly attend to one of two overlapping stimuli by presenting either mixed vertical and horizontal lines or moving dots. We here propose a new covert stimulation pattern which enables a better discrimination between two stimuli. In the following, we will compare the results obtained with this new pattern with the results of a lines pattern. We will also study the influence of different feature extraction algorithms, the duration of the stimulation and the number of harmonics. Use of higher harmonics has been shown to positively influence classification in overt SSVEP [2] but has never been tested in covert SSVEP.

2 Materials and Methods

Five healthy subjects (two males) aged between 19 and 43 years old participated in the study. All had normal or corrected-to-normal vision. Subjects were seated about 1 m from a 17"CRT monitor with a refresh rate of 85 Hz on which was displayed three different stimulation patterns. Each pattern contained a red and a yellow stimulus flashing at $f_1 = \frac{85}{8} \approx 10.625$ Hz and $f_2 = \frac{85}{6} \approx 14.167$ Hz respectively. First, a block pattern [4, 5] with two 8 by 8 cm squares, one of each color, separated by 12 cm with a white fixation cross in between (Figure 1, first row). Second, a lines pattern [6, 7] with 7 vertical yellow 13 by 1 cm strips separated from each other by 1 cm, and 7 horizontal red strips with a white fixation square in the middle (Figure 1, second row). Third, a "checkerboard" pattern where all overlapping segments of the lines pattern were removed. This pattern is a 13 by 13 cm "checkerboard" made of red, yellow and black 1 by 1 cm squares with a white fixation square in the middle (Figure 1, third row). Subjects were instructed on which stimulus to attend by either an arrow indicating the left or the right stimulus (block pattern) or a matched color strips surrounding the pattern (Figure 1, last column). EEG signals were recorded from 12 Ag/AgCl rings electrodes at location P_3, P_1, P_2, P_4, PO_7, PO_3, PO_z, PO_4, PO_8, O_1, O_z and O_2, referenced to P_z based on the international 10-20 electrode system, with a BrainVision V-Amp amplifier (band pass filter between 0.01 and 250 Hz with a sampling frequency of 1 kHz). Eye movements were monitored with four electrodes: two on the left and right temples; the other two over and under the supra-orbital ridge.

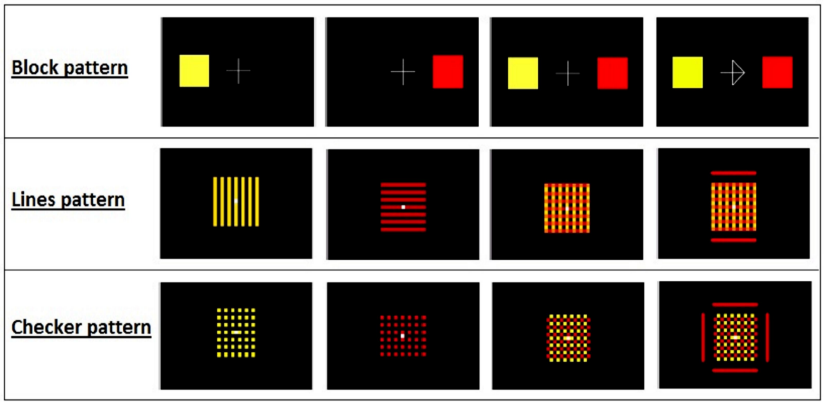


Figure 1: The different patterns used. Columns 1 to 3 shows the two different stimuli separately and together, and column 4 illustrates how a stimulus is indicated to the subject.

Each subject underwent a total of 5 runs, each lasting around 6 min. Each run contained 30 7 s-trials separated by 5 s rest period. One of the three patterns was continuously flashing at the screen during a whole run. During a run, an equal number of both stimuli was presented in random order. The first run consisted of an overt block pattern where the subject was instructed to direct his/her gaze to the stimulus. This run ensured subject response to SSVEP stimuli. Next four runs alternated between covert lines and checkerboard pattern. The subject was instructed to fix his/her gaze on the white square in the middle and to concentrate on one of the stimulus. Half of the subjects started with the lines pattern. Instruction stayed on the screen for the whole trial duration. Inter-run periods were at the discretion of the subject and lasted 2 to 5 minutes. Runs from the same pattern were concatenated before analysis and trials order was randomized to remove all time effects.
For each trial, the signal at EEG channels were individually decomposed in overlapping windows of length T with T ranging from 1 to 7 s: seven 1 s-windows starting from 0 to 6 s, six 2 s-windows starting from 0 to 5 s, ..., one 7 s-window starting at time 0. Frequency features were extracted from each window with three feature extraction algorithms proposed in the literature: discrete-time

Fourier transform (DFT), canonical correlation analysis (CCA) [8, 9] and lock-in analyzer system (LAS) [3]. First and second harmonics of each stimulation frequency were extracted. Features set included either both harmonics or just the first. Classification performance were computed with a linear discriminant analysis and assessed with a 10-folds cross validation. The cross validation was repeated 5 times, then the results were averaged first for each time window, second for all subjects and time window of the same length. Analysis was done with Matlab and Cogent 2000 toolbox developed by the Cogent 2000 team at the FIL and the ICN and Cogent Graphics developed by John Romaya at the LON at the Wellcome Department of Imaging Neuroscience.

3 Results

Overt block pattern data from 2 s-window could be classified with more than 99 % of accuracy for all subjects. Visual inspection of EOG showed no eye movement during the covert trials. The checkerboard pattern outperformed the lines pattern for every feature extraction algorithms and time windows (Figure 2, right). Classification accuracy increased with window length until reaching a plateau at 5 s for CCA and 6 s for LAS. DFT accuracy increased constantly. CCA has better accuracy than LAS and DFT for window smaller than 5 s. CCA accuracy was 70±3 % and 75±3 % at 2 and 4 s respectively. For 6 s and 7 s windows, LAS gave the best results reaching 79±7 % in average with 4 subjects at more than 81 % (Figure 2, left). Using the first two harmonics instead of only the first one decreased the classification (Figure 2, left).

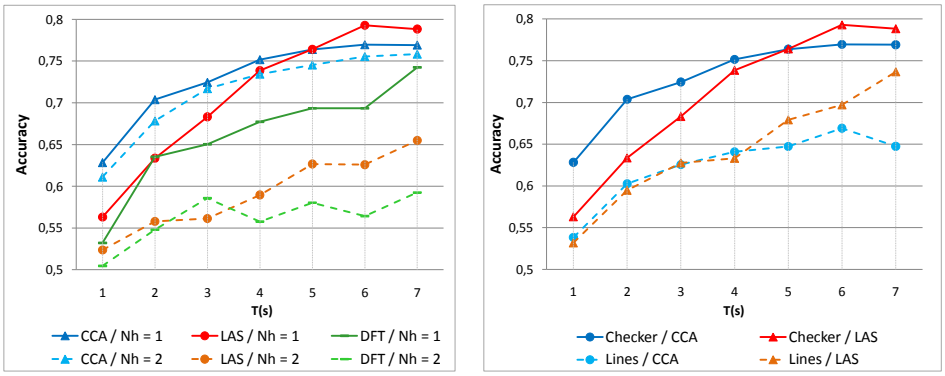


Figure 2: Relationship between recognition accuracy with the epoch length T for different number of harmonic N_h and feature extraction algorithms (left) and different patterns with a single harmonic (right).

4 Conclusion

In the present paper, we studied the design of independent BCI based on covert SSVEP and showed that we can get around 80 % accuracy. Due to the small number of subjects, all these results should be seen as preliminary. We proposed a new stimulation pattern which increased the classification of nearly 10 % in comparison to a lines pattern [6, 7] and seems to do better than moving dots pattern [9] but with different conditions. The new pattern has no overlap between stimuli which facilitates the concentration as reported by 4 out of 5 subjects; the last one being neutral. We also showed that CCA is the most suitable method for features extraction with window smaller or equal to 5 s while LAS may be preferred for longer time window. These two algorithms gave better results than the classical DFT. Surprisingly and contrary to results from overt SSVEP [2], the use of the second harmonic leads to a decrease of classification accuracy. This could be due to over training of the classifier or to the use of a CRT screen instead of LEDs. However, a power spectrum analysis did not show any peak besides the ones of the first harmonics. No influence of higher harmonics has also been reported in overt SSVEP analysis with CCA [8].

These results acquired with only the first harmonic and advanced frequency feature extraction algorithms demonstrate that SSVEP BCI could be used independently of gaze control but also independently of training. Indeed, using the first harmonic and the CCA method give only two features for each stimuli. The maximum of the features gives the user command [8].
The short concentration time, 5 s or 6 s, the absence of training, the robustness make this method very well suited for detecting command following and testing communication in unresponsive post-comatose patients. Futures studies will concentrate on implementing an online version of the BCI and on testing on more subjects. The use of a restricted set of electrodes will also be tested.

5 Acknowledgments

Steven Laureys is Senior Research Associate; Andrea Soddu and Quentin Noirhomme are Post-doctoral Fellows at the Fonds de la Recherche Scientifique (FRS). This work is supported by the European ICT Programme Projects FP7-247919 DECODER, DISCOS, Mindbridge, COST, McDonnell Foundation, Mind Science Foundation, FRS, Reine Elisabeth Medical Foundation and University and University Hospital of Liège. 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] J. R. Wolpaw, N. Birbaumer, D. J. McFarlandm, G. Pfurtscheller, and T. M. Vaughan. Brain-computer interfaces for communication and control. *Clinical Neurophysiology*, 113:767–791, 2002.

[2] G. R. Müller-Putz, R. Schere, C. Brauneis, and G. Pfurtscheller. Steady-state visual evoked potential (SSEVP)-based communication: impact of harmonic frequency components. *Journal of Neural Engineering*, 2:123–130, 2005.

[3] G. R. Müller-Putz, E. Eder, S. C. Wriessnegger, and G. Pfurtscheller. Comparison of DFT and lock-in amplifier features and search for optimal electrode positions in SSVEP-based BCI. *Journal of Neuroscience Methods*, 168:174–181, 2008.

[4] S. P. Kelly, E. C. Lalor, B. Reilly, and J. J. Foxe. Visual spatial attention tracking using high–density ssvep data for independent brain–computer communication. *IEEE Transactions On Neural Systems And Rehabilitation Engineering*, 13(2):172–178, June 2005.

[5] S. P. Kelly, E. C. Lalor, C. Finucane, G. McDarby, and B. Reilly. Visual spatial attention control in an independent brain-computer interface. *IEEE Transactions on Biomedical Engineering*, 52(9):1588–1596, September 2005.

[6] Y. Chen, A. K. Seth, J. A. Gally, and G. M. Edelman. The power of human brain magnetoencephalographic signals can be modulated up or down by changes in an attentive visual task. *Proceedings of the National Academy of Sciences*, 6:3501–3506, March 2003.

[7] Z. Allison, D. J. McFarland, G. Schalk, S. D. Zheng, M. Moore Jackson, and J. R. Wolpaw. Towards an independent brain–computer interface using steady state visual evoked potentials. *Clinical Neurophysiology*, 119:399–408, 2008.

[8] G. Bin, X. Gao, Z. Yan, B. Hong, and S. Gao. An online multi-channel SSVEP-based brain-computer interface using a canonical correlation analysis method. *Journal of Neural Engineering*, 6(4):46002–46008, June 2009.

[9] D. Zhang, A. Maye, X. Gao, B. Hong, A. K. Engel, and S. Gao. An independent brain-computer interface using covert non-spatial visual selective attention. *Journal of Neural Engineering*, 7(1):16010–16021, 2010.

SSVEP BCI Control of a Hand Orthosis: First Evaluation in Persons with Tetraplegia

D. S. Klobassa[1,*], T. Solis-Escalante[1,*], H. Gaggl[2], G. Korisek[2], G. Pfurtscheller[1], C. Neuper[1]

[1]Institute for Knowledge Discovery, Laboratory of Brain-Computer Interfaces, Graz University of Technology, Graz, Austria
[2]Rehabilitationsklinik Tobelbad, Tobelbad, Austria
[*]These authors contributed equally to this work

daniela.klobassa@tugraz.at
teodoro.solisescalante@tugraz.at

Abstract

In this work, we report on the first evaluation of a brain-computer interface (BCI) based on the steady state visual evoked potentials (SSVEP) for control of a hand orthosis, to assist grasping in persons with tetraplegia. Two persons with spinal cord injury operated the BCI to grasp, move and release a plastic bottle. Operation of the BCI was self-paced (asynchronous); meaning that the participants controlled the orthosis at their own pace, to complete as many grasp-move-release sequences as possible in a given time interval. Participants completed 16 and 15 runs of such a task, respectively, over a number of sessions (four and five) in separate days. Performance was estimated with the positive predictive value (PPV). Average PPV values were 70 % ± 14 % and 71 % ± 12 %. Previous results with healthy participants led to an average PPV of 76.5 % ± 15 %.

1 Introduction

Tetraplegia is the partial or total paralysis of all four limbs, caused by damage to the spinal cord at cervical level. Depending on the severity of the lesion and the level where it occurs, sensation and control are reduced or lost. Generally, persons with tetraplegia have an impaired hand function that affects grasping. Previous research has aimed to restore hand function by electroencephalogram (EEG)-based neuroprostheses [1–3] and hand orthoses [4, 5]. Such brain-computer interface (BCI) systems allow the control of prostheses or orthoses by modulations of the sensorimotor rhythms [1–4], or the steady state visual evoked potentials (SSVEP) [5]. Recently, Ortner et al. [5] demonstrated an SSVEP BCI system for control of a hand orthosis operated by healthy users. Seven naive participants were able to operate the SSVEP BCI to move a hand orthosis in a self-paced (asynchronous) mode. In this study our aim is to evaluate the control of a hand orthosis with two tetraplegic users operating this SSVEP BCI.

2 Methods

2.1 Hand Orthosis

Two electrically driven, reciprocal wrist-extension, finger-flexion hand orthoses [6] were used in this study. The first hand orthosis (HO1) can be adapted to different shapes/sizes of the human hand by adjusting its individual parts (see [5] for details). The second hand orthosis (HO2) is a simpler, lighter orthosis that must be constructed to match the hand size. Both orthoses have a

pair of LEDs attached that deliver the visual stimulation generated in an external unit. In HO1, the distance of the LEDs varies from finger extension (12.5 cm) to finger flexion (9 cm); whereas in HO2, this distance remains constant (15.5 cm). Besides these differences, HO1 and HO2 have the same functionality.

2.2 SSVEP Classification

SSVEPs are detected and classified from the ongoing EEG using the harmonic sum decision (HSD) [7]. EEG signals were recorded with a biosignal amplifier (g.USBamp, Guger Technologies, Schiedelberg, Austria) at a sampling rate of 256 Hz, with digital filters set between 0.1 and 100 Hz in combination with a notch filter (50 Hz). Power spectral density is estimated (zero padded 512-point discrete Fourier transform) every 250 ms from the previous time interval of one second. A class label is computed according to

$$\text{label} = \arg\max \left(\sum_{i=1}^{3} H_i f_1 \cdot BLf_1^{-1}, \sum_{i=1}^{3} H_i f_2 \cdot BLf_2^{-1} \right) \tag{1}$$

where H_i is the i-th harmonic, and BLf_j is the normalization factor of frequency f_j (the average harmonic sum from a baseline recording). In our experiments, $f_1 = 8$ Hz and $f_2 = 13$ Hz, were used to open and close the orthosis, respectively. A control command was generated every time that a certain class label was continuously detected during a dwell time of 1.5 s. After each command a refractory period of 4 s was added; in such period control of the orthosis was disabled. Consecutive commands could occur every 5.5 s, fastest.

2.3 Participants and Data Acquisition

Participant S1 is a 36-year-old woman with tetraplegia since April 1994 (17 years). Her lesion is subtotal below C6, with spastic partial paralysis of the upper extremities and torso, and spastic total paralysis of the lower extremities. Participant S2 is a 29-year-old man with tetraplegia since September 2010 (6 months). He has an incomplete lesion sub C4 and a complete lesion below C5, with partial paralysis of the upper extremities and total paralysis of the lower extremities. Both participants were recruited for this study while attending rehabilitation therapy at the Rehabilitationsklinik Tobelbad, Austria. Participant S1 attended a follow-up rehabilitation training, whereas participant S2 attended primary rehabilitation training.

A pair of Ag-AgCl electrodes were placed at 2.5 cm above and below the electrode positions O1, impedance was kept below 5 kΩ. Recordings were monopolar, with reference and ground electrodes located at the left and right mastoid, respectively. A bipolar derivation was computed online from the pair of electrodes.

2.4 Experimental Paradigm

During the experiments, participants were seated on their own wheelchair in front of a table, wearing one of the orthoses on the left hand, with the flickering stimulation switched on. In front of them, two blue circles (19 cm diameter with a distance of 50 cm between both centers) were fixed to the table. These circles were used as targets for the grasp-move-release task described below. An experimental session consisted of:

Baseline recording (1 min): wearing a hand orthosis (open), the participants were instructed to focus on a point in between both LEDs without attending neither of the flickering stimuli. EEG recordings from this run yielded the normalization factor BLf_i.

Grasp-move-release runs (5 min each): in these runs the participants completed the following sequence:

1. To move the orthosis to the bottle placed on the left target.

2. To grasp the plastic bottle by closing the orthosis through a BCI command (i.e., focus on the 8 Hz stimuli).
3. To move the bottle toward the target on the right.
4. To release the bottle by opening the orthosis through a BCI command (i.e., focus on the 13 Hz stimuli).

After completing a sequence (i.e., moving the bottle from left to right), the experimenter put the bottle back on the left area. The participants were instructed to repeat the sequence as many times as possible within one run.

Participant S1 completed 16/18 runs in four sessions (four runs per session), operating the orthosis HO2; and participant S2 completed 15/17 runs in five sessions (three runs per session). In the first session, S2 operated HO2 but grasping was not possible due to hand position; thus, for this session only, the task was conducted without the bottle. In the next four sessions, participant S2 operated the orthosis HO1. Prior to the experiments, participant S1 completed three sessions (seven runs in total) with HO1 to set up the SSSVEP BCI system; those runs are not included in this work.

3 Results

Table 1 presents the average performance (across runs) obtained for each experimental session. Performance is summarized by the true positives per minute (TP/min), false positives per minute (FP/min), the corresponding positive predictive value PPV = $\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}}$, and the number of grasp-move-release sequences completed. It also shows the session number and the recording day (relative to the beginning of the experiments per participant), as well as the orthosis operated in that session. Figure 1 displays, for each participant and experimental session, the PPV obtained for individual runs. Average PPV values were 70 % ± 14 % for participant S1, and 71 % ± 12 % for participant S2. In individual runs, the maximum PPV were 97 % (S1) and 86 % (S2). Furthermore, the average TP/min were 4.3 % ± 0.9 % (S1) and 4.5 % ± 1.0 % (S2), and the average FP/min were 1.9 % ± 1.0 % (S1) and 1.8 % ± 0.7 % (S2). In comparison, the average performance across seven healthy participants lead to an average PPV of 76.5 % ± 15 % (best run was 98 %), TP/min of 5.4 % ± 1.4 %, and FP/min of 1.5 % ± 0.9 % [5].

4 Discussion

Evaluation of the SSVEP BCI controlled hand orthosis with tetraplegic persons showed similar results to those obtained in previous experiments with healthy participants. S1 and S2 had their best performance shortly after the beginning (session 2 or 3). A performance decrease followed long pauses in between sessions (25 and 47 days, respectively); a plausible indication of a learning effect. In general, the performance was stable during a single session, with a small increase between the first and the last run. However, in some sessions the very first run was more accurate than the last one (S1 sessions 3 and 4, S2 sessions 1 and 3).

ID	Session (day)	TP/min	FP/min	Sequences completed	PPV (%)	Orthosis
S1	1 (1)	3.6	2.3	9.0	61	HO2
	2 (9)	3.9	2.4	9.7	64	HO2
	3 (16)	5.1	1.0	12.5	83	HO2
	4 (41)	4.7	1.7	11.5	73	HO2
S2	1 (1)	3.3	2.3	8.0	59	HO2
	2 (5)	6.0	1.0	14.7	86	HO1
	3 (52)	4.4	1.7	11.0	72	HO1
	4 (56)	4.2	2.2	10.3	66	HO1
	5 (63)	4.5	2.0	11.3	70	HO1

Table 1: Average performance for the different experimental sessions.

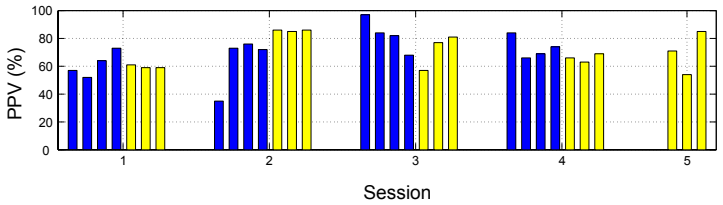


Figure 1: Individual run performance for each experimental session: (blue) participant S1 and (yellow) participant S2.

5 Conclusion

This work suggests that persons with tetraplegia can operate a SSVEP BCI orthosis with a performance similar to that of healthy persons, but due to the reduced number of participants, it is not possible to draw conclusions from these results. Further experiments involving more patients are necessary for an appropriate comparison. Both participants operated the BCI to control a hand orthosis to assist grasp of the left hand using a single EEG bipolar derivation. It could be expected that optimization of the electrode positions, EMG detection/reduction, and the use of subject-specific frequencies would lead to an improved performance. Further work will address these issues and continue the evaluation of both the SSVEP BCI and the orthoses with more persons with tetraplegia.

6 Acknowledgments

This study was supported by the Austria "Allgemeine Unfallsversicherung AUVA" and the "Lorenz-Böhler Gesellschaft". Hand orthoses HO2 was built by Pischler Wolfgang from the Rehabilitationsklinik Tobelbad, Tobelbad, Austria.

References

[1] J. M. Heasman, T. R. D. Scott, L. Kirkup, R. Y. Flynn, V. A. Vare, and C. Gschwind. Control of a hand grasp neuroprosthesis using an electroencephalogram-triggered switch: demonstration of improvements in performance using wavepacket analysis. *Medical and Biological Engineering and Computing*, 40:588–593, 2002.

[2] 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.

[3] M. Tavella, R. Leeb, R. Rupp, and J. del R. Millán. Towards natural non-invasive hand neuroprostheses for daily living. In *Proceedings of the 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 126–9, 2010.

[4] G. Pfurtscheller, C. Guger, G. Müller, G. Krausz, and C. Neuper. Brain oscillations control hand orthosis in a tetraplegic. *Neuroscience Letters*, 292:211–214, 2000.

[5] R. Ortner, B. Z. Allison, G. Korisek, H. Gaggl, and G. Pfurtscheller. An SSVEP BCI to control a hand orthosis for persons with tetraplegia. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 19:1–5, 2011.

[6] T. J. Engen. *Instructional manual for a reciprocal wrist extension-finger flexion orthosis.* Houston, TX, USA. Baylor University, 1968.

[7] G. R. Müller-Putz, R. Scherer, C. Brauneis, and G. Pfurtscheller. Steady-state visual evoked potential (SSVEP)-based communication: impact of harmonic frequency components. *Journal of Neural Engineering*, 2:1–8, 2005.

A Hybrid Brain-Computer Interface Based on P300 and M-VEP

J. Jin [1], B. Z. Allison [2], X. Wang [1], C. Neuper [2,3]

[1]Key Laboratory of Advanced Control and Optimization for Chemical Processes, Ministry of Education, East China University of Science and Technology, Shanghai, 200237 P.R. China
[2] Laboratory of Brain-Computer Interfaces, Institute for Knowledge Discovery, Graz University of Technology, 8010 Graz, Austria
[3] Department of Psychology, University of Graz, 8010 Graz, Austria

Jinjingat@gmail.com

Abstract

Hybrid brain-computer interface (BCI) systems could combine two different BCIs to improve information throughput relative to a conventional simple BCI. Here, we introduce a new hybrid BCI that combines BCIs based on P300 measures and motion-onset visual evoked potentials (M-VEPs). Subjects first participated in offline data collection to train the classification model, then we validated performance with an online BCI. All ten subjects who were tested could use the system with a high classification accuracy and ITR.

1 Introduction

Brain-computer interface (BCI) systems translate direct measures of brain activity into command and control signals [1]. In this paper, we introduce a new hybrid BCI based on P300 and M-VEP measures. Stimuli both flash and move, which could improve the classification accuracy, information transfer rate, literacy rate, or other factors relative to a P300 or M-VEP BCI. Our principal goal is to validate this approach in an online BCI with realtime feedback.

2 Methods

2.1 Stimuli

Subjects viewed a 6×6 matrix containing blue English letters and numbers against a black background (see Figure 1 A). Subjects were asked to count each time a target stimulus flashed. We use the term "flash" throughout this paper to refer to the "moving flash" phenomenon in this study. In this design, each character changed green and immediately moved 1.5 cm to the right for 50 ms at a constant speed of 3 m/s or about 20 degrees/s (see Figure 1 B, C). Next, the characters reverted to blue and immediately moved another 1.5 cm to the right for 20 ms at a constant speed of 7.5 m/s or about 50 degrees/s. Next, the characters moved 3 cm to the left (to their original positions) for 30 ms at a constant speed of 10 m/s or about 66.6 degrees/s. While we considered various ERP phenomena when designing this paradigm, feedback from pilot subjects was also a major consideration. For example, pilot subjects reported that they disliked an earlier paradigm in which the initial (green) flash and movement lasted 100 ms, because the stimuli were annoying. However, they reported less annoyance when the initial (green) flash and movement was reduced to 50 ms, and there was a 150 ms before the next flash began. While prior work used a 140 ms moving time [2], pilot subjects reported that 100 ms was adequate.

There are 12 flash patterns, which are labeled "$flash_1, flash_2, ..., flash_{12}$". We chose these flash patterns to minimize errors that could result from repetition or other factors [3]. Figure 2

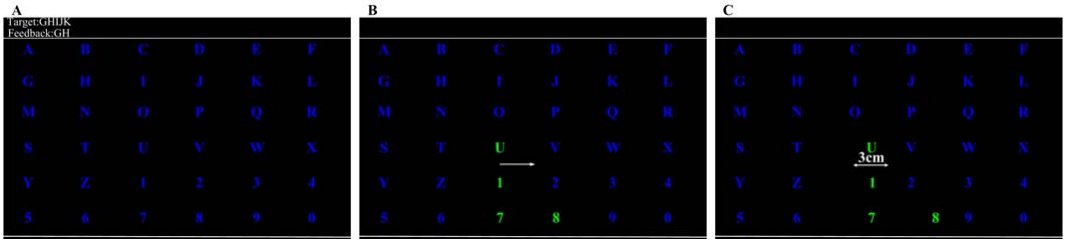


Figure 1: The display during the online runs. Panel A presents the matrix when no flashes occur. Panels B and C show four characters that flash green and shift 3 cm to the right. The white arrows and text in these panels were added to this figure.

1, 4	1, 5	1, 6	1, 7	1, 8	1, 9
2, 10	2, 5	2, 6	2, 7	2, 8	2, 9
3, 10	3, 11	3, 6	3, 7	3, 8	3, 9
4, 10	4, 11	4, 12	4, 7	4, 8	4, 9
5, 10	5, 11	5, 12	1, 10	5, 8	5, 9
6, 10	6, 11	6, 12	3, 12	2, 11	6, 9

Figure 2: The flash pattern configuration. The numbers indicate which of the 12 flashes contained the target character. For example, the top left element (the letter "A") was illuminated during the first and fourth of the 12 flashes. The grey boxes indicate characters that were illuminated during $flash_{12}$.

depicts which of the 36 characters were flashed during each of the 12 flash patterns. For example, the character "A" was flashed in $flash_1$ and $flash_4$ (see Figures 1 and 2). Stimuli were not grouped in rows or columns during each flash.

2.2 Experiment Set Up and Online Protocols

EEG signals were recorded with a g.USBamp and a g.EEGcap (Guger Technologies, Graz, Austria) with a sensitivity of 100 µV, band pass filtered between 0.1 Hz and 30 Hz, and sampled at 256 Hz. We recorded EEG from electrode positions Fz, Cz, Pz, Oz, C3, C4, P7, P3, P4, P8, O1, and O2 from the extended International 10-20 system. The right mastoid electrode was used as the reference, and the front electrode (FPz) was used as a ground.

Ten healthy subjects (6 male and 4 female, aged 21–25, mean 23) participated in this study. Six subjects had used a BCI before this study. During data acquisition, subjects were asked to relax and avoid unnecessary movement. Subjects first participated in three offline runs to obtain data to train the classifier that was used in the online system. We only report the online procedure and data here. Our journal paper [4] contains additional details concerning the offline procedure, prior references, a comparison to P300 and mVEP BCIs, scalp topographies over more sites, discussion of gaze dependence and additional future directions, etc.

We defined one trial as one flash (i.e., sub-trial) of each of the twelve flash patterns. A trial block was two or more consecutive trials with the same target character. The number of trials per trial block was variable, because the system adjusted this number to optimize performance as described in Section 2.6. Each run consisted of five trial blocks, each of which involved a different target character.

At the beginning of each run, subjects viewed the 6×6 matrix with a five-letter sequence in the "Target" line at the top of the screen. The subject was instructed to silently count each time the target character flashed. The first letter of the five-letter sequence was the target until the classifier decided on the correct target and feedback was presented (see Section 2.6). Then, that trial block ended and the next letter became the target for the next trial block. There was a one second delay between each trial, a five second delay between each trial block, and a variable delay (usually about 30 seconds) between each run. There were four online runs, which occurred on the same day as the offline runs.

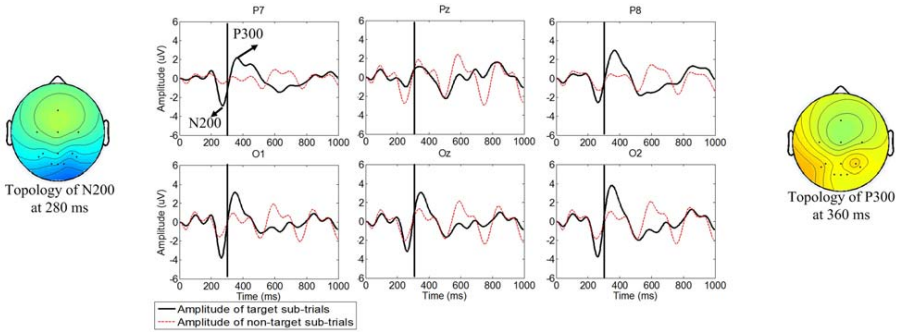


Figure 3: The grand averages of ERP amplitudes over sites P7, Pz, P8, O1, Oz and O2 averaged over all ten subjects, and topology of sites Fz, C3, Cz, C4, P7, P3, Pz, P4, P8, O1, Oz and O2 averaged over all ten subjects.

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	Average
Acc (%)	100	85	95	95	100	100	95	95	95	100	96.0 ± 4.6
RBR	38.3	28.6	27.6	31.9	35.1	42.2	38.6	33.1	33.7	47.0	35.6 ± 6.0
PBR	34.1	17.8	22.6	28.6	31.5	37.1	30.5	26.6	27.0	40.8	29.7 ± 6.8
TA	2.7	2.7	3.4	2.9	3.0	2.5	2.4	2.8	2.8	2.2	2.7 ± 0.3

Table 1: Performance from online feedback runs. In this table, "Acc" refers to classification accuracy, "RBR" is raw bit rate and "PBR" is practical bit rate, measured in bits/min. "TA" refers to the mean number of trials per average. The BCI could usually identify the target character with 2–3 flashes.

2.3 Feature Extraction Procedure

A third order Butterworth band pass filter was used to filter the EEG between 0.1 Hz and 12 Hz. Although the P300 is primarily in the band 0.1–4 Hz, it can also be found in higher bands. The EEG was downsampled from 256 Hz to 36.6 Hz by selecting every seventh sample from the filtered EEG. Single sub-trials lasting 800 ms were extracted from the data. The size of the feature vector was 12×29 (12 channels by 29 time points).

2.4 Classification Scheme

BLDA is an extension of Fisher's linear discriminant analysis (FLDA). BLDA was selected because of its demonstrated classification performance in P300 BCI and motor imagery BCIs [5], and because it avoids overfitting. Here, each trial yields 12 classifier outputs, one output for one flash. The first and second maximum classifier outputs are selected to identify the target character.

2.5 Practical Bit Rate

We used two bit rate calculation methods called practical bit rate and raw bit rate. The practical bit rate provides a more realistic estimate of realworld performance, because it incorporates the fact that every error results in a penalty of two additional selections. The practical bit rate includes the one second delay between character selections, and is calculated as RBR*(1–2*P), where RBR is the raw bit rate and P is the online error rate of the system [6]. We present both raw and practical bit rate in Table 1 to facilitate comparison to other studies.

2.6 Adaptive System Settings

The number of trials per average was automatically adjusted based on the classifier output. After each trial, the classifier would determine the target character based on data from all trials in that

trial block. If the classifier decided on the same character after two successive trials, then no new flashes were presented, and that character was presented as feedback to the subject.

3 Results

Figure 3 shows the grand averages of event related potentials' amplitude over sites P7, Pz, P8, O1, Oz and O2, and topology from sites Fz, C3, Cz, C4, P7, P3, Pz, P4, P8, O1, Oz and O2 from all ten subjects. M-VEP activity is especially prominent over P7, P8, O1, O2 and Oz. Table 1 shows the results from online feedback. It also shows that high classification accuracy and bit rate could be obtained from hybrid BCI using both P300 and M-VEP activity.

4 Conclusion

In this paper, a hybrid BCI that combined P300 and M-VEP was introduced and validated. Results show that this hybrid approach is possible, and that both the P300 and earlier ERP components show activity that could be useful in a BCI. This conference paper is part of a larger research study with a parametric comparison between the hybrid BCI described here and two "simple" BCIs based on the P300 or M-VEP. Results show that our new hybrid approach results in a statistically significant improvement in classification accuracy and bit rate [4]. Future work should explore other task and stimulus changes, with attention to the neuroelectrophysiology literature as well as subjective preferences, robustness to different lighting conditions, practicality for persons with different visual and other deficits, and other practical factors.

5 Acknowledgments

This work was supported in part by the Grant National Natural Science Foundation of China, under Grant No.61074113 and supported part by Shanghai Leading Academic Discipline Project, Project Number: B504. This work was also supported by the Information and Communication Technologies Collaborative Project action "BrainAble" within the Seventh Framework of the European Commission, Project number ICT-2010-247447.

References

[1] 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.

[2] B. Hong, F. Guo, T. Liu, X. Gao, and S. Gao. N200-speller using motion-onset visual response. *Clin Neurophysiol*, 120: 1658–1666, 2009.

[3] J. Jin, B. Z. Allison, E. W. Sellers, C. Brunner, P. Horki, X. Y. Wang, and C. Neuper. Optimized stimulus presentation patterns for an event-related potential. EEG-based brain computer interface. *Med Biol Eng Comput*, 49: 181–191, 2011.

[4] J. Jin, B. Z. Allison, X. Wang, and C. Neuper. A comparison of three brain-computer interfaces based on P300 Potentials motion-onset visual evoked potentials or a hybrid approach using both signals. *Under review*, 2011.

[5] X. Lei, P. Yang, D. Z. Yao. An empirical bayesian framework for brain computer interface. *IEEE Trans Neural Syst Rehabil Eng*, 17: 521–529, 2009.

[6] 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*, 121: 1109–1120, 2010.

Single Trial Detection of Spatial Covert Visual Attention for BCI

L. Tonin[1], R. Leeb[1], J. del R. Millán[1]

[1]Chair in Non-Invasive Brain-Machine Interface,
Ecole Polytecnique Fédérale de Lausanne, Switzerland

luca.tonin@epfl.ch

Abstract

This paper discusses a time-dependent classification approach for single trial recognition of spatial covert visual attention for Brain-Computer Interface. Covert visual attention is a natural and intuitive mental task that does not require any external stimulation. The possibility to recognize it from single trials is essential for a future online close-loop BCI. Experimental results indicate the feasibility of the proposed approach and its high performance in an offline study. We achieved an accuracy of $84.1 \pm 8.9\,\%$ with a rejection of $6.5 \pm 6.6\,\%$ averaged across subjects.

1 Introduction

Some Brain-Computer Interface (BCI) paradigms rely on human visual processes to elicit specific patterns in the brain (i.e. P300, RSVP, SSVEP). In general, these paradigms require external stimulation and even overt attention whereby the user must gaze at the desired target. A much more natural, flexible and truly "brain" approach is to exploit voluntary covert visual attention that does not require any stimulation nor gazing.

Different neurophysiological studies have demonstrated the involvement of α-band in visual attention tasks [1–4]. The synchronization of the alpha band in the parieto-occipital regions seems to reflect an ipsilateral inhibition mechanism in the retinotopical spatial organization of the visual cortex [3]. Some works have shown the possibility to discriminate the spatial focus of covert visual attention on grand averages [5]. However, attempts to recognize it from single trials are rare [6–8] although essential for a future online closed-loop BCI based on covert visual attention. In this work, we propose a time-dependent classification approach for single trial recognition of covert visual attention and report its high performance on an offline study with three subjects.

2 Methods

Three healthy volunteers (age 28 ± 2.5 years) participated in this experiment. All subjects reported normal or corrected-to-normal vision. Subjects didn't have any previous experience with spatial covert visual attention paradigms.

2.1 Visual Paradigm and Task

In this study we focus on a two-class visual attention task. A white cross in the middle of screen and two target positions were displayed (white circles at a radius of $15\,^\circ$, position bottom left and bottom right). Subjects were instructed to gaze continuously at the white cross and focus their attention on one of the two targets according to the symbolic green cue (size $3.12\,^\circ$) appearing at the beginning of the trial. After 3000–5000 ms at the end of the attention period, a red circle appears at the correct target location. Figure 1 shows the schematic representation of the protocol with the time intervals adopted.

Each subject performed a total of 200 trials in 5 different runs along the same session day. Equal number of stimuli for both classes were presented.

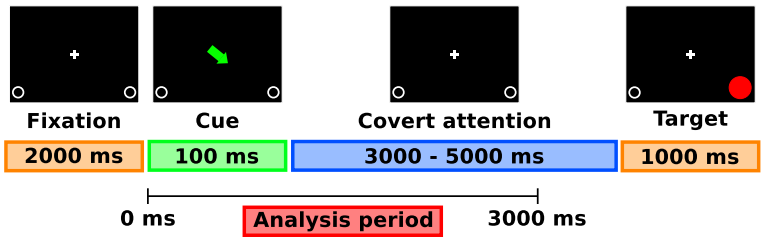


Figure 1: The visual attention protocol. The fixation period starts 2000 ms before the cue. The cue (100 ms duration) indicated the side where to focus attention. After 3000–5000 ms the circle corresponding to the correct target location becomes red (for 1000 ms). In the figure, the red circle is much larger than in reality for visualization purposes.

2.2 Data Acquisition and Processing

Signals were recorded with a 64-channel Electroencephalogram (EEG) system at 2048 Hz sampling rate. In parallel Electrooculogram (EOG) was recorded by means of 3 additional electrodes, two placed sideways to the eyes and one in the middle of the front. EEG was filtered (Butterworth filter, order 3, cut-off frequencies 5–35 Hz) and downsampled to 512 Hz. Afterwards, we applied a Laplacian spatial filter based on 8 neighbors (if possible).

We computed continuous wavelet transformation (complex Morlet family) on the data. The frequency range has been limited to 8–24 Hz. The mother wavelet has been selected in order to highlight the contribution in the α-band. This selection ensured a minimum frequency resolution of 2.65 Hz (at 24 Hz) and a minimum time resolution of 90 ms (at 10 Hz). Baseline has been computed trial by trial (baseline interval 300 ms before the cue) and has been subtracted to the trial period.

Vertical and horizontal eye movements were computed with a bipolar derivation of the EOG electrodes in the frequency range of 1–7 Hz. We discarded trials where any of the two EOG components had an amplitude higher than 30 μV. This value was selected by visual inspection. The average number of discarded trials across subjects was 8.3 ± 3.6 %.

2.3 Features Analysis

The first part of this work has been devoted to study the separability of the features distributions over time during the trial. The wavelet transformation allows us to see clearly this evolution for each pair frequency-channel (defined as feature). Based on neurophysiologycal evidences, we preselected the features coming from the parieto-occipital regions of the brain. We performed the analysis until 3000 ms after the cue in order to avoid target-related bias (see Figure 1). We divided each trial in 10 non-overlapping windows (length of each window 312.5 ms) in order to understand the evolution of the features over time. In each window, we computed the Fisher's score of each feature. This way, we obtained the most discriminable features for each time interval.

In addition, we have estimated a confidence level for each window. This value reflects how much we can trust the features in a given time interval. This confidence level has be computed by normalizing (window by window) the sum of the Fisher's score values of the selected features with respect to the minimum and maximum of the trial.

2.4 Time-dependent Classification

In this Section, we describe the classification method used in order to recognize the visual attention task on single trials. The analysis performed gave us the possibility to select different features according to the time interval. For each time window, we trained and tested a QDA classifier with the most discriminable features. In addition, for each time window t, we transformed the output of the corresponding QDA classifier into a posterior probability by using as prior the posterior at time $t-1$. Furthermore, we accumulated evidence over time by combining the posterior probabilities

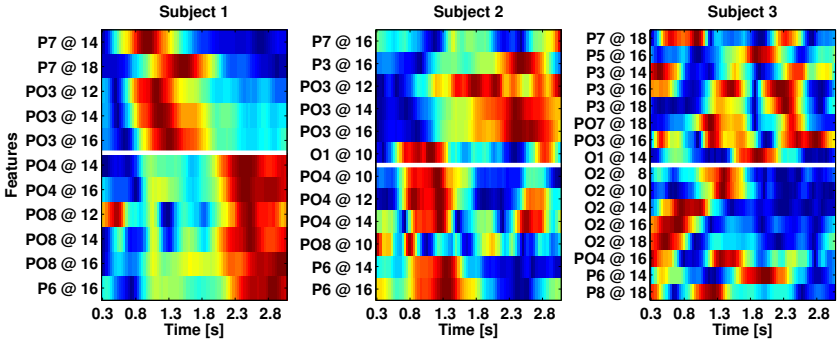


Figure 2: Evolution over time of the most discriminant features selected individually in each time window. The values have been normalized between 0 (blue) and 1 (red). The horizontal line represents the division between left and right hemisphere (top and bottom part, respectively).

by means of the following expression:

$$p(y_t) = \alpha(t) \times p(y_t|x_t) + (1 - \alpha(t)) \times p(y_{t-1}) \tag{1}$$

where $p(y_t|x_t)$ is the posterior probability, $p(y_{t-1})$ the evidence accumulated up to time $t-1$ and $\alpha(t)$ an integration parameter corresponding to the confidence level of each window (see Section 2.3). Afterwards, we integrated $p(y)$ over the whole time period (0–3000 ms) to make the decision (left, right or rejected) at the end of the trial. The rejection threshold was fixed to 0.8 for all subjects. We tested our method with 10-fold cross validation and averaged the performances.

3 Results

In this Section, we first describe the outcomes of our feature analysis and motivate the time-dependent classification approach introduced before. In the second part, we present the classification results of this method.

Figure 2 shows the evolution over time of the features with the highest Fisher's score values. These features were individually selected in each time window. The first clear result is that they are not consistent across subjects. Nevertheless, we can see that subjects 1 and 2 have a spatial consistency with respect to the brain regions (horizontal white line in the figure). For subject 1, the left region of the brain is more discriminable in the first part of the trial, while in the second part, the right region assures more separability between the two classes. Features of subject 2 present a specular behaviour. However, for subject 3 we cannot identify a clear spatial consistency.

The investigation on the separability of the features over time yields some preliminary conclusions. First of all, the time intervals with high discriminability are not stable across subjects. Second, for each subject it's not possible to identify features that guarantee high separability during the whole trial. It is for this reason that we decided to train individual classifiers for each time window, each with the most discriminant features of that window (see Section 2.4).

This time-dependent approach reaches high classification performances on single trial for all subjects. Table 1 shows the total number of correct, wrong and rejected trials for each subject across the whole dataset. In addition, last two columns report the accuracy performances computed on the accepted trials and the rejection rate across 10-fold cross validation.

The best classification rate is for subject 1 ($89.4 \pm 7.9\,\%$). The result is in line with the features map showed in Figure 2. In fact, the separability of the selected features is more consistent over time for this subject compared to the others. Nevertheless, the time-dependent feature selection and classification assures high performances also in case of subject 3 where the evolution in time of the discriminability between the two classes cannot be well defined.

Subject	# Correct	# Wrong	# Rejected	Accuracy	Rejection
1	157	19	8	$89.4 \pm 7.9\,\%$	$4.4 \pm 5.6\,\%$
2	131	34	21	$79.5 \pm 8.6\,\%$	$11.2 \pm 6.9\,\%$
3	141	28	5	$83.4 \pm 8.1\,\%$	$2.8 \pm 3.9\,\%$

Table 1: Results of the 10-fold cross validation study. Total number of correct, wrong and rejected trials across the whole dataset, as well as accuracy and rejection performances.

4 Conclusion

In this work we propose and demonstrate a time-dependent classification approach for single trial recognition of covert visual attention. The motivation for such an approach is that, based on our findings, it seems not to be possible to identify features with high discriminability that people can sustain over time. This result is supported by recent neurophysiological studies on time modulation of α-power during spatial visual attention [4]. Experimental results with three subjects indicate the feasibility of the approach in an offline study where grand average performance is $84.1 \pm 8.9\,\%$. Interestingly, our approach is a powerful mixture of neurophysiological findings and data-driven analysis. The former guide the selection of the parieto-occipital cortical regions and the frequency bands. The latter helps in exploiting natural fluctuations in the discriminant features so as to avoid any assumption about time intervals where to carry out classification.

The next challenges are twofold: to demonstrate this covert visual attention BCI in a closed-loop experiment and to show the stability of the selected features across different sessions (days).

Acknowledgments

This work is supported by the European ICT Programme Project TOBI FP7-224631.

References

[1] P. Sauseng, W. Klimesch, W. Stadler, M. Schabus, M. Doppelmayr, S. Hanslmayr, W. R. Gruber, and N. Birbaumer. A shift of visual spatial attention is selectively associated with human EEG alpha activity. *Eur. J. Neurosci.*, 22:2917–26, 2005.

[2] S. P. Kelly, E. C. Lalor, R. B. Reilly, and J. J. Foxe. Increases in alpha oscillatory power reflect an active retinotopic mechanism for distracter suppression during sustained visuospatial attention. *J. Neurophysiol.*, 95:3844–51, 2006.

[3] T. A. Rihs, C. M. Michel, and G. Thut. Mechanisms of selective inhibition in visual spatial attention are indexed by α-band EEG synchronization. *Neuroscience*, 25:603–10, 2007.

[4] T. A. Rihs, C. M. Michel, and G. Thut. A bias for posterior α-band power suppression versus enhancement during shifting versus maintenance of spatial attention. *NeuroImage*, 44:190–9, 2009.

[5] N. M. Schmidt, B. Blankertz, and M. S. Treder. α-modulation induced by covert attention shifts as a new input modality for EEG-based BCIs. In *Proc. 29th A. Int. Conf. IEEE Syst. Man Cybern. Soc.*, 2010.

[6] F. Galán, J. Palix, R. Chavarriaga, P. W. Ferrez, E. Lew, C. A. Hauert, and J. del R. Millán. Visuo-spatial attention frame recognition for brain-computer interfaces. In *Proc. Int. Conf. Cogn. Neurodyn.*, 2007.

[7] M. Van Gerven and O. Jensen. Attention modulations of posterior alpha as a control signal for two-dimensional brain–computer interfaces. *J. Neurosci. Methods*, 179:78–84, 2009.

[8] M. S. Treder, A. Bahramisharif, N. M. Schmidt, M. A. van Gerven, and B. Blankertz. Brain-computer interfacing using modulations of alpha activity induced by covert shifts of attention. *Journal of NeuroEngineering and Rehabilitation*, 8:24, 2011.

About to Fail! Detecting Subliminal Errors: a New Tool for BCI?

M. Dyson[1], C. Roger[2], L. Casini[1], B. Burle[1]

[1]Laboratoire de Neurobiologie de la Cognition, Université de Provence, Marseille, France
[2]Department of Experimental Psychology, Universiteit Gent, Ghent, Belgium

matthew.dyson@univ-provence.fr

Abstract

Detecting brain error potentials has attracted a lot of interest in BCI. Here we report classification results discriminating, not only errors from correct trials, but also "partial" errors from correct trials. Results demonstrate that partial errors can be detected by a classifier trained on pure correct and error trials when EEG component windows are adequately aligned. This research suggests partial errors, which are not necessarily consciously perceived, could be automatically identified, allowing for the detection of a possible precursor to errors.

1 Introduction

The possibility of detecting error potentials has become an important challenge in BCI, since it could allow a system to continuously learn in an unsupervised manner [1]. Falkenstein et al. [2] identified an EEG component peaking just after error commission in reaction time (RT) tasks. This component is commonly referred to as Error Negativity (Ne) or Error-Related Negativity (ERN). A similar activity is observed on "partial errors", characterized by subliminal small bursts of EMG activity indicative of an incorrect physical response, with amplitude levels too small to be considered a full error [3]. A similar wave (similar topographies and time courses, same cortical generators), although of much smaller amplitude, has also been reported on "pure correct" trials (*i.e.* those trials without any sign of incorrect response activation) [3, 4]. The existence of partial errors suggests discrimination between error and correct trials should not be considered as a binary problem, the presence or absence of a component, but rather as an amplitude discrimination task. Relative to errors, Ne/ERN in partial errors shares more similarity with that of pure correct activity [3,4]. In this paper we adopt BCI techniques to determine if Ne/ERN differences between pure correct trials, partial error trials and error trials are structured such that a classifier trained on pure correct and error trials is able to correctly identify partial error trials.

2 Methods

2.1 Data Acquisition, Preprocessing & Feature Extraction

Data was collected whilst subjects performed an Eriksen's flanker task. Subjects had to respond with a right or left thumb keypress as quickly as possible depending on whether a central target letter was 'S' or 'H'. The central target letter was flanked by two distractors which could be either compatible (e.g. SSS) or incompatible (e.g. HSH). After stimulus presentation subjects had one second to respond. Stimuli were extinguished as subjects responded. The next stimuli was presented one second after a response. All types of trials (HHH, HSH, SHS, and SSS) were presented in equiprobable pseudorandom order. Subjects performed 20 blocks of 128 trials.

EEG was recorded from 64 channels (10-20 system positions, BIOSEMI ActiveTwo). Electromyographic activity wasrecorded from the flexor pollicis brevis of each hand by paired surface

electrodes. Eye movement artefacts were corrected using a statistical method [5]. Other artefacts were rejected by visual inspection. The onset of EMG activity was manually marked by visual inspection. Trials were sorted into pure correct, error and partial error trials based on responses and patterns of EMG activity [4]. Data was sampled at 1024 Hz, and downsampled to 256 Hz.

Three referencing methods were tested: monopolar signals referenced to the left mastoid, surface Laplacian (based on four closest neighbours) and use of the common average. Nine channels were selected in the general proximity of the anterior cingulate cortex / supplementary motor area; F2, Fz, F1, FC1, FCz, FC2, C1, Cz and C2. Fourteen wavelets were tested, initially based on those used in [6]; db3, db4, db5, db6, db7, db8, db9, db10, sym2, sym3, sym4, coif3, coif4 and coif5. All wavelets were tested at levels two to five. Approximation coefficients were used as features.

2.2 Discrimination of Pure Correct and Error Trials

For each subject a random sample of pure correct trials of equal size to the number of error trials available were taken. Trials were epoched in 500 ms windows, retaining 100 ms prior to EMG response and 400 ms post. A grid search was performed over the feature space described in Section 2.1. For each step of the grid search, referencing was performed and wavelet approximation coefficients extracted. Data was divided into four splits. Sequential forward floating selection (SFFS) performed coefficient selection on three inner splits. A maximum of six coefficients were selected as features. Classification was based on FLD using ten fold cross validation. Mean accuracy classifying unseen data on the outer split, for the four folds, was used as test accuracy.

2.3 Discrimination of Pure Correct and Partial Error Trials

For discrimination of pure correct from partial error trials we select the feature parameters which performed best across the grid search described in Section 2.2, as measured by test accuracy obtained discriminating pure correct from error trials. SFFS was applied to the non split feature extracted data described in Section 2.2 to obtain six coefficients to use in the discrimination of partial error trials, and an FLD classifier was trained on the data described in Section 2.2.

In the case of a partial error trials we assume that the erroneous nature of an intended response has been correctly detected, leading to a suppression of the associated physical activity. When viewed as a grand average time locked to EMG onset, peak to peak differences exist between Ne/ERN ERPs induced in the case of partial errors and errors [4]. The peaks in the case of partial errors occurring earlier than those of errors. To use the FLD classifier trained on pure correct and error trials with partial errors we must temporally align Ne/ERN activity occurring with a partial error with that of the error trials used for training the classifier. This could be achieved via a subject specific analysis of the ERPs produced in error and partial error trials and the selection of a suitable offset value from EMG activity in either class to maximally align Ne/ERN.

In this analysis we instead chose to classify at 31 negative offset positions, corresponding to setting segmentation points 0 to 30 samples prior to the onset of subliminal EMG activity in partial error trials. Trials were epoched in the manner described in Section 2.2, substituting each offset position in place of the EMG response. By classifying at multiple offset positions we provide support for the hypothesis that Ne/ERN activity occurs in both error classes by observing classification as the ERP component in partial error trials moves from a position of low temporal alignment (with respect to the Ne/ERN of error trials used to train the classifier) at an offset value of zero through the tested 120 ms alignment window. Partial error trials had been epoched 400 ms prior and 400 ms post physical response in a previous analysis [4], trials which did not allow for a second epoching centred up to 120 ms prior to the partial EMG burst were discarded.

A new sample of pure correct trials of equal size to the number of partial error trials eligible for analysis were selected, excluding those trials used in Section 2.2. Feature extraction was performed on the new pure correct and partial error trial data based on the parameters used for training the FLD classifier. Classification was performed between pure correct trials and partial error trials at each of the offset values described, using the FLD trained on pure correct and error trials.

3 Results

When discriminating between pure correct and error trials, significant differences existed between the maximum accuracy achieved by subjects dependent on referencing condition (Kruskal-Wallis, $df = 2$, $P < 0.05$) Multiple comparison of mean rank revealed common average referenced data to perform better than Laplacian transformed. Separating the same data by channel showed Cz to perform best (Kruskal-Wallis, $df = 8$, $P < 0.01$), significantly outperforming the poorest performing channel (F2). The differences described were found in both training and test results.

Discrimination rates and features selected discriminating pure correct from error trials are shown in Table 1. Results in Table 1 represent accuracy achieved treating Section 2.2 as an independent investigation, with features selected based on training set accuracy. Results obtained applying a FLD trained on pure correct and error trials to pure correct and partial error trials, as described in Section 2.3, are shown in Table 2. The 'Fixed' column in Table 2 refers to rates obtained using an optimal across subject offset value of 75 ms, this offset was determined post-hoc via analysis of partial error discrimination results. Optimal accuracy rates and individual offset values are also presented. A graph depicting the relationship between offset values and classification accuracy for individual subjects is shown in Figure 1.

Subject	Trial Count	Accuracy Train	Accuracy Test	Site: Reference & Wavelet-Level	
1	130	91.35	89.23	FCz: MP	coif3-4
2	105	86.61	80.45	Cz: CA	db5-3
3	43	91.25	74.32	Cz: CA	sym2-2
4	39	97.08	84.58	Cz: CA	db8-2
5	145	86.93	82.07	Cz: CA	db8-2
6	100	91.25	86.00	Cz: CA	db7-2
7	63	94.37	89.53	FCz: CA	db3-5

Table 1: Classification of pure correct trials against error trials. Subject, trial count, accuracy on training set, accuracy on test set. Site, referencing method (CA: common average, MP: monopolar) and wavelet applied for features.

Subject	Trial Count	Accuracy Fixed	Accuracy Best	Accuracy Offset	Site: Reference & Wavelet-Level	
1	313	89.30	91.21	-55	Fz: MP	sym2-3
2	297	77.10	77.61	-86	Cz: CA	sym2-2
3	305	68.36	69.34	-90	Cz: CA	coif4-5
4	53	58.49	59.43	-117	C2: CA	db3-4
5	447	72.93	76.29	-47	Cz: CA	db9-4
6	464	78.34	78.66	-78	Cz: CA	db3-4
7	377	79.44	80.64	-94	Cz: CA	db3-5

Table 2: Classification of pure correct trials against partial errors. Subject, trial count accuracy at fixed across subject offset of 75 ms, best accuracy and optimal offset value. Site, referencing method and wavelet applied for features.

4 Discussion

The discrimination rates presented demonstrate that a classifier trained on pure correct and error trials can be used to detect individual partial error trials. These results support the proposition that partial errors are an important phenomenon enabling insight into the properties of Ne/ERN as a continuum, rather than tacitly approaching the component as being binary in nature. Partial errors were tested against pure correct trials, we justify this based on the fact that both responses share structurally similar Ne/ERN properties, whilst acknowledging that the within class variability of randomly sampled EEG could be greater than that of pure correct trials.

With respect to differences found between referencing methods it is important to consider that the results presented are based on single features. The Laplacian transformation is known to increase negativity observed in correct trials [3] and, as a single reference is not subtracted from all channels, may produce data which is more noisy on a trial by trial basis. However, it may be the case that the spatial resolution advantages of the Laplacian transform will become more apparent when complementary information from multiple channels are considered in classification.

The classification results obtained discriminating pure correct from error trials (Table 1) may be enhanced by the trial sorting methods applied. Contrasted with designating trials solely as either correct or error, the removal of partial errors and creation pure correct trials should enhance the distinction between classes. When testing pure correct trials against partial errors we note

only two subjects fail to obtain optimal classification accuracies above 75 % (Table 2, subjects 3 & 4). Inspection of Table 1 suggests poor classification obtained with subjects 3 and 4 is likely attributable to insufficient error trials for use as training data. An optimal across subject offset value of 75 ms for partial error trials was derived from classification data presented in Figure 1. This offset value corresponds very closely to the peak to peak difference presented in grand average between errors and partial errors in previous work based on the same data [4] (see Figure 1).

Ne/ERN in EEG could be used for the automated detection of partial errors, a class more structurally similar to correct responses than the typically defined error. The detection of partial errors online, using the current method of analysis of both EMG and EEG, would allow for subject specific adaptation of parameters in time-dependent behavioural experiments. With respect to interfaces requiring rapid responses, partial errors have the potential to provide diagnostic information by providing a measure of erroneous response activation at an earlier and lower level than that of outright error. As a BCI example, a P300 Donchin speller could be augmented with a variant of a partial error detection system, using EMG and EEG or potentially EEG alone. This method may quantify incorrect recognition responses to individual flashes, allowing detailed study of possible perceptual errors.

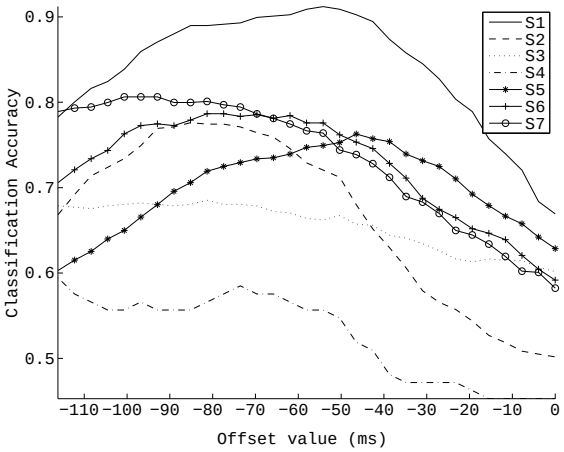


Figure 1: Classification of pure correct trials against partial errors. Accuracy rates obtained at differing offset delay values for individual subjects. Offset values from 0 to 120 ms.

5 Conclusion

Results indicate partial errors may be detected based on information derived solely from pure correct and error trials. This research demonstrates that Ne/ERN activity can be identified in single trial EEG without an outward error occurring.

References

[1] P. W. Ferrez and J. Del R. Millán. You are wrong!—automatic detection of interaction errors from brain waves. In *Proc. 19th Int. Joint Conference on Artificial Intelligence*, pages 1413–1418, Edinburgh, Scotland, 2005.

[2] M. Falkenstein, J. Hohnsbein, J. Hoorman, and L. Blanke. Effects of crossmodal divided attention on late ERP components. II. error processing in choice reaction tasks. *Electroencephalogr. Clin. Neurophysiol.*, 78:447–55, 1991.

[3] F. Vidal, T. Hasbroucq, J. Grapperon, and M. Bonnet. Is the 'error negativity' specific to errors. *Biol. Psychol*, 51:109–128, 2000.

[4] C. Roger, C. G. Benar, F. Vidal, T. Hasbroucq, and B. Burle. Rostral cingulate zone and correct response monitoring: ICA and source localization evidences for the unicity of correct- and error-negativities. *Neuroimage*, 51:391–403, 2010.

[5] G. Gratton, M. Coles, and E. Donchin. A new method for off-line removal of ocular artifact. *Electroencephalogr. Clin. Neurophysiol.*, 55:468–484, 1983.

[6] M. Salvaris and F. Sepulveda. Wavelets and ensemble of FLDs for P300 classification. In *Proc. 4th Int. IEEE EMBS Conf. on Neural Engineering*, pages 339–342, Antalya, Turkey, 2009.

The Rapid Serial Visual Presentation Paradigm Tested Online in a Brain-Computer Interface Speller

L. Acqualagna[1], B. Blankertz[1]

[1]Berlin Institute of Technology, Machine Learning laboratory, Berlin, Germany

lauraacqualagna@libero.it

Abstract

Rapid serial visual presentation (RSVP) can be used as a novel paradigm in Brain-Computer Interface (BCI) research for mental typewriting. Characters are displayed in a fixed central location. Hence, this kind of speller does not require any gaze shifts and can be operated by non-spatial feature attention. In the present study, we investigate three different variants of the RSVP speller in an online experiment with twelve participants. The results show for the best performing variant a mean online spelling speed up of 1.43 symbols/min and a mean symbol selection accuracy of 94.8 %, competitive with the other gaze-independent spellers in literature.

1 Introduction

A Brain Computer Interface (BCI) speller is a communication device, which can be used by patients suffering from neurodegenerative diseases (e.g. Amyotrophic Lateral Sclerosis) to select symbols in a computer display. This kind of BCI is often based on a class of brain signals named Event Related Potentials (ERPs), which arise when the user focuses the attention on the target symbol. Many of the ERP spellers use visual stimuli and require eye movements [1,2]. For patients unable to overtly fixate the target symbol, it is crucial to develop a speller independent of gaze shifts. Different groups investigated visual ERP spellers based on covert attention and non-spatial visual attention [3, 4]. In [3], three different variants of the two stages Hex-o-Spell are tested, with a stimulus onset asynchrony (SOA) of 200 ms. Online mean symbol selection accuracies of 88 % - 90 % - 97 % are reached. In [4], characters are grouped as the rows and columns of the classic Matrix Speller [5], and the groups are presented in a circular arrangement sequentially (SOA 400 ms), with an online selection accuracy up to 96.3 %. In the present online study, we investigate an ERP speller based on the rapid serial visual presentation paradigm (RSVP), which can be operated by non-spatial selective feature attention. We improved the design of the RSVP speller tested offline in our previous study [6]. In that study, we showed that the RSVP is a promising new paradigm.

2 Methods

2.1 Participants and Apparatus

12 subjects (6 f, 25–52 y) participated in the experiment. The study was performed in accordance with the declaration of Helsinki and all participants gave written consent. EEG was recorded at 1000 Hz using BrainAmp amplifiers and an actiCAP (Brain Products) with 64 electrodes.

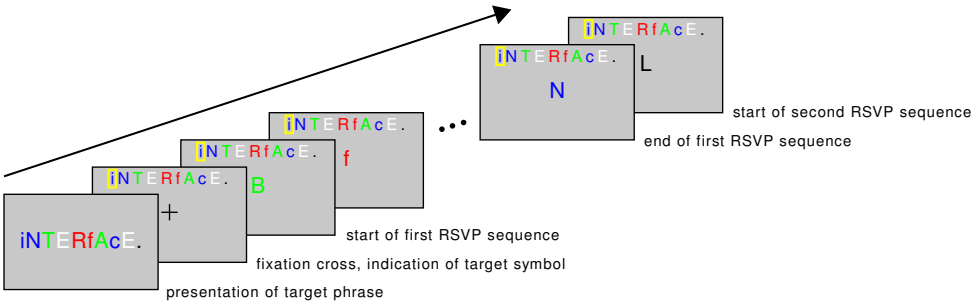


Figure 1: RSVP paradigm. First, the sentence is presented on the display. After the fixation cross, the RSVP of the symbols starts. The target letter is highlighted on the top of the screen.

2.2 Design and Procedure

In the RSVP, 30 symbols have been used: 26 letters of the English alphabet, the 2 punctuation marks ".", "!", the underscore, used as space symbol, and the backspace (Figure 1). In order to enhance the differences between the shapes of the different letters, some were uppercase and some lowercase. We investigated 3 different conditions: 1) no-color-116 ms; 2)color-116 ms; 3) color-83 ms. The order of the conditions was counterbalanced across participants. In the no-color condition all the letters were black, in the color they were divided into 5 groups: red (fRyGk<), white (pJUX!E), blue (iSwc_N), green (TBMqAH), black (LdvOz.). In this condition, two SOAs have been used: 83 ms and 116 s.

Participants sat a distance of approximately 80 cm from the screen. For each condition, they had to perform 3 phases: calibration, copy-spelling and free-spelling. In the calibration, they had to spell the sentence "BRAiN_cOMpUTER_iNTERfAcE". In the RSVP, the 30 symbols were randomly shuffled and presented 10 times on a gray background. Between the 10 sequences there was a short break of 0.3 s. In the color condition, the order of the colors was fixed to ease the allocation of attention to the target symbol. Participants had to silently count the number of occurrences of the target letter in the RSVP. In the copy-spelling, the subjects had to spell another sentence: "LET_yOUR_BRAiN_TALk" in the color-116, "wiNTER_iS_dEpRESSiNG" in the no-color-116, and "doNT_wORRy_BE_HAppy" in the color-83. In this phase, after the 10 sequences, a symbol was selected based on the classifier outputs and presented on the screen. In the free-spelling phase, participants were asked to conceive a sentence containing at least 15 symbols. They were instructed to use the backspace symbol if the classifier selected the wrong letter.

The RSVP speller was implemented in the open-source framework Pyff [7] with VisionEgg [8].

2.3 Data Analysis

For ERP analysis, EEG data were down-sampled to 200 Hz and divided into epochs ranging from -100 ms to 1200 ms relative to the onset of each stimulus. Classification was based on linear discriminant analysis (LDA) with shrinkage of the covariance matrix [9]. The time intervals for calculating the spatial-temproal features were determined by a heuristic searching for peaks based on the sgn r^2 [9]. During the training of the classifier, 5 different temporal windows were selected and occasionally adjusted by the experimenter before starting the on-line session.

3 Results

3.1 ERPs

ERP analyses show two main components time-locked with the target onset: the N2 in the occipital cortex, and the P3 in the central-parietal cortex. Table 1 reports the mean amplitudes and latencies

Condition	P3 Amplitude	P3 Latency	N2 Amplitude	N2 Latency
No-Color 116	6.65	498.75	-2.10	247.00
Color 116	7.16	490.00	-2.13	272.50
Color 83	5.76	480.83	-1.31	248.67

Table 1: P3 and N2 amplitudes (μV) and latencies (ms).

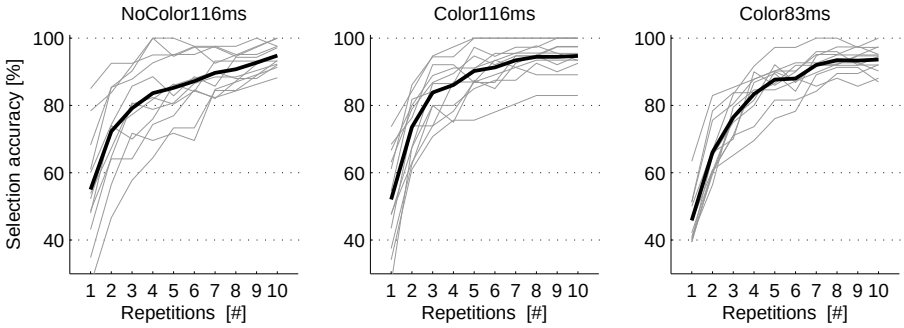


Figure 2: Symbol selection accuracies. The grey lines represent the different subjects, the black thick lines the mean classification accuracies over all the subjects.

of the P3 and N2 in the three conditions. Amplitudes and latencies were subjected to a two-way analysis of variance (ANOVA) with factors "Condition" (no-color 116, color 116, color 83) and "Status" (target, non-target). There is a significant effect of the Status on the P3 amplitude ($F = 244.67$, $p<0.001$) and latency ($F = 33.04$, $p<0.001$). The effect of the Condition ($p = 0.298$) is not significant, even if the P3 amplitude is smaller with a shorter SOA, consistent with earlier works [10]. Regarding the N2 component, there is a significant effect of both the Status ($F = 5.1$, $p<0.05$) and the Condition ($F = 8.63$, $p<0.001$) on the amplitude, but not on the latency (Status: $p = 0.68$; Condition: $p = 0.05$). Tukey-Kramer post-hoc test reveals that the N2 amplitude is significantly lower in the color 83 condition than in both the 116 conditions.

3.2 Classification

The mean online classification accuracy amounts to 94.8 % in the no-color-116 ms condition, 94.7 % in the color-116 ms and 93.6 % in the color-83 ms, after 10 sequences. One-way ANOVA analysis reveals that the differences among the conditions are not significant ($p = 0.756$). In the offline analyses, the data of the calibration phase were used as training set, and then the classifier was applied to the copy-spelling and the free-spelling. Figure 2 shows the classification accuracy as a function of the number of sequences. In all the conditions, the performance of the classifier increases significantly with the number of repetitions ($F = 103.58$, $p<0.001$). The evaluation is based on the accuracy of symbol selection, one out of 30. The chance level is 3,33 %. The mean online spelling speed was 1.16 symb/min in the no-color 116 ms (5.65 bits/min), 1.15 symb/min in the color 166 ms (5.66 bits/min), and 1.43 symb/min in the color 83 ms (7 bits/min), where all overhead (presentation of target symbol and feedback etc.) of the application was taken into account.

4 Discussion

All the subjects were able to operate successfully the RSVP speller. Robust ERP components and high mean symbol selection accuracies were achieved in all the conditions. The improvements made in the design of the RSVP speller increased both the P3 amplitude and the classification performance in respect to the previous offline study [6]. These results are also competitive com-

pared to the other gaze-independent visual ERP spellers described in [3,4]. The symbol rate of the RSVP speller is higher than the one achieved in [4], but still low in comparison to the throughput reached by gaze-dependent visual spellers [11]. The spelling speed could be increased reducing the number of sequences necessary for the target's selection, e.g introducing techniques of early stopping [12] or automatic error detection [13].

This study showed that the RSVP speller can be operated with high accuracy by healthy subjects. It has a large vocabulary, is fast-paced and independent on gaze-shift. Anyway, even a gaze-independent BCI could have low performances in extreme cases of oculomotor impairment, for example for those patients who can't differentiate the presented symbol anymore. The transfer of the results to a group as patients needs to be explored in the future, and other solutions have to be investigated if necessary.

Acknowledgments We are grateful to Matthias Treder for stimulating discussions concerning the experimental design and Torsten Schmits for programming the stimulus presentation.

References

[1] M. S. Treder and B. Blankertz. (C)overt attention and visual speller design in an ERP-based brain-computer interface. *Behav & Brain funct*, 6:28, May 2010.

[2] P. Brunner, S. Joshi, S. Briskin, J. R. Wolpaw, H. Bischof, and G. Schalk. Does the "P300" speller depend on eyegaze? *J Neural Eng*, 7(5), 2010.

[3] M. S Treder, N. M. Schmidt, and B. Blankertz. Towards gaze-independent visual brain-computer interfaces. *Abstract: Bernstein Conference on Computational Neuroscience*, 2010. doi: 10.3389/conf.fncom.2010.51.00117.

[4] Y. Liu, Z. Zhou, and D. Hu. Gaze independent brain-computer speller with covert visual search tasks. *Clinical Neurophysiology*, In Press, Corrected Proof, 2010.

[5] L. A. Farwell and E. Donchin. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Clin Neurophysiol*, 70:510–523, 1988.

[6] L. Acqualagna, M. S. Treder, M. Schreuder, and B. Blankertz. A novel brain-computer interface based on the rapid serial visual presentation paradigm. *Conf Proc IEEE Eng Med Biol Soc*, 1:2686–9, 2010.

[7] B. Venthur, S. Scholler, J. Williamson, S. Dähne, M. S Treder, M. T Kramarek, K. R. Müller, and B. Blankertz. Pyff—a Pythonic framework for feedback applications and stimulus presentation in neuroscience. *Frontiers in Neuroinformatics*, 4:179, 2010.

[8] A. D. Straw. Vision Egg: an open-source library for realtime visual stimulus generation. *Frontiers in Neuroinformatics*, 2(0), 2008.

[9] B. Blankertz, S. Lemm, M. S.Treder, S. Haufe, and K. R. Müller. Single-trial analysis and classification of ERP components – a tutorial. *Neuroimage*, 2011. in press.

[10] B. Z. Allison and J. A. Pineda. Effects of SOA and flash pattern manipulations on ERPs, performance, and preference: implications for a BCI system. 59(2):127–140, Feb 2006.

[11] E. Donchin, K. M. Spencer, and R. Wijesinghe. The mental prosthesis: assessing the speed of a P300-based brain-computer interface. *IEEE Trans. Rehab. Eng*, 8:174–179, 2000.

[12] A. Lenhardt, M. Kaper, and H. J. Ritter. An adaptive P300-based online brain-computer interface. *IEEE Trans. Rehab. Eng*, 16:121–130, 2008.

[13] N. M. Schmidt, B. Blankertz, and M. S. Treder. Online detection of error-related potentials boosts the communication speed of brain-computer interfaces. 2011. submitted.

On the Effect of ERPs-Based BCI Practice on User's Performances

P. Aricò[1,2], F. Aloise[1,2], F. Schettini[1,2], S. Salinari[2], S. Santostasi[2], D. Mattia[1], F. Cincotti[1]

[1]Neuroelectrical Imaging and BCI Lab, Fondazione Santa Lucia IRCCS, Rome, Italy
[2]Dept. of Computer Science, University. of Rome "Sapienza", Rome, Italy

f.aloise@hsantalucia.it

Abstract

The event related potentials are commonly used to control EEG-based Brain-Computer Interface (BCI) systems. Several evidences seem to converge on the fact that the usage (i.e. the practice) of BCIs based on P300 paradigm would lead to a overtime stable and even higher level of subject's performance possibly related to a more robust features (e.g higher P300 amplitude). In this study we investigate whether practicing a P300-based BCI task would results "per se" in an increase of subject's performance along with session repetition. The offine analysis of task accuracy, ERP amplitudes and reaction times was conducted on a data set obtained from six healthy naïve subject exposed to a P300-based BCI usage. Results showed that both the subjects' R-Square (R^2) index related to the amplitude differences between target and non-target classes and the users' performances increased across sessions of P300 Speller BCI application usage.

1 Introduction

Electroencephalographic (EEG)-based Brain-Computer Interface (BCI) represents an alternative channel of communication and control that allows the user to act on the external world bypassing the physiological output channels of the brain. Among the variety of features which are discernible from the EEG signal and thus are suitable to control an EEG-based BCI system for communication, the event-related potential (ERP) P300 has been extensively used in healthy subjects and patients [1, 2]. One of the advantages of this BCI paradigm relays on the fact that its usage does not require a training procedure and the users are able to communicate after a short-lasting system's calibration. Furthermore, it was demonstrated how user's performance did not degrade across repeated P300-based BCI sessions, even in the case of neurodegenerative diseases such as amyotrophic lateral sclerosis (ALS; [3]). A recent study reported that enhanced attentional focus and continuity that is characteristic of a short–lasting mental meditations (such as meditative mindfulness induction or MMI) positively affect the performances of healthy naïve subjects in mastering a P300 Speller BCI system over time [4]. Taken together these evidences, albeit different in nature, seems to converge on the fact that the usage (i.e. the practice) of BCIs based on P300 paradigm would lead to an overtime stable and even higher level of subject's performance possibly related to a more robust features (e.g higher P300 amplitude). This prompted us to investigate whether practicing a P300-based BCI task would results "per se" in an increase of subject's performance along with session repetition. In order to verify this assumption, we performed a preliminary study in which an offine analysis of task accuracy, ERP amplitudes and reaction times was conducted on a data set obtained from healthy naïve subject exposed to a P300-based BCI usage.

2 Methods

Six healthy volunteers (6 females, mean age = 26.3 ± 2.4) naïve to ERPs-based BCI paradigms took part to the current study. The stimulation was provided through the BCI2000 framework [5] by a 6 by 6 matrix containing alphanumeric characters (P300 Speller [6]). Rows and columns of the matrix were randomly intensified and subjects were instructed to focus their attention only on the desired stimuli (target) and to ignore the other flashing stimuli (non-targets). Scalp EEG potentials were recorded from 8 electrode positions (Fz, Cz, Pz, Oz, P3, P4, PO7, PO8; g.USBamp, g.Tec, Austria). Channels were referenced to linked earlobe and grounded to the right mastoid. Each subject performed 6 sessions. A session consisted of 3 runs and each run was composed of 6 trials. Each trial corresponded to the selection of a character displayed on the interface and consisted of 8 stimulation sequences, where a sequence is a complete random cycle of six rows and six columns flashes ([7]). At the beginning of each new trial, the system indicated to the user the letter he had to concentrate on. Every stimulus was intensified for 125 ms, with an Inter Stimulus Interval (ISI) of 125 ms. The experimental protocol included two types of alternate sessions: Reaction Time (RT) sessions, and EEG (BCI) sessions. During the RT sessions no EEG data were recorded and the subjects were presented with the same P3Speller application; they were required to attend to the stimulation and to push a button as fast as possible each time they recognized a target stimulus. These sessions aimed both to increase subject's familiarity with ERPs-based BCI application and the subjects' RT monitoring across sessions had the scope of quantifying the subjects' level of confidence with the application. The RT and BCI sessions were alternated (RT-BCI-RT-BCI-RT-BCI) and the characters that the user had to focus on were the same for every couple of RT and BCI sessions. EEG data collected during the BCI sessions were used to determinate the subjects' offline accuracy and to perform waveforms analysis. For each participant, BCI performances were determined offline, depending on the number of stimulation sequences mediated during each trial. In particular we performed a cross-validation exploring all the possible combinations of training (2 runs) and testing (1 run) data set. Also, for the waveform analysis the EEG signal was reorganized in overlapping epochs lasting 800 ms and following the onset of each stimulus [8]. Furthermore, each epoch of the signal was baseline corrected by removing the mean of the recorded signal from -200 ms to 0 ms. Therefore, the R-Square index (R^2) was employed to quantify separability between the ERP amplitudes related to the targets and non-targets classes, in according to the classification process.

3 Results

Reaction times analysis. We compared the subjects' mean reaction times detected during the 3 RT sessions: RT Session 1 = 354.1 ± 73.8 ms; RT Session 2 = 337.9 ± 65.1 ms; RT Session 3 = 335.1 ± 54.5 ms. We performed an one way repeated measures ANOVA ($CI = .95$), using *Sessions* as factor. Results showed a decrease in the subjects' mean reaction times across sessions that however, was not statistically significant ($F(2.51) = 1.63, p = .21$).
Offline user's accuracy. Figure 1 a) showed the offline accuracy per stimulation sequence and session. We performed a two way repeated measures ANOVA ($CI = .95$) using *Stimulation sequences* and *Sessions* as factors, in order to compare the accuracy values at each stimulation sequence across sessions. Overall, results showed an increase in the mean accuracy achieved by the subjects using the P300 Speller that however was not statistically significant ($F(2.51) = 1.29$, $p = .29$).
R-Square analysis. For statistical analysis, we used the unsigned R^2 index, because we were interested in absolute difference between the two target and non-target classes. We compared the overall contributions of the ERPs across sessions, in terms of amplitude by performing a two way repeated measures ANOVA ($CI = .95$) of the unsigned R^2 values on the overall length of the epoch (0–800 ms), using *Channels* and *Sessions* as factors (Figure 1 b). Results showed a significant increase in the unsigned R^2 values across sessions of P300 Speller usage ($F(2.369) = 22.894$, $p = 10^{-6}$).

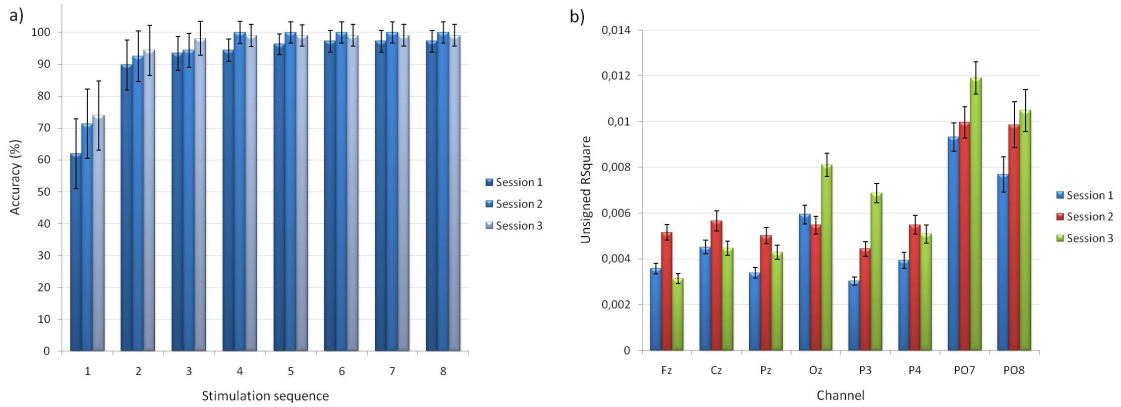


Figure 1: a) Two way repeated measures ANOVA ($CI = .95$) using *Stimulation sequences* and *Sessions* as factors about the mean subjects' accuracy calculated through offline cross-validations for each stimulation sequence using the P300 Speller interface. b) Two-way repeated measures ANOVA ($CI = .95$) using *Channels* and *Sessions* as factors about unsigned R^2 values calculated between ERP amplitudes related to the target and no-target classes.

4 Discussion

These preliminary findings mainly indicated a significant increase in the discriminability between target and non-target classes while using the P300 Speller, as revealed by the unsigned R^2 index that was significantly higher in the last sessions with respect to the first. The kind of analysis reported here highlights specific trends more accurately than the analysis performed only on the last stimulation sequence, especially if their number is significant. Using a higher number of stimulation sequences with respect to that the user really needs will lead to a "ceiling effect" resulting in a performance saturation that will prevent the detection of further changes in the performance level itself. This is of relevance in light of recent studies indicating that healthy users and potential BCI end-users could benefit of a system able to dynamically adapt the number of stimulation sequences across sessions [9, 10]. Furthermore, this approach could be useful in monitoring possible modulation of task performance related to attention level, as recently showed [4].

5 Conclusion

Our preliminary findings suggest that the practice of the ERP-based BCIs could have a positive effect on the user's performances over time, that were evaluated on the number of stimulation sequences mediated within the trial. This might indicate that a more frequent recalibration of the system could be required, especially in case of naïve users. Future study is need to monitor a long-term BCI practice, involving more subjects, in order to understand whether these effects are stable and more pronounced over time and which components of the ERPs are specifically potentiated.

6 Acknowledgments

The work is partly supported by the EU grant FP7-224631 "TOBI" project. 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] E. W. Sellers and E. Donchin. A P300-based brain-computer interface: initial tests by ALS patients. *Clinical Neurophysiology*, 117(3):538–548, 2006.

[2] F. Nijboer, N. Birbaumer, and A. Kübler. 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, 2010.

[3] F. Nijboer, E. W. Sellers, J. Mellinger, M. A. Jordan, T. Matuz, A. Furdea, S. Halder, U. Mochty, D. J. Krusienski, T. M. Vaughan, J. R. Wolpaw, N. Birbaumer, and A. Kübler. A P300-based brain-computer interface for people with amyotrophic lateral sclerosis. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 119(8):1909–1916, 2008.

[4] C. E. Lakey, D. R. Berry, and E. W. Sellers. Manipulating attention via mindfulness induction improves P300-based brain-computer interface performance. *Journal of Neural Engineering*, 8(2):025019, 2011.

[5] 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 Bio-Medical Engineering*, 51(6):1034–1043, 2004.

[6] 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.

[7] 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.

[8] 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. *Journal of Neural Engineering*, 3(4):299–305, 2006.

[9] F. Aloise, F. Schettini, P. Aricò, S. Salinari, C. Guger, J. Rinsma, M. Aiello, D. Mattia, and F. Cincotti. Asynchronous P300-based BCI to control a virtual apartment: initial tests on patients. *Clinical Electroencephalography and Neuroscience*. accepted for pubblication.

[10] F. Aloise, F. Schettini, P. Aricò, F. Leotta, S. Salinari, D. Mattia, F. Babiloni, and F. Cincotti. P300-based brain computer interface for environmental control: an asynchronous approach. *Journal of Neural Engineering*, 8(1):016001, 2011.

Comparing Efficiency for Synchronous and Asynchronous P300-Based BCIs

F. Schettini[1,2], F. Aloise[1,2], P. Aricò[1,2], S. Salinari[2], S. Petrichella[2], D. Mattia[1], F. Cincotti[1]

[1]Neuroelectrical Imaging and BCI Lab, Fondazione Santa Lucia IRCCS, Rome, Italy
[2]Dept. of Computer Science, University. of Rome "Sapienza", Rome, Italy

f.aloise@hsantalucia.it

Abstract

The efficiency for P300-based Brain computer interfaces (BCIs) can be evaluated from different points of view considering both the specific application and the system features such as user interface, usability, speed of selection and reliability. Regarding to the latter two issues, in our previous work we proposed an asynchronous BCI system which proved very strong in avoid both false positives (0.26 FP/min when subjects were engaged in other tasks) and misclassifications rate(1.5 % when they were attending to the stimulation). However, despite the asynchronous system was able to adapt the speed of selection to the user's current state, its overall information transfer rate assessed by the Wolpaw's bit-rate metric did not significantly enhance with respect to a synchronous system. In this study, we analyze the performances of the two systems using a different metric, which takes into account both the actual cost of misclassification and of abstentions. So assuming a lower cost for an unwanted abstention with respect to a wrong selection the asynchronous system resulted more significantly efficient with respect to the synchronous one in terms of time needed to achieve a selection.

1 Introduction

Electroencephalographic (EEG)-based brain-computer interface (BCI) systems potentially represent an alternative/additional communication channel which do not require muscular activity. For this reason, this technology is suitable for people suffering from severe motor disabilities. P300-based BCIs allow a relatively fast selection of letters/icons and it does not require to learn extensive training. Classic P300-based BCIs work in a synchronous mode: after a well defined number of stimuli the system always makes a "decision", assuming that the user is constantly attending to the stimulation. This mode of operation conditions works well in a laboratory context but it may have some limits when applied in a real-life condition. In a previous study a P300-based BCI implemented to control home automation was tested under asynchronous and synchronous modality of operation [1]. The asynchronous mode was proved reliable in avoiding false positives ($< 2\,\%$) when users diverted their attention from the stimulation interface, with a mean accuracy of 88.7 % comparable to that observed under synchronous modality (92.8 %). However, from the bit rate point of view the two systems did not show statistically significant differences. being the former evaluated by means of the Wolpaw's metric which considers errors and abstentions in the same way [2]. This study aims to evaluate the efficiency of the asynchronous system using a more appropriate metric, which distinguishes misclassifications from unwanted abstentions, and to compare it with the efficiency of the synchronous one.

2 Methods

2.1 Recording Protocol

Eleven healthy volunteers (4 females, 7 males; mean age and std 26,45 +/- 4,05 years) were involved in the study. The acquisition protocol was based on the P300 Speller interface [3] adapted to control a home automation system by using a 4 by 4 matrix containing 16 B&W icons representing the achievable actions on the environment. Stimuli consisted in the rows and columns intensification. Each stimulus was intensified for 125 ms with an inter stimulus interval (ISI) of 125 ms. Scalp EEG potentials were recorded (g.MobiLab, gTec, Austria, sampling rate 256 Hz) from 8 positions according to 10-10 standard (Fz, Cz, Pz, Oz, P3, P4, PO7 and PO8). Each channel was referenced to linked earlobe and grounded to the left mastoid. Here we indicate with the term Sequence a complete cycle of four rows and four columns flashes and 10 Sequences constituted a single Trial [4]. For each subject we acquired 4 Control runs of 8 Control trials each one and 12 Alternate runs during which Control and No-Control trials alternated for a total of 10 trials per run. During the No-Control trials the subjects voluntarily diverted their attention from the stimulation performing three simple no control tasks:

- Fixation Cross, 30 trials: Subjects were instructed to fixate the cross in the centre of the interface and to ignore the stimulation;
- Watch & Listen, 15 trials: Subjects were instructed to watch a movie displayed on the half of the screen beside the matrix;
- Computation, 15 trials: Subjects had to answer simple arithmetic questions posed by the operator while fixating the cross.

The Control runs were used to assess Synchronous system accuracy through a 6 round cross-validation using 16 trials to extract significant control features by Stepwise linear discriminant analysis (SWLDA) and 16 trials for testing them. We performed a 6 round cross-validation also with the alternate runs in order to assess asynchronous system performance. For each round the training dataset was composed of 35 Control trials, and 35 No-Control trials (15 Fixation, 10 Watch & Listen and 10 Computation), while the testing dataset was composed of remainder 25 Control trials. Significant control features and thresholds for asynchronous system were defined by a procedure relaying on SWLDA and ROC curves as explained in Aloise et al. [1].

2.2 Efficiency Evaluation

In order to evaluate synchronous and asynchronous systems efficiency we used the metric proposed by Bianchi et al. [5]. This metric starts from the extended confusion matrix (ECM), that consists in a N by N+1 matrix, where N is the number of the available symbols. The additional column indicates when the classifier abstains from take a decision. In order to estimate the probability that the classification of a symbol was incorrect or indeterminate, a misclassification probability matrix (MPM) can be defined from the ECM. Next, from MPM, the extended overtime matrix (EOM) can be built, representing the costs associated with errors and abstentions in terms of further steps that have to be done to correct mistakes. In the case of a system for environmental control, we associated a cost of 1 to the abstentions because the user has just to repeat the trial, while we associated a cost of 2 to misclassifications, as them can be corrected by selecting the respective UNDO actions. The latter assumption is still valid if the desired symbol is misclassified with the UNDO one, so that a correct symbol is deleted. Then it can be defined the super tax vector (ST) as:

$$\sum_{j=1}^{N+1} MPM[i,j] * EOM[i,j]$$

Where i denotes the desired class and j indicates the predicted class. Considering all symbols on the matrix equally likely, we can define the expected mean selection cost ($\overline{ESC}$), that is the number of classification required to generate a correct logical symbol:

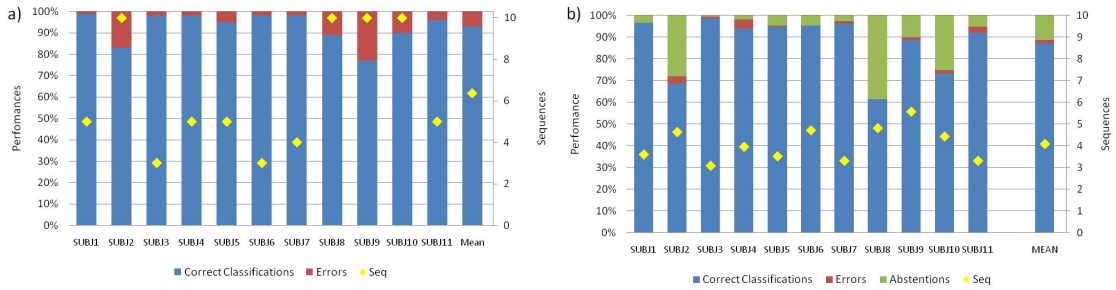


Figure 1: a) Accuracy for the synchronous system. The yellow dots indicate the number of stimulations sequences needed to reach the 95 % of accuracy. These values were used to estimate the efficiency of the synchronous system. b) Performances for the asynchronous system. The yellow dots denote the number of stimulation sequences needed to exceed the thresholds of the asynchronous classifier. These values were used to estimate efficiency for the asynchronous system.

	Subj1	Subj2	Subj3	Subj4	Subj5	Subj6	Subj7	Subj8	Subj9	Subj10	Subj11	Mean	Std
Async	0.27	0.14	0.32	0.23	0.27	0.20	0.29	0.13	0.16	0.16	0.27	0.22	0.06
Sync	0.20	0.07	0.32	0.19	0.18	0.32	0.24	0.08	0.05	0.08	0.18	0.17	0.09

Table 1: Efficiency values for the asynchronous and the synchronous system.

$$\overline{ESC} = \sum_{i=1}^{N} \frac{1}{1-ST(i)}$$

The efficiency of a system considering the time needed to achieve a classification, expressed in number of stimulation sequence (NumSeq), will be:

$$Eff = \frac{1}{NumSeq * \overline{ESC}}$$

3 Results

Figure 1 a) illustrates results about synchronous cross-validations. Performances related to the stimulation sequence needed to reach the 95 % of accuracy according to the false positive rate set to define thresholds in the asynchronous system (FPR < 0.05, see [1]). On average there was a 92 % of correct classifications using around 6 stimulation sequences. Figure 1 b) depicts results about asynchronous crossvalidations. Correct classifications reached a mean value of 87 % while errors did not exceed the 2 % using, on average, 4 stimulation sequences. Table 1 contains efficiency values both for the synchronous and the asynchronous system: the asynchronous system exhibited on average a higher efficiency with respect the synchronous one, and this difference resulted statistically significant with a paired t-test ($\alpha = .05$) confirmed (p-value = .029, t-value = 2.59)

4 Discussion

Understand the user's intentions from the ongoing EEG such as when he/she desires to suspend the control or when he recognize an error represents an important issue which could improve usability and reliability of BCI system. For instance dictionaries and grammar rules could improve spelling applications, or the error-related potentials can be used for automatic error detection [6]. However this kind of application was not extensively explored for P300 based BCI [7]. The proposed asynchronous system aims to prevent errors through abstentions, feature that also allows users to suspend the control without the need of an explicit pause. In fact, despite the synchronous system exhibited an higher percentage of correct classification (92 %) with respect to the asynchronous

system, the latter exhibited a lower error rate ($< 2\,\%$) than the synchronous system ($8\,\%$) reaching a significantly higher efficiency.

5 Conclusion

In the actual work we evaluated the efficiency of an asynchronous P300-based BCI assigning different cost to errors and to unwanted misclassifications. Particularly, considering that we tested the asynchronous system for environmental control, we assumed that an abstention has a lower cost than an error. With these general assumptions we demonstrated that the asynchronous system is statistically more efficient than a synchronous system in terms of time needed to achieve a selection, stressing the advantages of the former with respect to the latter.

6 Acknowledgments

This work was supported in part by a grant from the ARiSLA Foundation (Project BRINDISYS). 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] F. Aloise, F. Schettini, P. Aricò, F. Leotta, S. Salinari, D. Mattia, F. Babiloni, and F. Cincotti. P300-based brain-computer interface for environmental control: an asynchronous approach. *Journal of Neural Engineering*, 8(2):025025, 2011.

[2] J. R. Wolpaw, N. Birbaumer, W. J. Heetderks, D. J. McFarland, P. H. Peckham, G. Schalk, E. Donchin, L. A. Quatrano, C. J. Robinson, and T. M. Vaughan. Brain-computer interface technology: a review of the first international meeting. *IEEE Transactions on Rehabilitation Engineering*, 8(2):164–173, 2000.

[3] 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.

[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), 2008.

[5] L. Bianchi, L. R. Quitadamo, G. Garreffa, G. C. Cardarilli, and M. G. Marciani. Performances evaluation and optimization of brain computer interface systems in a copy spelling task. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 15(2):207–216, 2007.

[6] R. Chavarriaga and J. D. R. Millan. Learning from EEG error-related potentials in noninvasive brain-computer interfaces. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 18(4):381–388, 2010.

[7] B. Dal Seno, M. Matteucci, and L. Mainardi. Online detection of P300 and error potentials in a BCI speller. *Computational Intelligence and Neuroscience*, page 307254, 2010.

Stimulation Speed Boosts Auditory BCI Performance

J. Höhne[1], M. Tangermann[1]

[1]BBCI Lab, Berlin Institute of Technology, Berlin, Germany

j.hoehne@tu-berlin.de

Abstract

For most ERP-based BCI paradigms, the stimuli are presented with a predefined and constant speed. Based on the idea to boost BCI performance by optimizing the parameters of stimulation, this offline study investigates the impact of the stimulus onset asynchrony (SOA) on ERPs and the resulting classification accuracy. Therefore, a simple auditory oddball paradigm was tested in eight SOA conditions. It was found that the binary classification accuracy is not directly correlated to the SOA. A variability within subjects ($n = 5$) was observed, which indicates a potential increase in BCI performance, if the SOA is specified for each subject individually.

1 Introduction

Using a Brain-Computer Interface (BCI), users can send control signals to an application even if they are unable to control any muscle. Recent research aims to develop novel BCI paradigms with a high rate of communication. Most of these approaches are based on event related potentials (ERPs), which are the responses visible in the EEG upon a perceived event or stimulus. Various paradigms were proposed using the visual [1] or auditory [2, 3] modality of stimulation. While following the oddball paradigm, they differ in the choice of stimuli, and how the stimuli are presented. Thus it is reasonable to boost the classification accuracy and BCI performance by optimizing the stimuli characteristics. For the visual and auditory modality, this can be achieved by finding stimulation procedures that elicit the strongest possible class-dicriminative components [4, 5].

Another parameter that can be modified is the stimulation speed, which is often described by the stimulus onset asynchrony (SOA) or inter stimulus intervals (ISI). The SOA specifies the time between the onsets of two consecutive stimuli. Most BCI paradigms are applied with a SOA value between 83 ms [1] and 500 ms [6]. Comparing the visual BCI performance of two SOA levels (175 ms and 350 ms), [7] found that the choice of SOA highly impacts the BCI performance and conclude that "it appears to be worthwhile to test multiple ISI values and thereby determine the optimal value for each user". But until now, the exact choice of stimulation speed is not considered to be crucial, thus it was usually not optimized by any means.

In the present study, the parameter SOA was investigated with respect to the impact on the classification performance in a simple auditory oddball paradigm. The classical ERP literature [8] describes decreasing amplitudes of class-discriminative ERP components like P300 for decreasing SOA values and target-to-target intervals (TTI). Consequently, the main hypothesis of this study is that the binary classification accuracy (target vs. non-target) correlates with the SOA, such that fast SOA conditions result in a lower accuracy than slow SOA conditions.

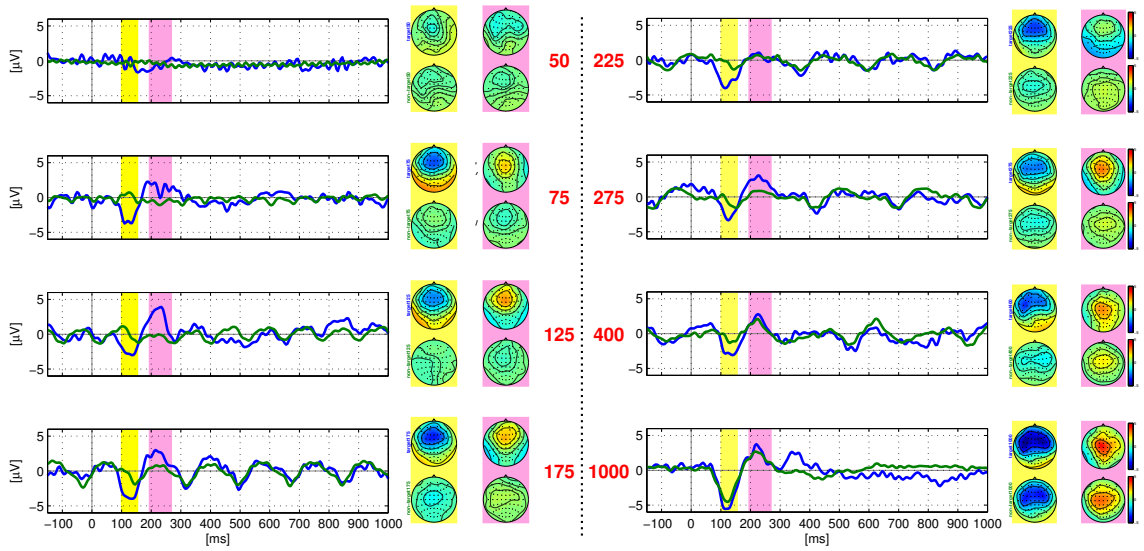


Figure 1: ERPs of *VPkab* for all eight SOA conditions. ERPs of targets (blue) and non-targets (green) and scalp maps are plotted with the same amplitude range. The scalp distributions for targets and non-targets are depicted for two intervals (yellow: [100 160] and purple: [190 270]).

2 Methods

2.1 Experimental Design

Five healthy volunteers (three male, all non-smokers) participated in a single session of an offline BCI study. Participants were not paid for participation.

Within a session, a simple auditory oddball paradigm was tested in eight conditions. In each condition, participants had to concentrate on a rare target tone while neglecting the frequent (83 %) non-target tone. Both stimuli were sinusoidal with a duration of 50 ms. The target tone had a high pitch (1000 Hz) and the non-target tone had a low pitch (500 Hz). Depending on the condition, the SOA was set to $\{50, 75, 125, 175, 225, 275, 400, 1000\}$ ms. In the following, these conditions are abbreviated with e.g. SOA_{225}.

The experiment was divided into four parts, each consisting of eight blocks with randomized order of conditions. Within one block, there were four consecutive trials of the same condition. Each trial consisted of 72–90 stimulus presentations (16.6 % targets), and the participant had the task to count the occurrence of the target stimulus. Finally, this leads to 1296 events (216 targets and 1080 non-targets) in each condition. Within one trial, the sequence of targets and non-targets was randomized, while it was assured that there were at least three non-targets between two consecutive target stimuli.

2.2 Data Acquisition

EEG signals were recorded monopolarly using a Fast'n Easy Cap (EasyCap GmbH) with 61 wet Ag/AgCl electrodes placed at symmetrical positions. Channels were referenced to the nose. Electrooculogram (EOG) signals were recorded in addition. Signals were amplified using two 32-channel amplifiers (Brain Products), sampled at 1 kHz and band-pass filtered between 0.4 and 40 Hz. The data was epoched between -150 ms and 1000 ms relative to each stimulus onset.

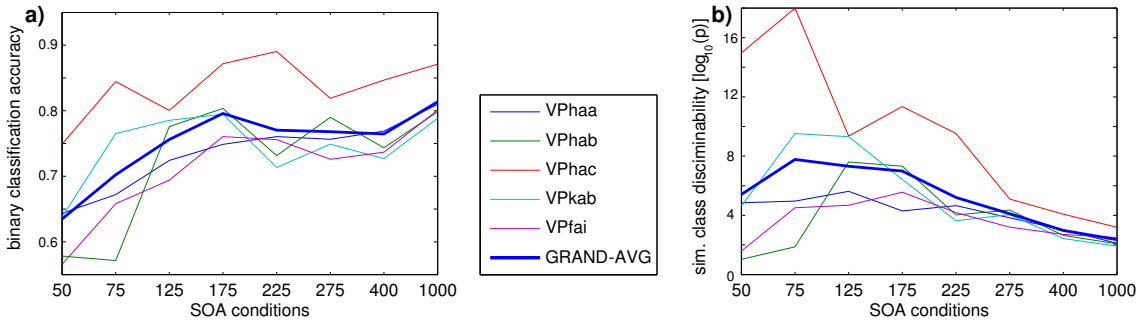


Figure 2: BCI performance for SOA conditions. The class-wise balanced binary decision accuracy (target vs. non-target, chance level 50 %) was estimated with cross-validaton (a). The simulated class discriminability (b) specifies the separability of distributions of target and non-target classifier outputs after 15 seconds of stimulation.

2.3 Classification

Binary classification of target and non-target epochs was performed using a (linear) Fisher Discriminant Analysis (FDA) using shrinkage regularization [9]. The mean potentials in selected intervals of epochs were taken as features. The intervals were selected manually for each participant and condition based on a discriminance analysis of the data. The binary classification accuracy was estimated by $[5 \times 5]$ cross validation.

Moreover, discriminability between the two classes (target vs non-target) was simulated after a fixed duration of stimulation: the number of classifier outputs accumulated within this time (here 15 sec) is higher for fast conditions (e.g. $\mathrm{N}(\mathrm{SOA}_{75}) = 200$ vs. $\mathrm{N}(\mathrm{SOA}_{400}) = 37$), but the discriminative information of a single epoch might be lower compared to slow conditions. The accumulated classifier outputs were used to estimate, how well one can separate between the two classes. Here, discriminability is quantified with the $(-log_{10})$ p-value of a one-sided t-test, since this measure is independent of the number of data points (here epochs).

3 Results

ERP plots of one participant are shown for all conditions in Figure 1. Typically for auditory oddball paradigms, an early class discriminative component (here 110–160 ms after stimulus, called N1 in the following) is present. In addition, a later positive class discriminative component (here 190–270 ms after stimulus onset, called P2 in the following) can be observed as well as a rhythmic pattern (induced by the repetitive stimuli) in the non-target ERPs. The amplitudes of the ERPs increase with an longer SOA, which is in line with the classical ERP literature [8]. This holds true particularly for the non-target ERPs (green lines). Interestingly for this subject, a second (positive) class discriminative component (300–400 ms after stimulus onset, P3) is observed for SOA_{1000} only. Thus the P3 component was not present in the faster conditions.

On average, the binary classification accuracy is highest for SOA_{1000}. Although this observation is in line with the main hypothesis, the accuracy is not decreasing monotonously with an faster SOA. For example, SOA_{175} outperforms SOA_{225}, SOA_{275} and SOA_{400} (Figure 2 a). Moreover, it can be seen that small variations in the SOA may lead to a substantial change in the classification accuracy: the performance of *VPkab* for example drops from 79 % (SOA_{175}) to 71 % (SOA_{225}). Inspection of the ERPs of this participant (Figure 1) reveals a significantly decreased P2 component in SOA_{225}. Reasons for this remain unclear, but the high number of repetitions and block-randomized order rule out any artifactual explanations.

Figure 2 b depicts the simulated class discriminability, which represents a compromise between stimulation speed and binary accuracy. For the hypothetical BCI setup with trials lasting

15 seconds, the optimal individual choice of SOA varies between 75 ms and 175 ms for the subjects and for none of the participants, a SOA greater than 175 ms should be chosen.

4 Discussion

This study investigates the impact of stimulation speed in BCI performance. Although expected, a direct correlation between SOA and classification accuracy could not be found. Instead, a surprisingly high variation of classification accuracy for the SOA conditions was present even within subjects. This fact leads to a substantial variation of the simulated class discriminability within subjects, which can be seen as surrogate for the online BCI performance.

Although the choice of SOA was mostly not considered important for ERP-based BCIs so far, the present results for an auditory ERP application underline the findings of [7], that it is reasonable to investigate methods that determine the optimal SOA for each subject.

Acknowledgments

This work is supported by the European ICT Programme Project FP7-224631 and by GRK 1589/1.

References

[1] L. Acqualagna, M. S. Treder, M. Schreuder, and B. Blankertz. A novel brain-computer interface based on the rapid serial visual presentation paradigm. In *Conf Proc IEEE Eng Med Biol Soc*, volume 1, pages 2686–2689, 2010.

[2] 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.

[3] J. Höhne, M. Schreuder, B. Blankertz, and M. Tangermann. A novel 9-class auditory ERP paradigm driving a predictive text entry system. *Front Neuroscience*, 2011. submitted.

[4] M. Tangermann, M. Schreuder, S. Dähne, J. Höhne, S. Regler, A. Ramsey, M. Queck, J. Williamson, and R. Murray-Smith. Optimized stimulation events for a visual ERP BCI. *Int J Bioelectromagnetism*, 13, 2011. in press.

[5] 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. *Advances in Neural Information Processing Systems*, 21, 2009.

[6] L. A. Farwell and E. Donchin. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalogr Clin Neurophysiol*, 70:510–523, 1988.

[7] E. W. Sellers, D. J. Krusienski, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw. A P300 event-related potential brain-computer interface (BCI): the effects of matrix size and inter stimulus interval on performance. *Biol Psychol*, 73:242–252, Oct 2006.

[8] C. Gonsalvez and J. Polich. P300 amplitude is determined by target-to-target interval. *Psychophysiology*, 39(3):388–396, 2002.

[9] B. Blankertz, S. Lemm, M. S. Treder, S. Haufe, and K. R. Müller. Single-trial analysis and classification of ERP components – a tutorial. *Neuroimage*, 56:814–825, 2011.

Introducing the Detection of Auditory Error Responses Based on BCI Technology for Passive Interaction.

T. O. Zander[1,2], D. M. Klippel[1], R. Scherer[3]

[1]Max Planck Institute for Biological Cybernetics, Tuebingen, Germany
[2]Team PhyPA, Chair of Human-Machine Systems, Berlin Institute of Technology, Berlin, Germany
[3]Institute for Knowledge Discovery, Laboratory of Brain-Computer Interfaces, Graz University of Technology, Graz, Austria

tzander@gmail.com, reinhold.scherer@tugraz.at

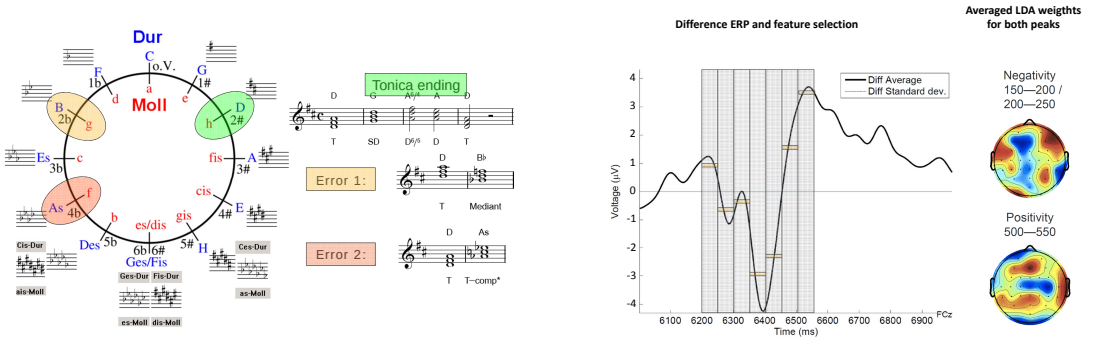
Abstract

In this article we will introduce a passive BCI system for detecting responses of the human brain on the perception of errors in music. We played cadences, sequences of chords, to subjects who are experts in music theory and actively play an classical instrument. As the expectancy is highest at the ending chord, we randomly introduced cadences with erroneous ending. In consistence with previous studies from the neurosciences we evoked an event-related potential, mainly consisting of an early right anterior negativity reflecting syntatcic error processing followed by a stronger negativity in erroneous trials at 500 ms, induced by semantic processing. We could identify single trials of these processes with a standardized, crossvalidated offline classification scheme, resulting in an accuracy of 75.7 %.

1 Introduction

Contemporary research in the field of Brain-Computer Interfaces (BCIs) is facing the problem that the visual modality might not be the best way to display information to the user. Patients in 'locked in' state lose control over their eye movements, and those who are completely locked in have completely lost it by definition. The assumption that perception with covert attention [1] is sufficient for gaining control over a BCI system, e.g. a P300 speller, was proven wrong in two recent independent studies [2, 3]. Besides more elaborate approaches for displaying sufficient information without the need for covert attention [2], a shift to other modalities might give hope. Information might be displayed through vibro-tactile interfaces [4, 5] as well as through auditory interfaces [6, 7]. Hence, it might be valuable to investigate how much of the knowledge about BCIs based on the visual domain can be transferred to other modalities. One interesting and relevant approach is the detection of error responses of the brain and it's utilization for setting up an automated correction system based on a passive BCI [8–12]. A first step for investigating error responses with a passive BCI from the tactile domain has been done in [5]. Here, we are going to discuss a similar approach for the auditory domain.

The presented experimental paradigm is based on sequences of chords, namely cadences. In cadences a chord will be determined by some degree by its predecessors. With processing the cadence further, this dependency grows stronger, such that the expectancy of an auditor for a specific last chord is very high. A violation of this expectancy initiates an error processing in the human brain, leading to the potentials mentioned above. How such a violation can be deduced from western music theory can be found in the next section.



(a) (b)

Figure 1: (a) (Cadence endings) This figure shows the three types of endings for the cadences: D-major (Tonica), As-major (Subdominant), Bb-major (Dominant). Left: The circle of fifths. Right: Chords on staves. Deviant endings (marked orange and red) occurred with a chance of 15 % while standars had a probability of 70 %. (b) Left: (Difference plot) This figure shows the Grandaverage of the differences between correct and erroneous trials. Grey marked areas indicate the time windows used for feature extraction. Orange bars in these time windows reflect estimates of the feature values for this channel (FCz). Right: (LDA weights) This figure shows the Grandaverage of the weights optimized by LDA for classification oround the ERP peaks. Upper plot: Grandaverage for classification on averages from 50 ms time windows starting at 150 and 200 ms. Lower plot: Grandaverage for classification on averages from a 50 ms time window starting at 500 ms.

1.1 A Glimpse on Western Music Theory

The study presented here investigates the effects of violating semantical rules in musical theory. The tempered intonation was used in a major scale. All trials were carried out as a sequence of in-key chords building the most simple cadence. A classical cadence in simple form is the result of combining chords in a way that in the end the subdominant is followed by the dominant which itself is then followed by the concluding tonica. According to western musical theory, the highest expectation (harmonically) appears at the end of a cadence. A special form of conclusion to a cadence was investigated here. In this case the cadence is ending on a chord referred to as *mediant*. A chord is called mediant when either its middle tone is a tone of the tonica or the middle tone of the tonica is a tone of the chord. In the case of this work the key was D-major, chosen because it is the personally preferred key of the author. The mediant chosen for the end was B(#)-major, which can be referred to as the large under mediant. Details can be found in Figure 1a.

2 Experimental Setup

We invited 10 subjects with strong background in music theory who actively play a classical instrument to our experiments. They were equipped with 32 channels of Brain Products ActiCap system, distributed in a standard way along the international 10/20 system. Impedances were kept below 20 kΩ for each electrode and data was recorded with a sampling frequency of 1000 Hz. Additionally, they were equipped with a standard, fully closed headset of Sennheiser. An experimenter instructed each subject before the starting of the recordings about the purpose of the experiment, it's duration and the planned course of events. Subjects sat in a sound insulation cabin during the experiments, for ensuring no distraction from environmental noise. During the experiment there was no feedback provided and no action required except for concentrated listening. Subjects were

positioned in a comfortable chair and asked to either stare at one point or keep their eyes closed during the experiment. The experiment's auditory stimulus was a a standard classical cadence consistent of three kinds of different endings. See Figure 1a for details.

3 Classification Scheme

We adopted the pattern matching method described by Blankertz in [13] in combination with a linear discriminant analyses (LDA). It showed best performance for detection of ERPs as can be seen in [14]. The data recorded was bandpassfiltered with a fast Fourier transform between 0.5 and 15 Hz and downsampled to 200 Hz. Then epochs of 1 second length were extracted after the marker indicating that the last chord of a cadence was initiated. Features were extracted for seven sequential time windows of 50 ms length, starting at 200 ms relative to the marker, by averaging the EEG within each window (see Figure 1b). This results in a $7 * 32 = 224$ dimensional feature space. We equalized the number of trials for each of the two extracted classes (standard and deviants) to 70 trials to avoid class specific overfitting of the LDA. The LDA was trained for each subject with a $10 \times (5, 10)$-crossvalidation (10 times repeated, 5 folded nested, 10 folded outer), and regularized with shrinkage [13] for dealing with the bad dimensionality ratio between features and trials.

4 Results

The relevance of features extracted at single electrode sites can be visualized by examining the Grandaverages of the respective LDA weights. Figure 1b shows Grandaverages for prominent features – those extracted at both peaks of the ERP. Data of two subjects was discarded because of technical problems with the recordings. The average classification accuracy across subjects is 75.4 %, with equally distributed errors between the classes.

5 Discussion and Conclusion

We found evidence that ERPs related to processing of syntactic and semantic errors can be detected by the presented approach. The accuracy of this detection is a little bit below that reported in studies from the visual domain. Further analyses of the experimental series could bring more insight into the underlying reasons. We hope that these results will be beneficial for the development of new auditory BCI-systems, as it can be used for enhancing communication and control in already existing and future BCI systems. In similarity to the visual domain, error responses can be used to correct unreliable classification and speed up communication or enhance control over the given system, like described in [10]. Furthermore, this passive BCI gives to some degree insight into internal interpretation of the subject. It allows us to see whether he or she accepts the presented cadence ending or not. This might be especially interesting for ongoing neuroscientific studies investigating cultural differences in music perception. In further studies we will focus on the development of new classification schemes that could improve the accuracy of the detection, as well as the evaluation whether this approach can be extended to other auditory signals, not embedded so strongly in music theory.

Acknowledgments

This work was supported by the ICT Collaborative Project BrainAble (247447) and the PhyPA project (TU Berlin).

References

[1] M. I. Posner and Y. Cohen. Components of visual orienting. *Attention and performance X: Control of language processes*, 32:531–556, 1984.

[2] M. S. Treder and B. Blankertz. (C) overt attention and visual speller design in an ERP-based brain-computer interface. *Behavioral and Brain Functions*, 6(1):28, 2010.

[3] P. Brunner, S. Joshi, S. Briskin, J. R. Wolpaw, H. Bischof, and G. Schalk. Does the P300 speller depend on eye-gaze? In *Integrating brain-computer interfaces with conventional assistive technology*, TOBI-Workshop, Graz, 2010.

[4] F. Aloise, I. Lasorsa, F. Schettini, A. Brouwer, D. Mattia, F. Babiloni, S. Salinari, M. Marciani, and F. Cincotti. Multimodal stimulation for a P300-based BCI. *International Journal of Bioelectromagnetism*, 9(3):128–130, 2007.

[5] M. Lehne, K. Ihme, A. M. Brouwer, J. van Erp, and T. O. Zander. Error-related EEG patterns during tactile human-machine interaction. In IEEE Computer Society Press, editor, *Proceedings of ACII 2009*, Los Alamitos, CA, 2009.

[6] F. Nijboer, A. Furdea, I. Gunst, J. Mellinger, D. J. McFarland, N. Birbaumer, and A. Kübler. An auditory brain-computer interface (BCI). *Journal of neuroscience methods*, 167, 1:43–50, 2008.

[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), 2010.

[8] T. O. Zander and C. Kothe. Towards passive brain–computer interfaces: applying brain–computer interface technology to human–machine systems in general. *Journal of Neural Engineering*, 8:025005, 2011.

[9] T. O. Zander, C. Kothe, S. Welke, and M. Rötting. Enhancing human-machine systems with secondary input from passive brain-computer interfaces. In *Proceedings of the 4th International BCI Workshop & Training Course. Graz University of Technology Publishing House, Graz*, 2008.

[10] T. O. Zander, C. Kothe, S. Jatzev, and M. Gärtner. Enhancing human–computer interaction with input from active and passive brain–computer interfaces. In D. Tan and A. Nijholt, editors, *The Human in Brain–Computer Interfaces and the Brain in Human–Computer Interaction*, pages 24–29. Springer, 2010.

[11] P. W. Ferrez and J. del R. Millán. Error-related EEG potentials generated during simulated brain-computer interaction. *IEEE Trans Biomed Eng.*, 55(3):923–9, 2008.

[12] B. Blankertz, G. Dornhege, C. Schäfer, R. Krepki, J. Kohlmorgen, K. R. Müller, V. Kunzmann, F. Losch, and G. Curio. Boosting bit rates and error detection for the classification of fast-paced motor commands based on single-trial EEG analysis. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11(2):127–131, 2003.

[13] B. Blankertz, S. Lemm, M. Treder, S. Haufe, and K. R. Müller. Single-trial analysis and classification of ERP components-a tutorial. *NeuroImage*, 2010.

[14] T. O. Zander, K. Ihme, M. Gärtner, and M. Rötting. A public data hub for benchmarking common brain–computer interface algorithms. *Journal of Neural Engineering*, 8:025021, 2011.

Calibration of the P300 BCI with the Single-Stimulus Protocol

S. L. Shishkin[1,2], A. A. Nikolaev[2], Y. O. Nuzhdin[2], A. Y. Zhigalov[1], I. P. Ganin[1,2], A. Y. Kaplan[1,2]

[1]Lomonosov Moscow State University, Faculty of Biology, Laboratory for Neurophysiology and Neuro-Computer Interfaces, Moscow, Russia
[2]National Research Nuclear University MEPhI, Moscow, Russia

sergshishkin@mail.ru

Abstract

For successful calibration of a P300 based brain-computer interface the user should not attend non-target stimuli. However, non-target stimuli automatically draw user's attention. To overcome this problem, we propose the use of the single-stimulus paradigm for calibration, at least in fresh users. In this simple paradigm targets are presented in absence of non-targets.

In the current study classifiers were trained on recordings made with either the single-stimulus protocol or with the usual P300 BCI protocol in the same participants, and applied online to detect an attended cell in a 3×3 matrix. Similar accuracy was observed for the new and the standard classifiers. This result are promising, as the single-stimulus paradigm should be easily understandable even without any previous BCI experience.

1 Introduction

In a typical visual P300 BCI discrimination between different possible inputs is based on the user's attention to one stimulus (the target) and his/her inattentiveness to the others. However, the non-target stimuli, acting as distractors, can automatically draw attention and evoke a brain response similar to the responses to targets [1].

This problem is especially serious during calibration of the BCI, when the user has no feedback showing whether the task is fulfilled correctly or not. The user's task is different from everyday experience, thus, the distracting power of the non-target flashes can be especially strong in a person without previous P300 BCI experience. Moreover, fresh users may easily misunderstand their task. If the user is a paralyzed patient, detecting such problems and helping the user to overcome them may require too high level of skills from the personnel. If the user is a video game player, he/she may try their first calibration sessions without getting well into all the details of the manual. Hence, there is a need for making the calibration protocol more robust in the face of real life situations.

Frye et al. [2] reported that suppressing stimuli at locations surrounding the target leads to the improvement of the calibration results (i.e., the accuracy of the subsequent BCI use). At further positions the non-target stimuli still were presented, so they, in principle, could capture attention.

The most radical solution can be removing the non-target stimuli completely from the calibration protocol. In addition to the absence of the distracting non-targets, this design and the related instructions could be much easier understood by the subjects, as they are required only to attend to stimuli which are themselves capturing their attention.

Although looking, for the first glance, not acceptable in the framework of the P300 BCI, this approach has a basis in psychophysiological studies, were the "single-stimulus" paradigm was found comparable to the standard oddball paradigm in its capacity to elicit the P300 wave [3, 4]. In our studies of modifications of the P300 BCI stimuli design the target-non-target difference in the P300 and occipital N1 were almost the same in single-stimulus and in row/column stimulation in a 3×3 matrix [5]. In addition, in a P300 BCI puzzle game proposed in [6] and implemented

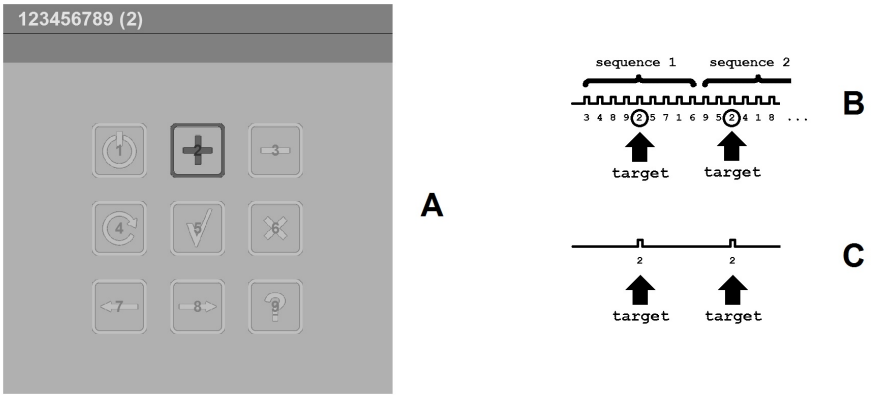


Figure 1: Stimuli in the single cell P300 BCI and in the new single-stimulus calibration protocol. (A) Stimuli matrix used in the current study, same in multistimulus (M-) and in the single-stimulus (S-)modes. (B) Stimuli time course in M-mode. (C) Stimuli time course in S-mode.

by A.Y.Z., S.L.S. and A.Y.K., the calibration was done with the usual protocol, but during the game the number of non-target items gradually decreased down to zero. It was tried by tens of novice players in public demonstrations, and the vast majority of them successfully assembled a full 4×4 puzzle matrix using it.

For classifier training, the EEG recorded at the positions where no stimuli are presented can be used instead of the EEG epochs following non-targets. The ground for this approach is provided by the fact that the amplitude of the responses to non-targets is usually low, and, therefore, the classifier's major task seems to be discrimination between the responses to targets and the background activity.

In the current study, we implemented, by modifying the open source BCI2000 software [7], and tested a single-stimulus modification of the P300 BCI calibration protocol.

2 Methods

Seven healthy participants (age 20–27) without previous experience with BCIs took part in the study after signing the informed consent. They viewed a 3×3 matrix with symbols and digits 1..9 overlapping the symbols (Figure 1 (A)).

In this study we compared the single-stimulus protocol with the single cell variation of the P300 BCI (e.g., [9]), in which the cells are highlighted independently of rows or columns, i.e., a stimulus is presented each time in one cell. Stimulation consisted of decreases of brightness (as, e.g., in [8]) of single cells for 128 ms with interstimulus interval 64 ms (see example in Figure 1 (A)).

We used two stimulus presentation modes: the multistimulus (M-) mode, and the single-stimulus (S-) mode. In M-mode, the stimuli were presented in all cells in a random order. A series of 9 stimuli presented once in each cell formed a sequence (Figure 1 (B)). In S-mode, the stimuli were presented only in the target cell (Figure 1 (C)). The distributions of target-to-target intervals (the time between the target and the preceding target) in S- and M-modes were approximately equal.

The current target cell number was indicated in the parentheses above the matrix (Figure 1 (A)). The order of targets was always 1,2, ..., 9; this sequence formed one block. The participants were asked to find in the matrix the digit shown in the parenthesis and to count silently "blinks" of the cell. Four stimuli sequences were presented per target in M-mode. In S-mode, four stimuli were presented per target. In the calibration phase, two blocks were presented in S-mode and two blocks in M-mode (participant #7 received three blocks per mode). In the test phase, all blocks were in M-mode.

Participant #	M	S	M*	S*	Sign(S-M)*	Abs(S-M)*
1	83	72	86	72	-	14
2	94	100	97	92	-	5
3	93	93	91	83	-	8
4	81	81	76	85	+	9
5	67	63	67	65	-	2
6	67	74	74	67	-	7
7	85	89	89	91	+	2
Mean (SD)	81 (11)	82 (13)	83 (11)	79 (11)		6 (4)
Median	83	81	86	83		7

Table 1: Accuracy (in %) for classifiers M (standard) and S (new). * - offline accuracy, which corresponded to the results applying the given classifier both to the data to which it was applied online and to the rest of the data.

EEG was recorded at Cz, Pz, PO7, PO8, O1, O2 against a joint reference at the earlobes with 250 Hz sampling rate. The data were low pass filtered, decimated down to 20 Hz and segmented in 50..750 ms epochs relative to stimulus onset. Channel amplitude data in each epoch were concatenated and formed a feature vector. Fisher Linear Discriminant Analysis was applied to these vectors for computing the classifier weights in S- and M-modes separately. We will further refer to the classifiers using these sets of weights as to S and M classifiers, respectively. In the case of S classifier training, the "non-target" epochs corresponded not to actually presented stimuli but to the positions at which they would be presented in M-mode.

In the test stage of the experiment, the task was the same as in the calibration phase, but this time the index of the cell recognized as attended was typed in the line below the target line. Thus, the participants could see whether they "typed" the digits correctly or not. M and S classifiers were randomly assigned to the blocks. The participants were not aware of which classifier was used in each block. Participants #1 and #2 received two blocks per classifier, typing 18 digits. The other five participants received three blocks per classifier and typed 27 digits. The S and M classifiers were also applied offline to the data from the same participants' test blocks recorded while the opposite types of classifiers (M and S, respectively) were applied in the online classification. This way, the size of the data used for computing each classifier's accuracy was doubled.

3 Results

Participant accuracies in the online and offline tests using both classifiers is shown in Table 1. On average, this group showed very similar accuracy for the standard and new protocols. The new protocol provided slightly lower accuracy comparing to the standard one, however, the difference was not significant according to Student's test for paired data ($t(6) = 1.27$, $p = 0.25$).

4 Discussion

The preliminary results obtained in this study are fully in line with our expectation that the P300 BCI accuracy would not deteriorate significantly if the classifier is trained using the single-stimulus protocol. They were obtained using the single cell variant of the P300 BCI and more testing is needed to determine if the single-stimulus protocol is also applicable to the more widely used row-column or new checkerboard protocols [10]. However, the single cell variant of the P300 BCI is also important, as being most practical when the number of commands is small (e.g., in wheelchair or robot arm control panels, or in video games).

In this study, we used relatively small number of trials for both classifier training and for online classification, to imitate regimes where the speed is more important than accuracy (e.g., in games). Additional testing might be needed to check the ability of the new protocol to provide

high accuracy under conditions when higher number of stimuli repetitions is allowed. Attempts can be made to significantly reduce the trial-to-trial intervals for shortening time for calibration.

In addition, the similarity between the performance of classifiers trained in single-stimulus and in multistimulus modes suggest that single-stimulus paradigm should be also studied as a possible online classification paradigm, e.g., for BCI switches.

5 Conclusion

Our first results of testing the single-stimulus calibration protocol suggest that it may have a significant value for calibration of the P300 BCI, at least when it is used by fresh BCI users.

6 Acknowledgments

This study was partly supported by the Federal Targeted Program "Scientific and Scientific-Pedagogical Personnel of Innovative Russia in 2009-2013" (contract P1087). The authors thank Andrey R. Nikolaev and the anonymous reviewers for their useful comments and suggestions.

References

[1] R. Fazel-Rezai. Human error in P300 Speller paradigm for brain-computer interface. *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, 2516-2519, 2007.

[2] G. E. Frye, C. K. Hauser, G. Townsend, and E. W. Sellers. Suppressing flashes of items surrounding targets during calibration of a P300-based brain-computer interface improves performance. *Journal of Neural Engineering*, 8:025024, 2011.

[3] J. Polich and M. R. D. Heine. P300 topography and modality effects from a single-stimulus paradigm. *Psychophysiology*, 33:747-752, 1996.

[4] C. J. Gonsalvez and J. Polich. P300 amplitude is determined by target-to-target interval. *Psychophysiology*, 39:388-396, 2002.

[5] I. P. Ganin. The N1 component of brain potentials and the spatial factors in the P300 brain-computer interface. *Proc. of the XIV young scientist conference on the physiology of higher nervous activity and neurophysiology. 21-22 Oct. 2010, IHNA RAS*, p. 37, 2010 (in Russian).

[6] A. J. Kaplan and S. V. Logachev. Game and method of its playing. *Patent RU 2406554 C1*, 2009.

[7] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw. BCI2000: a general-purpose brain-computer interface (BCI) system. *IEEE Trans. Biomed. Eng.*, 51:1034-1043, 2004.

[8] M. Salvaris and F. Sepulveda. Visual modifications on the P300 speller BCI paradigm. *Journal of Neural Engineering*, 6:046011, 2009.

[9] B. Rebsamen, C. Guan, H. Zhang, C. Wang, C. Teo, M. H. Ang, and E. Burdet. A brain controlled wheelchair to navigate in familiar environments. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 18:590-598, 2010.

[10] 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*, 121:1109-1120, 2010.

Subject-Specific Selection of the Oxygenation Parameter Boosts Classification Accuracy of NIRS Signals

M. Stangl[1], C. Neuper[1,2]

[1]Department of Psychology, Section Neuropsychology, University of Graz, Austria
[2]Institute for Knowledge Discovery, BCI Lab, Graz University of Technology, Austria

matthias.stangl@uni-graz.at, christa.neuper@uni-graz.at

Abstract

Over the last decades near-infrared spectroscopy (NIRS) became an established neuroimaging technology for non-invasive optical measurement of functional brain signals. NIRS measures changes in the two blood oxygenation parameters oxygenated haemoglobin (oxy-Hb) and deoxygenated haemoglobin (deoxy-Hb). The acquisition of these parameters enables an assessment of localized changes in brain activation. More recently, an increasing number of studies dealt with the suitability of NIRS signals in brain-computer interface (BCI) systems. However, many of these studies focused on the investigation and classification of only oxy-Hb and neglected the possibility that usage of deoxy-Hb might increase the classification accuracy rates. Our study aimed to examine the suitability of both oxy-Hb and deoxy-Hb for NIRS signal classification. We carried out independent classification analyses for both oxygenation parameters on the NIRS signals of 22 subjects. A comparison between the resulting accuracy rates revealed that neither of the two parameters leads to clearly better results. Moreover, we found notable inter-individual differences: each of the two parameters may have advantage over the other one, depending on the specific subject. Consequently, a subject-specific selection of the oxygenation parameter significantly increases the accuracy of NIRS signal classification.

1 Introduction

Near-infrared spectroscopy (NIRS) is a non-invasive neuroimaging technology to measure functional brain activation. Near-infrared light is emitted through the scalp into cortical areas. The amount of light that is not lost due to absorption and scattering effects follows a crescent-shaped path from light sources to detectors. Measuring the attenuation of the near-infrared light enables a calculation of the blood oxygenation parameters oxygenated haemoglobin (oxy-Hb) and deoxygenated haemoglobin (deoxy-Hb) based on their different characteristic absorption specta [1]. Localized concentration changes of oxy-Hb and deoxy-Hb in cortical regions have been identified as typical indicators of regional brain activation [2].

Over the past decade, the applicability of NIRS in brain-computer interface (BCI) systems has been examined (for a review see [3]). However, as oxy-Hb is noted to exhibit larger and more evident signal changes than deoxy-Hb, numerous studies predominantly used oxy-Hb for the purpose of NIRS signal classification (e.g. [4–7]).

For the present investigation we acquired NIRS signals of subjects while they performed two mental tasks, namely motor imagery and mental arithmetic. The signals of both oxy-Hb and deoxy-Hb were classified independently. A comparison between the accuracy rates of separate classifications aims to examine if either of the two parameters leads to better classification results and, consequently, is more suitable for BCI purposes. Furthermore, we investigated whether the classification accuracy rates might be increased by a subject-specific selection of the oxygenation parameter.

2 Methods

2.1 Subjects and Experimental Paradigm

The subject sample consisted of 22 healthy subjects (11 female; mean age = 25.7±2.6 years). They performed 25 trials of a motor imagery task (imagination of right-handed ball squeezing) and 25 trials of a mental arithmetic task (serial subtractions of a one-digit number from a two-digit number) in a pseudo-random order. Every trial started with a 3 sec baseline-period followed by a 13.5 sec task-period. The inter-trial interval was set to a random duration between 10 and 12 sec.

2.2 Data Acquisition

We acquired NIRS signals with a continuous wave system (ETG-4000, Hitachi Medical Co., Japan) that uses two wavelengths to obtain concentration changes of oxy-Hb and deoxy-Hb. The NIRS signal was recorded at 24 channels with a sampling frequency of 10 Hz. We used two optode-grids, both of which covered a quadratic area of about 36 cm^2. One optode-grid (12 channels) was adjusted around the EEG electrode-position C_3 (Int. 10-20 System). These channels covered the left hemisphere above primary motor, pre-motor, supplementary-motor and primary somatosensory areas. Another optode-grid was adjusted posterior to electrode-position F_{PZ}, to cover frontopolar and dorsolateral prefrontal areas of both hemispheres.

2.3 Signal Processing and Classification

The influence of pulse on the NIRS signal was reduced by a 1 sec moving-average filter. Baseline-drifts were derived by a 15 sec moving-average filter and subtracted from the signal. Subsequently, the signal was baseline-corrected for every trial.

We carried out separate classification analyses for oxy-Hb and deoxy-Hb, respectively. The signal during task performance was classified by Fisher's linear discriminant analysis (LDA), in order to discriminate between the two mental tasks (motor imagery vs. mental arithmetic). The first five seconds of the task-period were excluded from analyses. Since the hemodynamic response has a typical response latency of a few seconds [8], we assumed that this early time period after task-onset has no practical relevance for classification purposes. Independent LDAs were applied to the signal amplitudes of every single datapoint between second 5.1 and 13.5 of the task-period.

For the feature-selection, firstly we classified the signal of every single channel separately. The Leave-One-Out-Method (LOOM) was applied to estimate the cross-validated classification accuracy for every classified datapoint. We averaged the accuracy rates of all classified datapoints to get an indicator of every single channel for its suitability as a classification feature. Secondly, based on this indicator, we applied a sequential forward selection algorithm (as described in [9]) individually for every subject, to identify an optimal subset of three channels for classification. The signal of these three channels was used as a three-dimensional LDA feature-set.

In the following sections we will differentiate between two accuracy values, which both resulted from the classification and LOOM cross-validation of the selected three-dimensional feature-set: 1) The *maximum-accuracy* rate of a specific datapoint within the task-period, namely at the point of time when the largest accuracy value could be achieved. 2) The *mean-accuracy* rate, averaged over the whole task-period except its first five seconds.

2.4 Statistical Analysis

We conducted two one-factorial ANOVAs to compare accuracy rates of all subjects for three conditions, separately for *mean-accuracy* and *maximum-accuracy* values. Conditions one and two were the accuracy rates from the classification of the oxy-Hb and deoxy-Hb signal, respectively. Condition three was defined subject-specifically: either the oxy-Hb or the deoxy-Hb accuracy rate was used, depending on which one of the parameters led to a higher accuracy rate for the specific subject. We applied Greenhouse-Geisser corrections for both ANOVAs, to account for violations in the sphericity assumption.

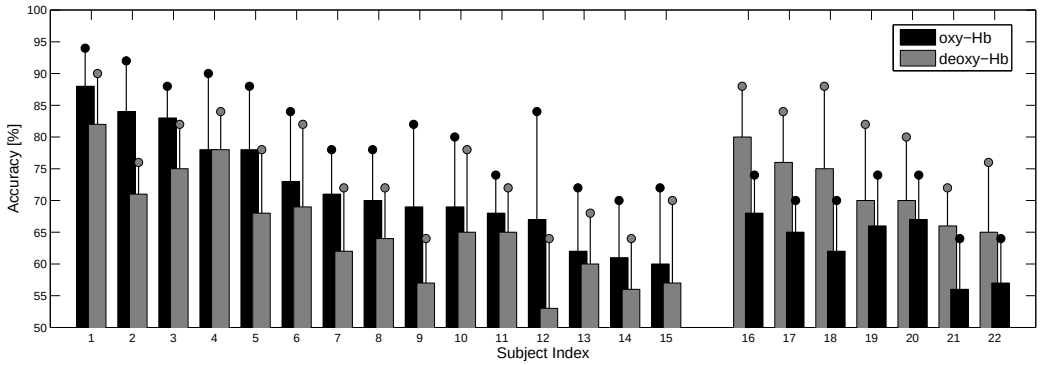


Figure 1: Classification accuracy rates. For subjects 1 to 15 (displayed in the left plot area) oxy-Hb classification (black bars) led to higher accuracy rates, whereas deoxy-Hb (gray bars) was the better parameter for subjects 16 to 22 (displayed in the right plot area). The height of a bar indicates the *mean-accuracy* rate, while the small circle indicates the *maximum-accuracy* rate.

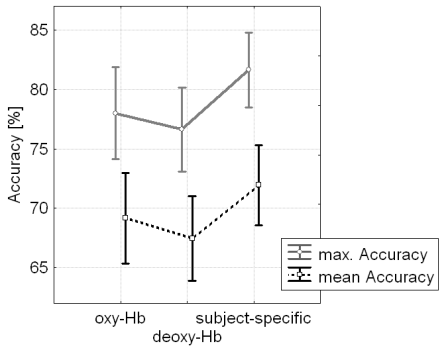


Figure 2: Classification accuracy rates averaged over 22 subjects. The gray solid line shows the *maximum-accuracy* rate, while the black dashed line shows the *mean-accuracy* rate. The vertical error bars visualize the standard error.

3 Results

For 15 subjects higher accuracy rates were derived from classification of oxy-Hb compared to deoxy-Hb, whereas for seven subjects deoxy-Hb was identified to be a better parameter for classification. The *maximum-accuracy* rates of all subjects were above the statistical chance level ($\alpha = 0.05$) of 64 % (according to [10]) for both oxy-Hb and deoxy-Hb. The *mean-accuracy* rates exceeded statistical chance level for 16 (oxy-Hb) and 15 subjects (deoxy-Hb), respectively. Classification accuracy rates for every single subject are illustrated in Figure 1.

Averaged over all 22 subjects, rates of 78 % (oxy-Hb) and 77 % (deoxy-Hb) could be ascertained for the *maximum-accuracy* at the best classification time points. They ranged from 64 % to 94 % (oxy-Hb) and 64 % to 90 % (deoxy-Hb). The *mean-accuracy* rates within the task-period were 69 % (oxy-Hb) and 67 % (deoxy-Hb), ranging from 56 % to 88 % (oxy-Hb) and 53 % to 82 % (deoxy-Hb).

The statistical analysis revealed significant differences between the three conditions (oxy-Hb, deoxy-Hb and subject-specific) for both ANOVAs (*maximum-accuracy*: $F_{2,21} = 4.89$, $p = 0.030$; *mean-accuracy*: $F_{2,21} = 6.01$, $p = 0.018$). Posttests revealed significantly higher accuracy rates for the subject-specific oxygenation parameter compared to both oxy-Hb and deoxy-Hb. However, no difference between oxy-Hb and deoxy-Hb accuracy rates was found. Figure 2 shows, averaged over all 22 subjects, the *maximum-accuracy* and *mean-accuracy* rate as well as the accuracy rate that was determined by a subject-specific selection of the oxygenation parameter.

4 Discussion and Conclusion

The results of the present study emphasize the substantial potential of NIRS as signal acquisition technology for BCI systems. Classification of the oxygenation parameters oxy-Hb and deoxy-Hb in order to discriminate between motor imagery versus mental arithmetic tasks led to notable accuracy rates.

Moreover, the results underline the importance of both oxy-Hb and deoxy-Hb for the classification of NIRS signals. It is a common approach in NIRS-BCI research to focus solely on the oxygenation parameter oxy-Hb, because it is presumed to reflect changes in regional cerebral blood flow better than deoxy-Hb [6]. By contrast, the present results do not support this assumption. Our findings suggest that neither the exclusive utilization of oxy-Hb nor deoxy-Hb leads to better classification results. However, we found remarkable subject-specific differences between the suitability of each parameter. While for a number of subjects the usage of oxy-Hb leads to higher classification accuracy, for some other subjects deoxy-Hb is better suited for the purpose of classification. Consequently, the subject-specific selection of either oxy-Hb or deoxy-Hb significantly boosts classification accuracy rates.

5 Acknowledgments

This work is supported by the NeuroCenter Styria (Land Steiermark, PN 4055).

References

[1] F. F. Jobsis. Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters. *Science*, 198:1264–1267, 1977.

[2] A. Maki, Y. Yamashita, Y. Ito, and H. Koizumi. Spatial and temporal analysis of human motor activity using noninvasive NIR topography. *Medical Physics*, 22:1997–2005, 1995.

[3] R. Sitaram, A. Caria, and N. Birbaumer. Hemodynamic brain-computer interfaces for communication and rehabilitation. *Neural Networks*, 22:1320–1328, 2009.

[4] 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*, 4:219–226, 2007.

[5] M. Naito, Y. Michioka, K. Ozawa, Y. Ito, M. Kiguchi, and T. Kanazawa. A communication means for totally locked-in ALS patients based on changes in cerebral blood volume measured with near-infrared light. *IEICE Trans. on Information and Systems*, 90:2028–1037, 2007.

[6] I. Nambu, R. Osu, M. Sato, S. Ando, M. Kawato, and E. Naito. Single-trial reconstruction of finger-pinch forces from human motor-cortical activation measured by near-infrared spectroscopy (NIRS). *NeuroImage*, 47:628–637, 2009.

[7] K. Utsugi, A. Obata, H. Sato, R. Aoki, A. Maki, H. Koizumi, K. Sagara, H. Kawamichi, H. Atsumori, and T. Katura. GO-STOP control using optical brain-computer interface during calculation task. *IEICE Transactions on Communications*, E91.B:2133–2141, 2008.

[8] S. C. Wriessnegger, J. Kurzmann, and C. Neuper. Spatio-temporal differences in brain oxygenation between movement execution and imagery: A multichannel near-infrared spectroscopy study. *International Journal of Psychophysiology*, 67:54–63, 2008.

[9] I. Guyon, S. Gunn, M. Nikravesh, and L. A. Zadeh. *Feature Extraction: Foundations and Applications*. Berlin: Springer, 2006.

[10] G. R. Müller-Putz, R. Scherer, C. Brunner, R. Leeb, and G. Pfurtscheller. Better than random? A closer look on BCI results. *International Journal of Bioelectromagnetism*, 10:52–55, 2008.

Classification of Focal Frontal Oxyhemoglobin Responses During Mental Arithmetic

G. Bauernfeind[1], R. Scherer[1], G. Pfurtscheller[1], C. Neuper[1,2]

[1]Institute for Knowledge Discovery, Laboratory of Brain-Computer Interfaces, Graz University of Technology, Graz, Austria
[2]Department of Psychology, University of Graz, Graz, Austria

g.bauernfeind@tugraz.at

Abstract

Near-infrared spectroscopy (NIRS) is a non-invasive optical technique for the assessment of functional activity in the human brain. Task-specific hemodynamic responses, i.e., concentration changes of oxy- and deoxyhemoglobin (oxy-Hb, deoxy-Hb) during cognitive, visual, visuo-motor, and motor tasks can be detected and used for brain-computer interfacing (BCI). For this study we used the data of a previous investigation on simple mental arithmetic (MA) tasks where we found antagonistic activation patterns (focal bilateral increase of oxy-Hb in the dorsolateral prefrontal cortex (DLPFC) in parallel with a decrease in the medial area of the anterior prefrontal cortex (APFC) in eight of ten subjects. The oxy-Hb responses from the eight subjects were used to perform a cue-paced BCI off-line simulation. Therefore we searched for the best antagonistic feature combination from the left or right DLPFC and the APFC and compared it to individual features from the same regions. Our results indicate that the use of antagonistic features significantly increases the classification accuracy compared to individual features, that a mean classification accuracy of around 80 % by using antagonistic hemodynamic response patterns is possible, and that the use of these patterns may be a suitable control strategy for optical BCIs.

1 Introduction

Brain-computer interface (BCI) systems give users the possibility to communicate or interact through thought processes alone [1,2]. The required methods are normaly based on the detection of neural activity by electroencephalography (EEG), magnetoencephalography (MEG), and recently, with functional magnetic resonance imaging (fMRI) [3, 4]. However, the use of these techniques may be restricted due to system size, ambient requirements, and user preparation (conductive gel, sensor placement, ...). As a more promising method the near-infrared spectroscopy (NIRS) technique can also be used for BCI systems [5]. However, there are some drawbacks to using NIRS for optical BCI applications. Beside the temporal resolution of the hemodynamic response - in the range of several seconds - a robust single-trial classification is essential. The identification of brain patterns that most naïve users can reliably generate and that are stable over time may significantly counteract this issue.

In [6] we investigated changes of oxy- and deoxyhemoglobin (oxy-Hb, deoxy-Hb) during the performance of simple mental arithmetic (MA) tasks in ten subjects. We found that eight out of ten subjects displayed a significant focal bilateral increase of oxy-Hb accompanied by a deoxy-Hb decrease in the dorsolateral prefrontal cortex (DLPFC). In parallel, they showed a significant decrease of oxy-Hb, accompanied by a deoxy-Hb increase, in most channels (the largest and most stable decreases are localized around the FP1 position) overlaying the medial area of the anterior prefrontal cortex (APFC). Such hemodynamic responses could be explained in the context of "focal activation/surround deactivation" and can be seen as antagonistic activation patterns [6]. We

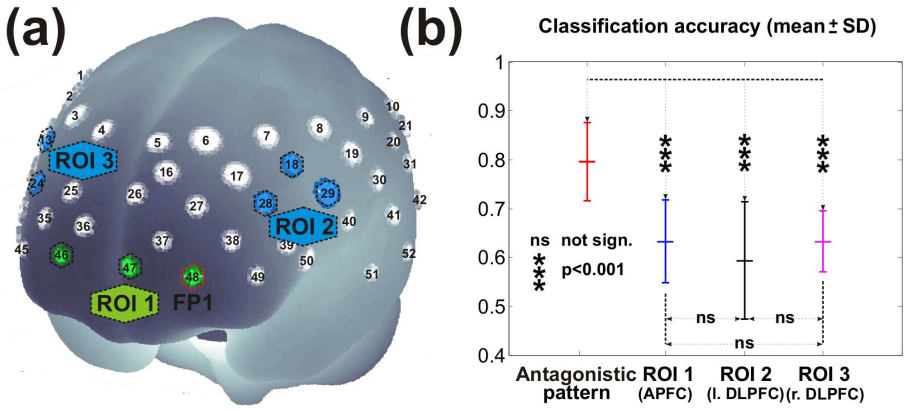


Figure 1: (a) Projections of the 52 NIRS channel positions (3×11 grid) on the cortical surface. Positions are overlaid on an MNI-152 compatible canonical brain which is optimized for NIRS analysis [7]. Also indicated are the regions of interest (ROI). (b) Significant contrasts of the classification accuracy between the antagonistic and individual features.

hypothesized that these focal antagonistic hemodynamic response patterns during MA may be reliably detected by placing only two NIRS sensors over the prefrontal cortex. Here we evaluate the above hypothesis by means of cue-based BCI off-line simulations.

2 Methods

2.1 Subjects, Experimental Paradigm and Data Collection

The investigations were carried out on a group of eight healthy subjects (three males and five females, all right-handed) aged 26.0±2.8 years (mean±SD). The subjects were asked to perform a series of MA tasks. During the task they had to sequentially subtract a one-digit number from a two-digit number (e.g., $97 - 4 = 93$, $93 - 4 = 89$, ...; the initial subtraction was presented visually on a monitor) as quickly as possible for 12 seconds, afterwards a 28 s resting period was given. Subjects performed 3 or 4 runs resulting in 18 or 24 trials per class, respectively.
A continuous wave system (ETG-4000, Hitachi Medical Co., Japan) was used to record brain oxygenation. The multi-channel system measures the change of oxy-Hb and deoxy-Hb in the unit of mM*mm and consists of 16 photo-detectors and 17 light emitters (3×11 grid), resulting in a total of 52 channels. The lowest line of channels was arranged along the FP1- FP2 line of the international EEG 10-20 system, thus the grid covered the whole frontal lobe (for further details see [6]).

2.2 Data Analysis

The mean task-related concentration changes of oxy-Hb and deoxy-Hb, referenced to a 10 s baseline interval prior to the task (seconds -10 to 0) were calculated for each channel after removing baseline drifts by using a 0.01 Hz high-pass filter (for further details see [8]). To capture the antagonistic oxy-Hb patterns three regions of interest (ROI): ROI_1 over APFC, ROI_2 over the left DLPFC, and ROI_3 over right the DLPFC (see Figure 1 a)) were defined according to the results of [6].
For classification a Fisher's linear discriminant analysis (LDA) classifier was used. To evaluate the LDA generalization, data recorded from each subject was split into a training and evaluation set. The former, consisting of trials 1 to 10 (16), was used to train and test the discriminative power of oxy-Hb feature combinations selected from the different ROIs. The best performing features were selected and used to train the LDA. The evaluation set, composed of the last eight trials, was then used to assess the performance of the trained LDA.

Subj.	Anta. oxy-Hb feature	ROI_1 oxy-Hb feature	ROI_2 oxy-Hb feature	ROI_3 oxy-Hb feature
AG1	68.75	62.50	56.25	56.25
AH3	**87.50**	62.50	50.00	62.50
AK4	**75.00**	62.50	50.00	68.75
AK8	**87.50**	**81.25**	50.00	62.50
AL2	**81.25**	50.00	68.75	56.25
AL6	68.75	62.50	50.00	62.50
AM4	**87.50**	62.50	**81.25**	62.50
AL2	**81.25**	62.50	68.75	**75.00**
mean	79.69	63.28	59.38	63.28
SD	8.01	8.48	12.05	6.19

Table 1: Classification accuracies (in %) for the antagonistic (Anta.) and individual features for all subjects. Bold numbers indicate classification accuracies above the chance level (71.9 % for 8 trials) [9].

Concentration changes (1-s mean) of the oxy-Hb response in the time windows from second 10 to 14 were labeled as class MA. Samples at seconds 26 to 30 were labeled as class REST. Features consist of an individual oxy-Hb value of one channel at a fixed time. For each subject, independent classifiers were trained and tested (leave-one-out cross validation) with individual oxy-Hb responses for each possible combination. Exhaustive Search, i.e., all possible feature combinations were evaluated, was used in the above procedure to identify the best performing antagonistic feature combination (ROI_1, ROI_2) or (ROI_1, ROI_3).

3 Results

The best performing classifiers calculated from the training set were used to compute an off-line simulation with the evaluation set (8 trials per class, see Table 1). Six out of the 8 subjects (75 %) performed better than the chance level (71.9 % for 8 trials) [9] using antagonistic patterns. In contrast only one subject performed better than random when using individual features from ROI_1, ROI_2, or ROI_3, respectively. An analysis of variance (ANOVA) and the Newman-Keuls posttest revealed that antagonistic features performed significantly better than individual features ($F_{(3/21)} = 8.74$; $p < 0.001$; Figure 1 b).

4 Discussion

The aim of the study was to underpin the hypothesis of the usefulness of antagonistic oxy-Hb patterns (found in eight of ten subjects in [6]) in the context of single trial classification for optical BCIs. The given results state that only two NIRS sensors placed over predefined brain areas, i.e., left or right DLPFC and APFC, respectively, may significantly increase the performance of optical BCIs compared to the more common approach to use only one sensor (e.g., [10, 11]). By using the best antagonistic oxy-Hb features 75 % of the subjects reached accuracies above the chance level and a mean±SD classification accuracy of 79.69±8.01 % (Table 1) was computed. In contrast, individual features performed worse (classification accuracies of 63.28±8.48 % (ROI_1), 59.38±12.05 % (ROI_2) and 63.28±6.19 % (ROI_3), Table 1). In each case only one subject reached accuracies above chance level.

5 Conclusion

The results of our study suggest that the use of antagonistic hemodynamic features may significantly increase the classification accuracy compared to individual features from the same regions. The off-line simulation results confirmed our hypothesis that two prefrontal NIRS channels can capture antagonistic hemodynamic patterns during a mental arithmetic task that can be detected reasonably well without the need of time consuming user-adaptation. In combination with a self paced paradigm the use of antagonistic pattern may be an important contribution for optical BCIs.

6 Acknowledgments

The authors' BCI research has been supported by the EU project PRESENCCIA (IST-2006-27731), "Land Steiermark" (project A3-22. N-13/2009-8) and the Neuro Center Styria (NCS) in Graz, Austria.

References

[1] 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):94–109, 2002.

[2] N. Birbaumer. Breaking the silence: brain-computer interfaces (BCI) for communication and motor control. *Psychophysiology*, 43(6):517–532, 2006.

[3] N. Birbaumer, C. Weber, C. Neuper, E. Buch, K. Haapen, and L. Cohen. Physiological regulation of thinking: brain-computer interface (BCI) research. *Prog Brain Res*, 159:369–391, 2006.

[4] R. Sitaram, A. Caria, and N. Birbaumer. Hemodynamic brain-computer interfaces for communication and rehabilitation. *Neural Netw*, 22(9):1320–1329, 2009.

[5] S. Coyle, T. Ward, C. Markham, and G. McDarby. On the suitability of near-infrared (NIR) systems for next-generation brain-computer interfaces. *Physiol Meas*, 25(4):815–822, 2004.

[6] G. Pfurtscheller, G. Bauernfeind, S. C. Wriessnegger, and C. Neuper. Focal frontal (de)oxyhemoglobin responses during simple arithmetic. *Int J Psychophysiol*, 76:186–192, 2010.

[7] A. K. Singh, M. Okamoto, H. Dan, V. Jurcak, and I Dan. Spatial registration of multichannel multi-subject fNIRS data to MNI space without MRI. *Neuroimage*, 27:842–841, 2005.

[8] G. Bauernfeind, R. Leeb, S. C. Wriessnegger, and G. Pfurtscheller. Development, set-up and first results for a one-channel near-infrared spectroscopy system. *Biomed Tech (Berl.)*, 53:36–43, 2007.

[9] G. R. Müller-Putz, R. Scherer, C. Brunner, R. Leeb, and G. Pfurtscheller. Better than random? A closer look on BCI results. *Int J Bioelectromagn*, 10:52–55, 2008.

[10] M. Naito, Y. Michioka, K. Ozawa, Y. Ito, M. Kiguchi, and T. Kanazawa. A communication means for totally locked-in ALS patients based on changes in cerebral blood volume measured with near-infrared light. *IEICE Trans Inf Syst*, E90-D:1028–1037, 2007.

[11] K. Sagara, K. Kido, and K. Ozawa. Portable single-channel NIRS-based BMI system for motor disabilities' communication tools. *Conf Proc IEEE Eng Med Biol Soc*, 10:602–605, 2009.

Towards a Brain Computer Interface-Based Rehabilitation: from Bench to Bedside.

F. Pichiorri[1], F. Cincotti[1], F. De Vico Fallani[1,2], I. Pisotta[3], G. Morone[4], M. Molinari[3], D. Mattia[1]

[1]Neuroelectrical Imaging and BCI Lab, Fondazione Santa Lucia, IRCCS, Rome, Italy
[2]Department of Human Phisiology, Sapienza University, Rome, Italy
[3]Laboratorio di Neuroriabilitazione Sperimentale, Fondazione Santa Lucia, IRCCS, Rome, Italy
[4]Movement and Brain Laboratory, Fondazione Santa Lucia, IRCCS, Rome, Italy

f.pichiorri@hsantalucia.it

Abstract

The application of Brain Computer Interface (BCI) technology in stroke rehabilitation represents one of the most challenging matters in the BCI field. With the aim of developing a specific BCI system for stroke rehabilitation of the upper limb we monitored the EEG sensorimotor reactivity to actual and imaged hand grasping in a group of stroke patients consecutively enrolled from a rehabilitation clinic. A subgroup of patients underwent a one-month BCI training with a system developed specifically for the purpose of stroke rehabilitation and installed in the rehabilitation hospital ward.

1 Introduction

One of the most recent and promising application fields of the Brain Computer Interface (BCI) technology is represented by the motor rehabilitation of stroke patients [1]. The practice of motor imagery (MI) has been suggested to improve motor recovery after stroke, by inducing use-dependent plastic changes in the lesioned hemisphere [2, 3]. In this respect, EEG-based BCI systems operated via MI appear a unique option to potentiate the restoration of motor function after stroke. Moreover, BCI technology could potentiate post-stroke rehabilitation by exploiting the neuroplasticity phenomena induced by the BCI training per se [4]. In an effort to deploy a practical EEG-based BCI system as an effective post-stroke rehabilitation training tool, it is crucial to first define patients' neurophysiological profiles; only EEG sensorimotor patterns harnessed to a "good" induced-neuroplasticity (e.g. patterns originated from the affected hemisphere) should be reinforced via a BCI training. Moreover, to effectively encourage training and practice the BCI design should incorporate principles of current rehabilitative settings, apt to stimulate patients' engagement during the exercise. We coped with the aforementioned points by means of a multimodal neurophysiological assessment of post-stroke movement-related brain activity (in 21 consecutively enrolled stroke patients). This preliminary step guided the application of a novel EEG-based BCI system specifically designed for the rehabilitation of the upper limb in stroke patients. The prototype system is currently installed in a rehabilitation hospital ward and allows stroke patients to perform daily sessions in which they practice MI of simple movements of their paralyzed hand, by controlling a visual representation of their own hands. Both clinical and neurophysiological assessment of the first 5 stroke patients who underwent a one-month BCI-training with this system revealed encouraging results for the future introduction of the BCI technology-assisted intervention in large scale clinical programs for stroke rehabilitation.

2 Methods

Data were collected from 21 patients with first ever monolateral stroke, consecutively enrolled from the Fondazione Santa Lucia where they were admitted for neurorehabilitation (mean age 58±16 y,

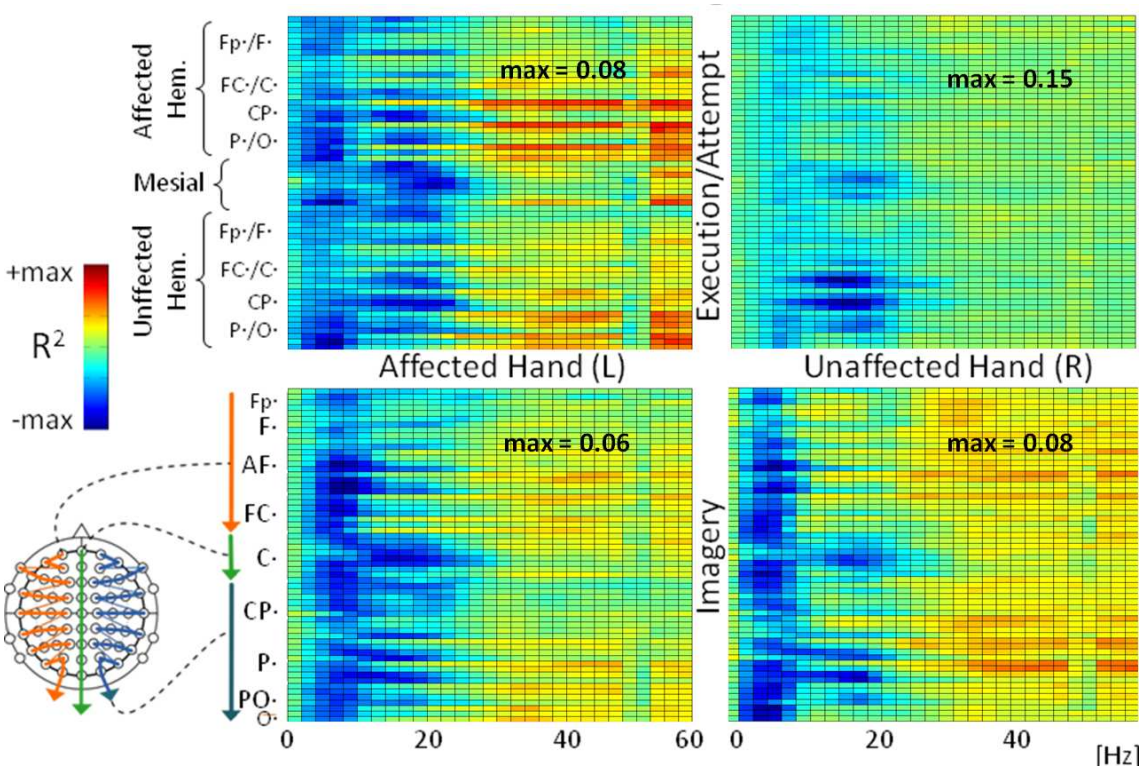


Figure 1: Spectral changes in the EEG oscillation during hand grasping movement and imagery in stroke patients. The channel–frequency matrices (horizontal and vertical axis, respectively) obtained by compiling the R-square values averaged across patients and relative to overt (upper panels) and covert (lower panels) motor tasks are shown. Note that before averaging, the electrode positions relative to each scalp side have been flipped in order to respect the non-homogeneity of patient lesion side (the top region of each panel represents the activity of the affected hemisphere; the unaffected hemisphere is represented in the bottom region; midline electrodes are represented between the two). The coloured bar codes for the decrease (blue) and increase (red) of the EEG spectral signal amplitude, quantified by the signed R-square values (indicated for each panel). On average, the overt and covert hand grasping performed with the unaffected hand (right panels) produced a decrease in the EEG alpha/beta power spectrum, with a lower magnitude for the covert condition. When patients were asked to focus on their affected hand (left panels) the attempted movement and the imagery were associated again with a decrease in the power spectrum with a lower magnitude as compared with the unaffected hand condition. These changes in the amplitude of the EEG oscillation were more broadly represented over the bilateral scalp centro-parietal areas, with a concomitant involvement of the midline frontal, central and parietal electrodes during imagery.

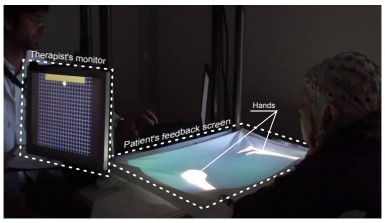


Figure 2: Training session with the novel EEG-based BCI system. In this session, two actors take part: the patient and the therapist. The patient is trained to gain control of the visual hand representation by imagining hand movements and receives as a feedback the congruent movements of the represented hand. The therapist is fed back with the real-time movement of a cursor on a screen that is actually controlled by the patient EEG relevant feature.

chronic/subacute 6/15, left/right lesion 7/14, subcortical/cortical 5/16). Scalp EEG potentials were collected from 61 positions and amplified by a commercial EEG system (band-pass 0.1–70 Hz; frequency sample 200 Hz). Patients were asked to either imagine (MI) or execute/attempt (ME) hand grasping with unaffected (UH) and affected hand (AH), being instructed by a visual cue. An offline analysis was performed to contrast the EEG signals relative to rest trials with those associated with motor task trials. All possible features in a reasonable range (i.e., 0–60 Hz in 2 Hz bins) were extracted and analyzed simultaneously. R-square values were compiled in a channel-frequency matrix and evaluated to identify theset of candidate features that separated best rest vs. a given motor task. Five of these patients also underwent a MI-based BCI training, during which they were asked to control the movement of a visual representation of their own AH by MI; training comprised 4 runs of 20 trials per session, 3 sessions per week, and lasted one month. During the training sessions patients were assisted by a therapist to whom the EEG features were fed back in real time together with bilateral EMG monitoring of muscles of the upper limb. Data acquisition, on-line EEG processing and feedback to the therapist was performed by using the BCI2000 software, while the feedback to the patient was provided with a software specifically designed in our laboratory.

3 Results

The averaged channel-frequency matrices are illustrated in Figure 1 R-square values were higher for the UH condition (0.15 ME, 15 subjects averaged; 0.08 MI, 18 subjects averaged) than the AH condition (0.08 ME, 14 subjects averaged; 0.06 MI, 20 subjects averaged). The prototype system developed for the rehabilitation of the upper limb in stroke patients is shown in Figure 2. Figure 3 shows the results of the BCI training of a representative stroke patient.

4 Discussion

The screening results indicated that EEG patterns observed in stroke patients with monohemispheric lesions partially preserve the movement and imagery-related desynchronization of the motor associated fronto-parietal alpha and beta rhythms. For the AH MI however, these patterns display a variable magnitude between patients and more spatially broad distribution, involving the contralateral unaffected hemisphere. This should be taken into account when considering BCI as a tool to engage neuroplasticity leading to a better post-stroke functional outcome. The results of the BCI training shown for one representative patient show that EEG features isolated from the screening session progressively shifted towards the affected hemisphere during training; an increase in performance was also observed, demonstrating how the training led to a progressive mastering of the BCI system.

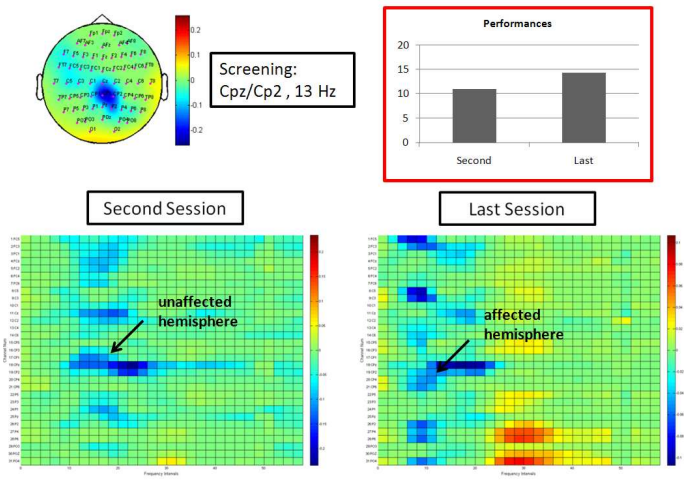


Figure 3: Results of BCI training in a representative stroke patient; the head topography on the left top shows the screening features extracted from the screening session and used as control features during training. BCI performances are shown in the histogram on the right top, expressed as average number of correct trials per run. The channel–frequency matrices of R-square values in the bottom part result from the contrast between the task and a baseline condition in the second and last training sessions.

5 Conclusion

The preliminary results of the BCI training with stroke patients are encouraging: first it has been shown that EEG features related to the MI task can be collected from the affected side of stroke patients and successfully adopted to control a BCI system; secondly, the analysis of EEG data showed that the affected hemisphere is progressively more involved during the BCI training.

Acknowledgments

This work is supported by the European ICT Programme Project FP7-224631(TOBI). 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] J. J. Daly and J. R. Wolpaw. Brain-computer interfaces in neurological rehabilitation. *Lancet Neurology*, 7(11):1032–1043, November 2008. PMID: 18835541.

[2] P. Langhorne, F. Coupar, and A. Pollock. Motor recovery after stroke: a systematic review. *Lancet Neurology*, 8(8):741–754, August 2009. PMID: 19608100.

[3] M. A. Dimyan and L. G. Cohen. Neuroplasticity in the context of motor rehabilitation after stroke. *Nature Reviews. Neurology*, 7(2):76–85, February 2011. PMID: 21243015.

[4] F. Pichiorri, F. De Vico Fallani, F. Cincotti, F. Babiloni, M. Molinari, S. C. Kleih, C. Neuper, A. Kübler, and D. Mattia. Sensorimotor rhythm-based brain-computer interface training: the impact on motor cortical responsiveness. *Journal of Neural Engineering*, 8(2):025020, March 2011. PMID: 21436514.

Evaluation of the OpenViBE P300-Speller in a Locked-In Patient

E. Maby[1,2], M. Perrin[1,2], D. Morlet[1,2], P. Ruby[1,2], O. Bertrand[1,2], S. Ciancia[3], N. Gallifet[3], J. Luauté[1,2,3], J. Mattout[1,2]

[1] INSERM U1028, CNRS UMR5292, Lyon Neuroscience Research Center, Lyon, France
[2]University Lyon 1, Lyon, France
[3]Hospices Civils de Lyon, Henry Gabrielle Hospital, Lyon, France
manu.maby@inserm.fr

Abstract

This study evaluates the efficacy of a P300-based brain–computer interface (BCI) implemented with the OpenViBE software, in a patient suffering from a locked-in syndrome (LIS). The highest accuracy achieved online reached 24.4 % and was obtained for the only session, out of four, for which both training and testing data were recorded on the same day. By contrast, other testing sessions, based on the same training, yielded accuracies no greater than chance level. Subsequent offline analysis and complementary data from an auditory oddball paradigm confirmed that the patient understood and performed the task correctly. Nevertheless and although we increased the offline training set with test data from each session, the overall accuracy reached 20.6 % only, which remains substantially lower than healthy subject's performance with the same protocol [1]. Altogether, our results suggest that poor performance coincide with low signal-to-noise ratio and high trial-wise variability. This might be explained by the difficulty for this patient to focus attention and to avoid disturbances caused by stimuli adjacent to the target.

1 Introduction

Locked-in patients are characterized by complete motor paralysis, except for eye movements, with intact cognition and sensation. A communication tool independent of muscle control would enable them to regain autonomy and be less dependent on others for communication. It has been shown that some severely disabled patients can communicate through the P300-based BCI [2]. In this paradigm, symbols are displayed on screen, in a matrix form, whose rows and columns are flashed alternatively. A P300 response is elicited whenever the row or column that contains the target is being flashed. The present study evaluates the efficacy of our recent implementation of the P300-speller in a LIS patient, using OpenViBE (http://openvibe.inria.fr) [3].

2 Methods

The patient is a 38 years old woman who sustained a locked-in-syndrome following a basilar artery thrombosis on March 1^{st} 2009. She was diagnosed with a complete tetraplegia with anarthria. Cerebral MRI revealed a pontine infarction with a complete interruption of the cortico-spinal tract as well as lesions of the cerebellar pedoncules and tegmentum nuclei. Speech and language therapy rehabilitation progressively established a yes/no communication code and developed a words spelling setting using eye closure in response to letter enumeration. The patient underwent cognitive function evaluation through auditory paradigms. First, a passive novelty oddball paradigm [4] highlighted normal sensory N1 responses to all stimuli and a large novelty P3, suggesting the absence of abnormality in automatic attention orienting to rare salient stimuli. Second, the patient was presented with a classical auditory oddball paradigm with both a passive (diverted attention) and active condition (attention oriented toward deviants). Like healthy subjects, patient showed a larger P300 response to deviant tones in the active condition. This result confirms that she understood the instructions and was able to engage in a voluntary, active and sustained task.

In the P300-speller experiment, EEG referenced to the nose was recorded from a 32-electrode Acticap system (Brain Products, Germany), following a standard extended 10-10 system placement. Impedances were kept below 10 kΩ and amplified analog signals were digitized at 100 Hz.
To make a selection from the 6×6 item matrix, the patient attends to the target symbol and counts how many times it is flashed. While visual stimulations are sent to a CRT (Cathode Ray Tube) screen in random order, a trigger (labeled according to the flashed row or column) is sent to the EEG amplifier via parallel port (jitter < 0.1 ms). In addition to intensity, symbol size is enhanced, which we proved yields larger P300 responses and higher classification rates [5]. The flash duration was set to 100 ms and the time between two flash onsets to 250 ms. We used sixteen repetitions per item, meaning that we averaged the responses for each row and column over eight flashes, respectively. A 4 s interval separated each symbol, enabling the patient to process the online feedback and then focus attention on the next symbol to spell.
The patient completed one training (Train01) and four online testing sessions (named Sess01, Sess02, Sess03 and Sess04, respectively) within a period of two weeks, each separated by three or four days. The training and first test were performed on the same day. During the former, the patient was given a sequence of 30 characters to spell. Test sessions consisted in spelling 4-character (french) words. Sess01, Sess02, Sess03 and Sess04 were made of nine, eighteen, twelve and eighteen words, respectively.
All parameters learned from Train01 were subsequently applied during the four online test sessions. These parameters pertain to the feature selection and subsequent classification steps. The former consisted in spatial filters derived from the xDAWN Algorithm [6]. Then a two-class Naïve Bayes classifier was trained, based on the spatially filtered training data. The main processing steps of our OpenViBE P300-speller scenario include [1]:

- Bandpass filtering between 1 and 20 Hz
- Spatial Filtering using xDAWN algorithm
- 600 ms wide epoching, starting from flash onset
- Averaging over 8 single-sweep epochs for each stimulation type
- Selecting the row and column with the highest posterior probability for class target.

3 Results

To evaluate performance, we considered two types of accuracy: the item selection accuracy and the P300 classification accuracy. While the first one reflects the percentage of items spelled correctly, the second one deals with rows and columns independently and indicates the percentage of target stimuli correctly identified. Consequently, chance level performance for item detection and P300 classification are 2.8 % and 16.7 %, respectively.

	Training dataset	Test dataset	Item selec. accuracy	Classif. accuracy
Online	Train01	Sess01	24.4 %	53.7 %
	Train01	Sess02	4.2 %	19.5 %
	Train01	Sess03	0.0 %	13.5 %
	Train01	Sess04	2.8 %	18.4 %
Offline	Sess02.01	Sess02.02	15.6 %	37.8 %
	Sess03.01	Sess03.02	15.0 %	52.5 %
	Sess04.01	Sess04.02	27.3 %	48.9 %

Table 1: Item selection and classification accuracies for both the online and offline evaluation.

Table 1 shows the performance obtained online, in each session and, offline, after having split each session dataset into a training (Sess0x.01 made of the first thirty letters of Session 0x) and a test set (Sess0x.02 made of the remaining symbols of the same session). Note that those results emphasize the need for a session specific training, which suggests a high variability in the signals from one day recording to another.

Figures 1 and 2 represent the topographical distribution of errors with respect to target location, for session Sess02.02 and Sess03.02, respectively. The pixel at the center of those matrices indicates the target location. The darker the pixel, the higher the frequency of selection of the corresponding location. Note that Sess03.02 exhibits a much less dispersed distribution of errors around target location, compare to Sess02.02.

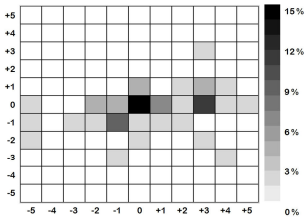


Figure 1: Error distribution for Sess02.02.

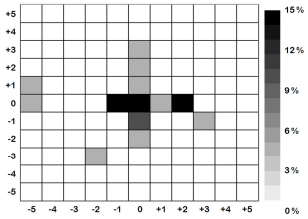


Figure 2: Error distribution for Sess03.02.

Finally, Figure 3 depicts the learned typical target and non-target responses with associated standard deviation, for the patient (a) and two healthy subjects (a poor (b) and a good (c) performer). Importantly, for comparison between individuals, item selection accuracies obtained in subsequent tests were 27.3 % (a), 43.8 % (b) and 99.4 % (c), respectively.

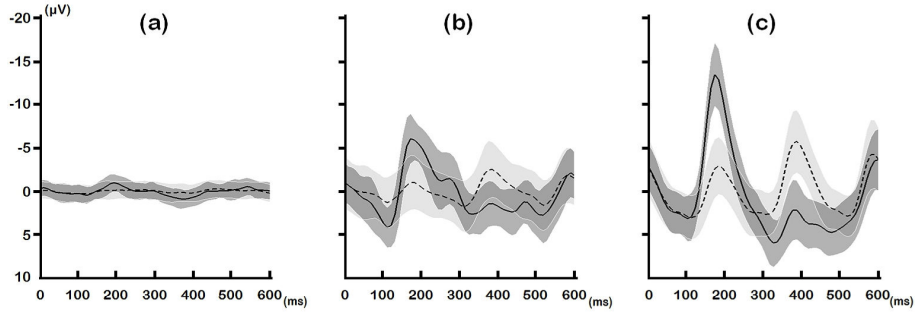


Figure 3: Averaged waveforms for target (solid line) and non-target (dashed line) stimuli, estimated from training datasets, having applied optimal spatial filtering in each individual (a patient (a) and two healthy subjects (b & c)). The shading areas represent the mean ±1 std.

4 Discussion

In this short paper, we evaluated for the first time, our OpenViBE implementation of the P300-speller BCI in a Locked-in patient. On average, online performance observed over four sessions were very poor, especially compared to previous evaluation in healthy subjects. Indeed, the LIS patient reached an item selection accuracy of 24.4 % at best, while 18 healthy subjects ranged between 73.3 and 100 % (mean 97.4 %) [1]. Offline analysis of the patient's data clearly revealed that the calibration phase could not be re-used from one day to another. Indeed, generalization from initial training to the last three sessions yielded performance at chance level.
After having ruled out technical explanations and the obvious possibility that the patient would not perform the task correctly, we conclude from this indepth evaluation that this patient's signals suffer from both low signal-to-noise ratio and high variability. This is supported by observed offline performance and associated spatial distribution of errors. Interestingly indeed, although Sess02.02 and Sess03.02 show similar (low) performance in terms of item selection, the former exhibit a significantly lower classification accuracy and a wider spread spatial distribution of errors. This speaks in favour of a difficulty to maintain attentional focus on the target.

Finally, the comparison between typical target and non-target responses, in the patient and two healthy subjects, is also striking. The patient's responses are an order of magnitude weaker. This most probably explains the dramatic difference in performance.

5 Conclusion

Although limited to a single individual so far, this evaluation is a rare case study in a truly locked-in patient. Performance obtained are significantly poorer than the ones observed in healthy subjects. This is in line with previous observations in BCI, suggesting that performance often decreases as physical impairment increases [7]. However, there is still room for improvement, at least on the computer side. For instance, we believe that adopting a new stimulation procedure to suppress adjacency distractions could boost the online performance [8].

6 Acknowledgments

This work is supported by the French ANR-DEFIS program under grant ANR-09-EMER-002 in the context of the CoAdapt project.

References

[1] E. Maby, G. Gibert, P. E. Aguera, M. Perrin, O. Bertrand, and J. Mattout. The OpenViBE P300-speller scenario: a thorough online evaluation. *Human Brain Mapping Conference*, 2010.

[2] F. Nijboer, E. W. Sellers, J. Mellinger, M. A. Jordan, T. Matuz, A. Furdea, S. Halder, U. Mochty, D. J. Krusienski, T. M. Vaughan, J. R. Wolpaw, N. Birbaumer, and A. Kübler. A P300-based brain–computer interface for people with amyotrophic lateral sclerosis. *Clinical Neurophysiology*, 119(8):1909–1916, 2008.

[3] Y. Renard, F. Lotte, G. Gibert, M. Congedo, E. Maby, V. Delannoy, O. Bertrand, and A. Lécuyer. OpenViBE: An open-source software platform to design, test and use brain-computer interfaces in real and virtual environments. *Presence Teleoperators & Virtual Environments*, 19(1):35–53, 2010.

[4] C. Fischer, J. Luaute, and D. Morlet. Event-related potentials (MMN and novelty P3) in permanent vegetative or minimally conscious states. *Clinical Neurophysiology*, 121(7):1032–1042, 2010.

[5] G. Gibert, V. Attina, J. Mattout, E. Maby, and O. Bertrand. Size enhancement coupled with intensification of symbols improves P300 Speller accuracy. *Proc. 4th International BCI Interface Workshop and Training Course*, pages 250–255, 2008.

[6] B. Rivet, A. Souloumiac, V. Attina, and G. Gibert. xDAWN algorithm to enhance evoked potentials: application to brain-computer interface. *IEEE Trans Biomed Eng*, 56(8):2035–2043, 2009.

[7] A. Kübler and N. Birbaumer. Brain-computer interfaces and communication in paralysis: extinction of goal directed thinking in completely paralysed patients? *Clinical Neurophysiology*, 119(11):2658–2666, 2008.

[8] 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 column. *Clinical Neurophysiology*, 121(7):1109–1120, 2010.

Continuous Control of a Mobile Robot with an Audio-Cued SMR-BCI

R. Ron-Angevin[1], F. Velasco-Álvarez[1], S. Sancha-Ros[1], L. Da Silva-Sauer[1]

[1]Dpto. Tecnología Electrónica, University of Málaga, Málaga, Spain

rra@dte.uma.es

Abstract

In this work, a sensorimotor rhythm-based, self-paced brain-computer interface (BCI) is proposed to control a mobile robot using four low-level navigation commands: turn right, turn left, move forward and move back. In order to reduce the number of mental tasks and then, the probability of misclassification, the BCI is to be controlled with only two mental tasks (relaxed state versus imagination of right hand movements). We hypothesized that three levels of adaptation could be useful in the training process: i) moving from discrete to continuous movements; ii) changing from a virtual to a real environment; iii) using firstly a graphical control interface before navigating with only an audio-cued interface. Three healthy subjects participated in the experiment; they all succeeded in controlling the robot through a goal-oriented maze.

1 Introduction

A brain-computer interface (BCI) is based on the analysis of the brain activity, recorded during certain mental activities, to control an external device. Recently, BCI research is targeted at rehabilitation of motor-disabled individuals, where many BCI applications based on mental task discrimination allow the user to control simulated [1] or real mobile robots [2, 3]. In all these BCI systems, the number of navigation commands is associated with the number of classes to discriminate. In [1] e.g., 2 mental tasks (left and right hand motor imagery, MI) are discriminated in order to execute 2 different commands. However, a higher number of commands is necessary to make the control of the device more effective. One option to increase the number of navigation commands is moving from a binary decision to a more diverse decision, giving a choice between more options [4], e.g., by increasing the number of mental tasks. In [2] 3 mental states are used to provide 3 different navigation commands (move right, left and forward). Very recently, *Barbosa et al.* have reported an EEG-based BCI that is able to discriminate among 4 different mental activities related to sensorimotor rhythms (imagined movements of feet, tongue, left arm and right arm) to provide 4 robot movements: stop, move forward 500 mm, and turn left or right 30 degrees [3].

In this work, an EEG analysis-based, self-paced (asynchronous) BCI is proposed to control a mobile robot using 4 different navigation commands: turn right, turn left, move forward and move back. A self-paced BCI system distinguishes between 2 states: (i) a non-control (NC) state in which subjects can be involved in a mental activity other than controlling the BCI, and (ii) an intentional control (IC) state, in which subjects can control the system by specific mental tasks. Subjects voluntarily switch the state they are in. As proposed in [5], in order to reduce the probability of misclassification, the BCI is controlled with only 2 mental tasks: one MI versus relaxed state. The main objective of the study is to validate the usefulness of the system that lets the subjects control the robot in any direction using 4 low level commands.

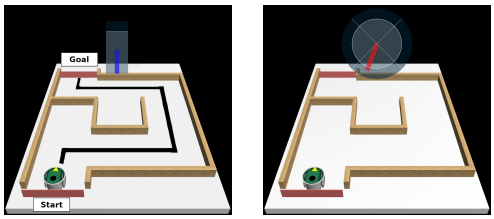


Figure 1: Simulated mobile robot in VE: NC interface (left) and IC interface (right).

2 Methods

2.1 Subject and Data Acquisition

Three healthy subjects participated in the study. None of them had any previous experience in BCI. The EEG was recorded from 2 bipolar channels. The active electrodes were placed 2.5 cm anterior and posterior to electrode position C3 and C4 (right and left hand sensorimotor area, respectively) according to the 10/20 international system. The ground electrode was placed at the FPz position. Signals were amplified by a 16-channel biosignal amplifier (g.Bsamp, Guger Technologies), and then digitized at 128 Hz by a 12-bit resolution data acquisition card (NI USB-6210, National Instruments).

2.2 Initial Training and Signal Processing

Before the online self-paced experiments, subjects participated in 2 initial training sessions, for calibration purposes. This training was based on the paradigm proposed by our group (UMA-BCI) in [6], in which subjects, immersed in a VE, had to control the displacement of a car right or left, according to the mental task carried out, in order to avoid an obstacle.

The training was carried out discriminating between 2 mental tasks: mental relaxation and imagined right hand movements. In the first session, the subjects did not receive any feedback, and it was used to set up classifier parameters for the second session, in which continuous feedback was provided. The same parameters were used to calibrate the system for the control of the robot in the VE exploration sessions. The offline processing was based on the procedure detailed in [6], and consisted of estimating the average band power of each channel in predefined, subject-specific reactive frequency bands (manually selected) at intervals of 500 ms. In the feedback session, the length of the bar was computed on line every 31.25 ms as a result of a Linear Discriminant Analysis (LDA) classification.

2.3 Control of the Robot

The main objective of the experiment was to control a real robot in a continuous way. To this end, a progressive training divided into 2 phases, and carried out in 6 different days, was proposed. During the first phase, subjects learned to control the robot in a discrete way. After this, subjects participated in a second phase to control the robot in a continuous way. In each phase, before controlling the real robot, which it was controlled with the only help of an audio-cued interface, subjects practiced controlling a simulated robot through a group of corridors, setting up a sort of small maze. The proposed VE is presented in Figure 1 and was designed with the same features as the real environment to control the real robot: 70×95 cm; the corridors were all 20 cm wide. The virtual robot, with the same features as the real robot (an EPFL educational e-Puck robot) is cylindrical-shaped with a diameter of 7.5 cm and a height of 4.7 cm. It was configured to stop automatically when it approached an obstacle at 2 cm, to move at a speed of 3.9 cm/s and turn at 42.9 degrees/s. During the first phase, the BCI commands were translated into 4 different movements of the robot: turn 90 degrees to the right or to the left, move forward or backward a fixed distance (14 cm). By default, when the robot moved back the distance was set to 14 cm;

Ses	T(s)	Col	F	R	B	L	TC	E	I	M
1	431.5	3	16.5	11.5	3.5	5.5	36	V	V-A	D
2	616.8	2.8	14.5	7.65	3	7.65	33.3	V	A	D
3	500.3	2.9	14.6	8.6	3.3	6.3	32.4	R	A	D
4	496.3	1.5	18.3	10.6	1.9	10.6	41.3	V	V-A	C
5	541.6	4.8	18.9	13.1	6.4	13.1	51.8	V	A	C
6	522.8	2.9	16.9	11.6	3.1	12.3	44.1	R	A	C

Table 1: Average results obtained from online self-paced BCI experiments for each run.

but if the previous movement ended up in a collision, it moved the same distance as the previous advance, in order to offer the subject the opportunity to rectify the error. The task was to drive the simulated robot from the start position to the goal as fast as possible, using the minimum number of navigation commands, trying to move always forward (the forward direction is indicated by an arrow on the top of the robot) and avoiding the collisions.

The procedure to control the simulated robot with the BCI system is similar to the one proposed in [7], where subjects had to navigate through a virtual environment (VE) using 3 different navigation commands. 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, see Figure 1 left). By extending the bar carrying out the MI task, if its length exceeds a subject-dependant selection threshold, the system switches to the 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 rotation speed was fixed to 24 degrees every second) (Figure 1 right). 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. The way selection works in this interface is the same as in the NC interface, with the same selection threshold. Subjects receive audio cues while they interact with the system. When the state switches from IC to NC, 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").

During this first phase of the experiment subjects were trained to switch from the graphical interface to the audio-cued interface. Each subject participated in 3 sessions, carried out in different days, with two experimental runs each. In the first session, the aim of the runs was to drive the simulated robot to the goal using the graphical and the audio-cued interfaces together. In the second session, only the audio-cued interface was used to drive the simulated robot in the 2 runs. Finally, in the third session, the experiments consisted of controlling the real robot in a discrete way using only the audio-cued interface. After these 3 sessions, subjects repeated the same training protocol to control the robot in a continuous way, participating in other 3 sessions denoted sessions 4, 5 and 6 (equivalent to session 1, 2 and 3 respectively). In a continuous way, once a command is selected, the robot starts moving, and the movement is maintained as long as the bar length is above the 'selection threshold' (this means that the subject is still carrying out the MI mental task). A detailed description about how to control the NC and IC interface in a continuous way can be found in [8].

3 Results

In Table 1, the average values obtained from different parameters from each session are shown. These parameters are: the time in seconds necessary to generate the desired trajectory (T), the number of times that the robot collided with the wall (Col), the number of selected commands of each type (F: Forward, R: Right, B: Back and L: Left) and the total number of commands used to drive the robot from the start position to the goal (TC). The final three columns indicate the kind of session: environment (E; that can be virtual 'V' or real 'R'), interface (I; visual and auditive

'V-A' or only auditive 'A') and movement (M; discrete 'D' or continuous 'C').

It is worth mentioning that the required average time to generate the trajectory controlling the robot in a continuous way (522.8 s) is similar to the average time spent to control the robot in a discrete way (500.3 s). However, when only the audio-cued interface was used, a clear improvement in performance can be observed, in terms of time, between the control of the simulated robot and the real robot (616.8 s and 500.3 s for session 2 and 3 respectively in discrete way, and 541.6 and 522.8 for session 5 and 6 respectively in continuous way). The number of collisions was very small for all sessions (except for session 5, with a mean value of 4.8 collisions).

4 Discussion and Conclusion

In general, the results obtained in only 6 sessions are remarkably good. It must be pointed out that the results are worse in session 2 and 5; maybe because it is the first time that subjects face an audio-only interface in both modalities (discrete and continuous way respectively). These sessions should be considered as a training for sessions 3 and 6, where subjects perform noticeably better. A progressive training has been proposed. The first 3 sessions are used to learn to control the robot in a discrete way and using only the audio-only interface. In a second phase of the training, subjects achieve control of the robot in a continuous way, increasing the freedom of movement, and trough the discrimination of only 2 mental tasks. The obtained results suggest that this paradigm is useful and allows providing several navigation commands without increasing the number of mental tasks to discriminate.

5 Acnowledgments

This work was partially supported by the Innovation, Science and Enterprise Council of the Junta de Andalucía (Spain), project P07-TIC-03310.

References

[1] C. S. L. Tsui, J. Q. Gan, and S. J. Roberts. A self-paced brain-computer interface for controlling a robot simulator: an online event labelling paradigm and an extended kalman filter based algorithm for online training. *Med. Biol. Eng. Comput.*, 47:257–265, 2009.

[2] J. D. R. Millan, F. Renkens, J. Mourino, and W. Gerstner. Noninvasive brain-actuated control of a mobile robot by human EEG. *IEEE Trans. on Biomed. Eng.*, 51:1026–1033, 2004.

[3] A. O. G. Barbosa, D. R. Achanccaray, and M. A. Meggiolaro. Activation of a mobile robot through a brain computer interface. *IEEE Int. Conf. on Rob. and Autom. ICRA*, 2010.

[4] B. Obermaier, C. Neuper, C. Guger, and G. Pfurtscheller. Information transfer rate in a five-classes brain-computer interface. *IEEE Trans. Neur. Syst. Rehab. Eng.*, 9:283–288, 2001.

[5] J. Kronegg, G. Chanel, S. Voloshynovskiy, and T. Pun. EEG based synchronized brain-computer interfaces: a model for optimizing the number of mental tasks. *IEEE Trans. Neur. Syst. and Rehab. Eng.*, 15:50–58, 2007.

[6] R. Ron-Angevin and A. Diaz-Estrella. Brain-computer interface: changes in performance using virtual reality techniques. *Neurosci. Lett.*, 449:123–127, 2009.

[7] F. Velasco-Alvarez and R. Ron-Angevin. Asynchronous brain-computer interface to navigate in virtual environments using one motor imagery. *LNCS*, 5517:698–705, 2009.

[8] F. Velasco-Alvarez, R. Ron-Angevin, and M. J. Blanca-Mena. Free virtual navigation using motor imagery through an asynchronous brain-computer interface. *Presence*, 19:71–81, 2010.

'Brain Invaders': a Prototype of an Open-Source P300-Based Video Game Working with the OpenViBE Platform

M. Congedo, M. Goyat, N. Tarrin, G. Ionescu, L. Varnet, B. Rivet, R. Phlypo, N. Jrad, M. Acquadro, C. Jutten

ViBS (Vision and Brain Signal Processing), GIPSA-lab, CNRS, Grenoble University. FRANCE

Abstract

We have developed the prototype of a pure-BCI video game based on the well known vintage video game 'Space Invaders'. In our 'Brain Invaders' a number of aliens are displayed in a grid and the player has to destroy a particular alien, the target, only by concentrating on it. The game makes use of a state-of-the art P300 oddball paradigm to select the alien to be destroyed at a regular pace, based on current probabilities assigned to each alien by a learning machine continuously analyzing and classifying the user's electroencephalographic stream (the Open-ViBE platform). As compared to the standard P300-speller paradigm our game may optionally use: 1) flashing items in random groups and no longer by rows and columns; 2) variable inter-stimulus intervals drawn from an exponential distribution; 3) magnification of flashed target items. Preliminary tests show an excellent transfer rate, since starting with 36 aliens on the screen, one to three repetitions typically suffice to destroy the target alien. Our development is completely open-source and will continue to improve further the signal processing and classification algorithms, besides the paradigm itself, the gameplay and the BCI ergonomics, in order to achieve a 'Plug & Play' video game suitable for the large public of gamers.

1 Introduction

A P300-based Brain Computer Interface (BCI) enables the user to successively select symbols among an available set, without relying on any motor command. The symbols can be of any kind, such as alphanumeric characters (e.g., for spelling) or icons (e.g., the elements of a menu in a computer application). These BCIs exploit the well-known oddball paradigm, in which an infrequent task-related item (the target symbol) elicits a P300 Event-Related Potential (ERP) [1]. By flashing symbols exhaustively, either one-by-one or in groups, it is possible to estimate the probability of each symbol being the one selected by the user. This is achieved evaluating the P300 elicited by each symbol once it has flashed. The complete set of flashes must be repeated a number of times to obtain reliable ERP estimations by means of trial averaging. The distinctive advantages of P300-based BCI are that the alphabet (the set of all available symbols) can be large (hundreds of symbols) and that 100 % accuracy can be obtained when allowing a sufficient number of repetitions. That is to say, with P300-based BCIs there is a direct trade-off between accuracy and transfer rate. In this work, the low transfer rate is not considered a limitation, rather a challenge for the player, along the line of the reasoning in [2]. Nonetheless, we aim at a video game flowing with a sustained pace. For this reason we have implemented three improvements over the basic P300 BCI paradigm.

2 Methods

Among the three (optional) improvements over the basic P300 paradigm [3], only the third takes advantage explicitly of the a-priori target knowledge; the first two can be applied to any P300 BCI:

1) In the original P300-speller paradigm symbols flash by rows and columns. Often detection errors arise because of the 'adjacency-distraction' phenomenon [4, 5]; non-target symbols in rows

or columns adjacent to the target attract the user's attention when they flash, producing a P300 that makes the detection of the target P300 more difficult. To mitigate this effect we flash the symbols by random groups. Let the pair (r, c), with $r \in \{1, \ldots, R\}$ and $c \in \{1, \ldots, C\}$ be the ordinal indexes of the row and column position of each symbol, with R rows and C columns in total. It suffices to create a supplementary index table with entries $(x[r], y[c])$, where $x[r] \in \{1, \ldots, R\}$ and $y[c] \in \{1, \ldots, C\}$ represent permutations and flash the symbols according to these new mapped indexes, where the permutation may change over repetitions. Not only the 'adjacency-distraction' effect is mitigated, we also obtain that the pattern of flashing becomes totally unpredictable, which is expected to sustain the attention of the gamer. Furthermore, by flashing in random groups there is no apparent difference between a 'row' and 'column', thus we can flash all rows then all columns (or vice versa) rather than rows and columns alternately. Hence, the probability that the target symbol is included in two successive flashes, which is known to worsen the target detection, is only $1/(R{\times}C)$. Of course, one may generate pseudo-random sequences to avoid also this little probability, but this is out of scope of this contribution. Noteworthy, random-group flashing allows arbitrary positioning of the symbols on the screen (no more need to arrange symbols on a grid), which greatly expand the usability of the P300 paradigm.

2) Usually, the stimulus interval (the flashing time) and the inter-stimulus interval (ISI; the time between two flashes) are kept constant. The periodic flashing is annoying and tiring because the visual cortex is driven to oscillate at the flashing frequency, which is usually far away from the natural talamo-cortical loop oscillation of this region, which is in the alpha range (8–12 Hz). Furthermore, the periodic flashing makes the flashing pattern predictable and boring. To eliminate all these effects we may use random ISI drawn from an exponential distribution. The exponential distribution (also called 'waiting-time' distribution) with parameter λ and population mean(sd)=$1/\lambda(1/\lambda)$ is a natural distribution for the time passing in between two events of random series following a Poisson process with the same parameter λ and population mean(sd)=$1/\lambda(1/\lambda)$, that is a process in which events occur continuously and independently at a constant average rate. Random ISI uniformly distributed in between 450 and 550 ms have been successfully used in a P300 BCI by [6].

3) In a previous research conducted within the OpenViBE ANR project it has been found that magnifying (increasing the size of) the flashed symbols augments the P300 amplitude and improves classification accuracy [7]. In this work, we only magnify the target symbol – by 30 % – during flashing.

The 'Brain Invaders' works by rounds. In each round a 6×6 grid of aliens is displayed on a black background. Aliens are all displayed grey with the exception of the target which is displayed in red. The group flashing is achieved displaying the non-target symbols with a lighter grey and the target symbol with its complementary color (cyan). The choice of the complementary color is natural as it creates maximum contrast on the color wheel. As in the original 'Space Invaders', aliens move altogether from side to side of the screen continuously and move from top to bottom by one step as they touch either screen border. Figure 1 (left) shows a random group flashing and target magnification. At each repetition the system assigns to each symbol the probability of being the target according to the signal processing and classification method implemented in the OpenViBE platform (the xDAWN spatial filter [8] followed by a Linear Discriminant Analysis classifier) and destroys the alien with the highest probability (Figure 1, right). If this alien is the target the round ends, otherwise this alien is eliminated and another repetition of flashes starts. The process is continued until the target alien is destroyed or until 14 non-target aliens have been destroyed, after which another round starts. Between two rounds, the points obtained in the last round and the cumulative score are shown to the player. The points obtained at each round are inversely proportional to the number of repetitions necessary to destroy the target (NRD). The flashing time is fixed and equals 60 ms. The ISI, as aforementioned, is based on a random variable drawn from an exponential distribution with $\lambda = 1$, postmultiplied by 100 to obtain a mean ISI of 100 ms and is further bounded to the range $[20 \ldots 500]$ ms by excluding samples that fall out of this range. The destruction is almost instantaneous after the last flash. Then a 3 s break is allowed to relax and move freely, after which the new round starts. A collection of rounds is tagged a

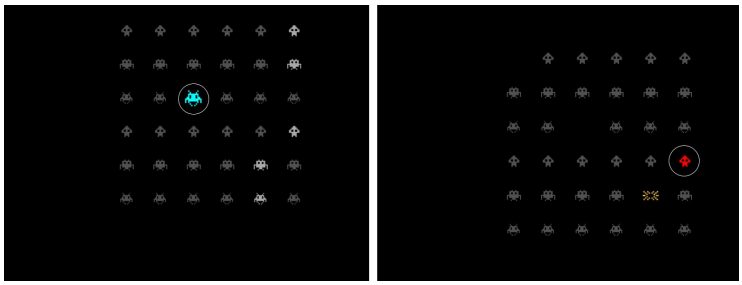


Figure 1: Screen shots of the 'Brain Invaders'. On the left it is shown the flashing of a random group including the target (circled in the figure, but not in the game), which is magnified by 30 % during flashing. The flash is within the first repetition (no alien has been destroyed yet). On the right it is shown the destruction of a non-target alien at the end of the third repetition (two non-target aliens have been already destroyed after the first two repetitions of this round).

block. Typically a block is comprised of six rounds. Between rounds a 30 s break is allowed. A game session is made of several blocks.

The implementation of the 'Brain Invaders' is achieved with three software modules, 1) acquisition, 2) processing and 3) rendering. Since they communicate to each other via a TCP/IP protocol, they may run on a single computer or on distinct computers in any combination:
Acquisition. This is the Open-ViBE acquisition server [9]. It is in charge of acquiring the data from the EEG machine, streaming the data, correcting for possible amplifiers drifts and sending the data to the Open-ViBE platform [9] (http://openvibe.inria.fr/) for analysis.
Processing. The Open-ViBE platform performs data analysis on-line. At the end of each repetition it computes the probability of each alien being the target and sends to the rendering application the indexes of the alien with the highest probability.
Rendering. The visual rendering module, written in C++ using the Ogre3D engine (www.ogre3d.org), runs the user-interface. It manages target destruction and the sequence of trials in blocks. It is in charge also of tagging the flash onset/offset via parallel port directly into the EEG acquisition machine. We use a 120 Hz monitor.

3 Results

We have tested the prototype on four BCI-naive healthy subjects (S1... S4). S1 and S2 played with the random group flashing set up, whereas S3 and S4 played with the usual row-column flashing. The training was composed of 10 times 8 repetitions, lasting approximately three minutes only. The test was composed of six blocks of six trials each. EEG was acquired on a Mitsar EEG-202 amplifier (Mitsar, St. Petersburg, Russia) using an elastic cap (Electro-Cap, Inc., Eaton, USA) holding 31 silver chloride electrodes uniformly positioned all over the scalp according to the extended 10-20 system and referenced to both earlobes linked digitally. The sampling rate was 500 samples per second and an acquisition band-pass filter in the range 0.1–70 Hz was enabled. The mean (sd) number of NRD over the 36 trials was 2.06 (1.07) for S1, 1.50(0.65) for S2, 1.33 (1.34) for S3 and 3.56 (3.43) for S4 (Fig. 2). For this experiment 89 % of trials required at most three repetitions to destroy the target. For S2 and S3 this was true for 100 % of trials.

4 Conclusion

Our development is open-source and available to interested peers. This research is a starting point of an ongoing program. Our goal is to achieve a 'Plug&Play' game suitable to the large public. To this end, our next priority is to get rid completely of the learning phase by devising an adaptive learning algorithm initialized using the average optimal setting obtained on a large database. Then, we will work on the gameplay, adding on the foreground animations becoming

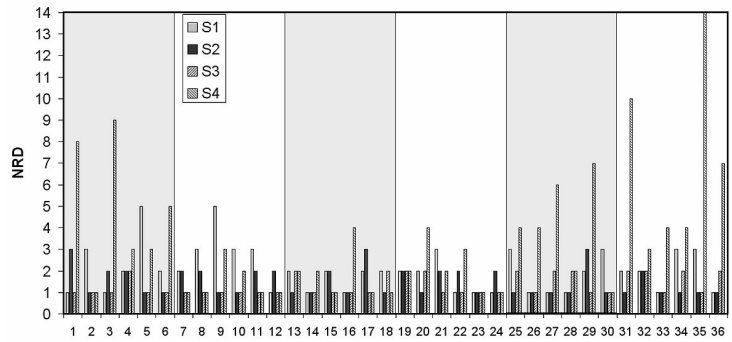


Figure 2: Number of repetitions necessary to destroy the target (NRD) on the y-axis for each trial (x-axis) for four BCI-naif subjects. The vertical bands indicates the six block, composed of six trials each.

progressively more distracting as the number of blocks progresses. Meanwhile, further research is being carried out at GIPSA-lab to improve the ERP single-trial extraction and classification, besides the BCI paradigm itself.

Acknowledgements
This research has been partially funded by French ANR (Agence Nationale de la Recherche) projects RoBIK and OpenViBE2 (http://www.irisa.fr/bunraku/openvibe2/wiki/).

References

[1] J. R. Wolpaw, N. Birbaumer, L. D. L. McFarland, G. Pfurtscheller, and Vaugan T . Brain-computer interfaces for communication and control. *Clin Neurophysiol*, 113:767–791, 2002.

[2] A. Nijholt, D. P.-O. Bos, and B. Reuderink. Turning shortcomings into challenges: brain-computer interfaces for games. *Entertain Comput*, pages 85–94, 2009.

[3] L. A. Farwell and E Donchin. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalogr Clin Neurophysiol*, 70:510–23, 1988.

[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. *Clin Neurophysiol*, 121(7):1109–20, 2010.

[5] J. Jin, B.Z. Allison, E.W. Sellers, C. Brunner, P. Horki, X. Wang, and C. Neuper. Optimized stimulus presentation patterns for an event-related potential EEG based brain-computer interface. *Med Biol Eng Comput*, 49(2):181–91, 2011.

[6] B. Z. Allison and J. A. Pineda. ERPs evoked by different matrix sizes: implications for a brain computer interface (BCI) system. 11(2):110–113, 2003.

[7] G. Gibert, V. Attina, J. Mattout, E. Maby, and O. Bertrand. Size enhancement coupled with intensification of symbols improves P300 speller accuracy. In *4th International Brain-Computer Interface Workshop and Training Course 2008*, 2008.

[8] B. Rivet, A. Souloumiac, V. Attina, and G. Gibert. xDAWN algorithm to enhance evoked potentials: application to brain-computer interface. 56(8):2035–43, 2009.

[9] Y. Renard, F. Lotte, G. Gibert, M. Congedo, E. Maby, V. Delannoy, O. Bertrand, and A. 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.

Inferring Mental States by Means of Functional Imaging Methods that Probe Brain Responses to Emotional Language

C. Herbert[1], J. Kissler[2]

[1]Department of Psychology, University of Würzburg, Germany
[2]Department of Psychology, University of Konstanz, Germany

cornelia.herbert@psychologie.uni-wuerzburg.de

Abstract

Internal, mental states such as the emotional state of a person are usually assessed by means of self-report measures. Standard self report measures often use emotional trait adjectives. The present functional imaging study investigated brain states underlying the processing of emotional and neutral adjectives in healthy subjects during a silent reading task. Results showed an increase in hemodynamic brain signals in visual and medial temporal cortex during reading of all types of emotional compared to neutral adjectives. Medial prefrontal cortex including ACC (anterior cingulate cortex) was specifically activated during reading of positive trait adjectives. Results are consistent with a self-positivity bias and suggest that information about an individual's internal emotional state can be inferred by means of functional imaging methods in the absence of any instructed or overt verbal or behaviorally visible response of the individual. These results have implications for research in behaviorally non-responsive patients and the development of novel types of brain computer interfaces that are based on individuals' emotional responses (eBCI).

1 Introduction

How can we infer a person's mental state, his/her emotional state or self concept without directly using self-report measures? Questions like these are of central concern for experimental research investigating behaviorally non-responsive individuals, who cannot reply to nor communicate with their environment by giving verbal or behavioral signs due to severe brain injuries or chronic neurodegenerative motor diseases like amyotrophic lateral sclerosis [1]. Neurophysiological methods based on electroencephalography (EEG) and functional magnetic resonance imaging (fMRI) could provide a solution to the above described problems: in combination with the appropriate experimental design these methods provide a means for the online investigation of a person's mental activities even without simultaneously recording his/her overt behavioral responses. Brain signals detected during such tasks could, together with assistive technologies such as brain computer interfaces, be used as predictors of an individual's current emotional state, even in samples of behaviorally non-responsive patients. Regarding emotional states, self-report measures [2–4] routinely use emotional trait adjectives for the assessment and evaluation of an individual's emotional state, his or her self-concept. Results show that healthy people evaluate positive trait adjectives (e.g., happy, cheerful, content, etc.) as more relevant and congruent with their self-concept compared to negative adjectives describing unpleasant emotional traits and states (e.g., unhappy, anxious, sad, etc.). This phenomenon has been called self positivity bias [5]. In contrast, depressed individuals and individuals at risk for depression evaluate negative trait adjectives as more descriptive of themselves. This hypersensitivity to negative words occurs in conjunction with a negative self concept, self-reproach and a pessimistic view about the self and the future [6, 7]. Recently, functional imaging studies investigated the neural correlates underlying these self-related positivity/negativity biases [8]. Most of these studies used functional imaging methods in combination with explicit

tasks of self-evaluation, i.e., participants were explicitly instructed to evaluate the words for their emotional significance or self-descriptiveness during scanning [9]. Results consistently showed an increase in brain activity patterns in cortical midline structures including medial prefrontal cortex (MPFC) in response to trait adjectives evaluated as self-descriptive by the participants. These results, amongst others, have led to the suggestion that MPFC might constitute an important converging zone for the online monitoring and appraisal of personally relevant information [10, 11]. Here, we investigated if changes in brain activity patterns such as the MPFC also occur in response to positive and negative adjectives when no explicit evaluation task is provided to the participants (i.e., silent reading). Brain signals indicating a person's emotional state during an implicit processing task like silent reading could be used either as evaluation indicators or as components in the interface loop of brain computer interfaces. This would allow for exhaustive training of the BCI classification algorithm under various emotional states and on-line emotional adaptation of the classification algorithm.

2 Methods

2.1 Participants

Fifteen right-handed healthy native German adults (7 females, mean age: 26 years) without history of psychiatric or neurological disorder or medication for any of these, participated in the study. Participants were recruited via the University of Tübingen and were paid € 10 for participation. Participants were informed in detail about the imaging procedure and gave written informed consent to participate in the study.

2.2 Stimulus Material and Experimental Design

Stimuli consisted of emotional (60 positive and 60 negative) and 60 neutral adjectives. Positive, negative and neutral adjectives did not differ in word-length or word frequency and positive adjectives were rated as more self-descriptive compared to both negative and neutral adjectives by an independent sample of healthy adults not participating in the present study. Stimuli were presented in three consecutive runs. In each run, words were presented for 1 second in 5 blocks of 12 words per word category. Block order was randomized across runs and participants. Stimuli were presented visually and participants' task was to read the words silently during scanning.

2.3 Signal Recording and Analysis

Functional images were recorded on a 1.5 T-whole body scanner (Siemens Vision, Erlangen, Germany) using T2*-weighted multislice echo-planar imaging (EPI) sequences (28 axial slices acquired in descending direction, 5-mm thickness, 1-mm gap, TR = 3 s, TA = 100 ms, $\alpha = 86\,^{\circ}$, FOV = $192 \times 192\,\mathrm{mm}^2$). Imaging data were analyzed with Statistical Parametric Mapping software (SPM99, Wellcome Department of Imaging Neuroscience, London, UK). After standard pre-processing, hemodynamic responses were modeled using a box car function convolved with the canonical hemodynamic response function (hrf). Individual contrast images were averaged across runs and entered into a second level random-effect analysis comparing (1) positive versus neutral, (2) negative versus neutral, and (3) positive versus negative using one-sample T Tests. Results are reported for clusters extending a spatial threshold of at least 20 contiguous voxels at a significance threshold of $P < 0.005$ ($T > 2.98$, uncorrected).

3 Results

Reading of emotional (positive or negative) compared to neutral adjectives significantly increased hemodynamic signals in visual (BA 18, BA 19) and temporal (BA 37) cortex. Reading of positive adjectives was associated with significantly larger hemodynamic signal changes in medial prefrontal

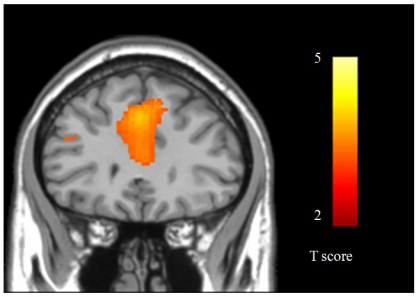


Figure 1: Hemodynamic signal change in MPFC / ACC during reading of positive adjectives relative to neutral adjectives and negative adjectives.

cortex regions (BA 32) including anterior cingulate cortex (BA 24) in comparison to both, reading of negative adjectives or neutral adjectives devoid of any emotional connotation (see Figure 1).

4 Discussion

The results of the present study demonstrate that brain regions in the visual, temporal and medial prefrontal cortex show more, or with respect to MPFC and ACC, preferential activation when healthy individuals read adjectives describing positive emotional states and traits. Medial prefrontal cortex and ACC play a role in self-referential processing [8–11] and the generation of subjective emotional experience [12]. The greater involvement of specifically MPFC and ACC during processing of positive adjectives compared to neutral or negative adjectives is consistent with previous imaging results using adjectives and explicit self-evaluation tasks [8, 9] and in support of a self positivity bias in healthy subjects [5, 13].

5 Conclusion

The results of the present study have a number of practical implications for research interested in improving treatment and communication in behaviorally non-responsive patients. First, our results suggest that information about internal, mental states can be assessed by means of functional imaging methods in the absence of any instructed or overt verbal or behaviorally visible response of the individual. Second, our paradigm could provide insight into changes in the emotional state and self-concept in behaviorally non-responsive patients, such as patients suffering from ALS. During the course of the disease, ALS patients are challenged with a number of critical and threatening life events that might impact their self-concept, their affect and emotional experience in a negative way. Interestingly, many studies using self-report techniques suggest that ALS patients do not experience themselves as severely depressed, nor do they differ in self-scored perception of mental health from healthy controls [14, 15]. Research using our paradigm could complement this body of literature and provide additional information of ALS patients' current mood and self-concept, even in later phases of the disease, in which almost no overt behavioral signs and responses can be obtained from the patient. Third, in non-responsive patients our paradigm could be used for communication of emotional states via brain computer interfaces: For an emotional BCI (eBCI) system the following steps could be integrated: Detection of different emotional states and changes thereof by the BCI system. In this respect, combination of our paradigm with EEG methods could also be promising and recently demonstrated to yield impressively reliable results in samples of healthy adults [5, 13].

References

[1] D. Lulé, V. Diekmann, S. Anders, J. Kassubek, A. Kübler, A. C. Ludolph, and N. Birbaumer. Brain responses to emotional stimuli in patients with amyotrophic lateral sclerosis (ALS). *Journal of Neurology*, 254(4):519–527, April 2007. PMID: 17401515.

[2] W. Janke and G. Debus. Die Eigenschaftswörterliste (EWL). Handanweisung. Göttingen: Hogrefe, 1978.

[3] D. Watson, L. A. Clark, and A. Tellegen. Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology*, 54(6):1063–1070, June 1988. PMID: 3397865.

[4] T. B. Rogers, N. A. Kuiper, and W. S. Kirker. Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology*, 35(9):677–688, September 1977. PMID: 909043.

[5] L. A. Watson, B. Dritschel, M. C. Obonsawin, and I. Jentzsch. Seeing yourself in a positive light: brain correlates of the self-positivity bias. *Brain Research*, 1152:106–110, June 2007. PMID: 17462610.

[6] P. A. Derry and N. A. Kuiper. Schematic processing and self-reference in clinical depression. *Journal of Abnormal Psychology*, 90(4):286–297, August 1981. PMID: 7264058.

[7] N. A. Kuiper and P. A. Derry. Depressed and nondepressed content self-reference in mild depressives. *Journal of Personality*, 50:67–79, 1982.

[8] P. Fossati, S. J. Hevenor, S. J. Graham, C. Grady, M. L. Keightley, F. Craik, and H. Mayberg. In search of the emotional self: an fMRI study using positive and negative emotional words. *The American Journal of Psychiatry*, 160(11):1938–1945, November 2003. PMID: 14594739.

[9] K. H. Lee and G. J. Siegle. Common and distinct brain networks underlying explicit emotional evaluation: a meta-analytic study. *Social Cognitive and Affective Neuroscience*, March 2009. PMID: 19270039.

[10] T. W. Schmitz and S. C. Johnson. Relevance to self: A brief review and framework of neural systems underlying appraisal. *Neuroscience and Biobehavioral Reviews*, 31(4):585–596, 2007. PMID: 17418416.

[11] G. Northoff and F. Bermpohl. Cortical midline structures and the self. *Trends in Cognitive Sciences*, 8(3):102–107, March 2004. PMID: 15301749.

[12] R. D. Lane, E. Reiman, G. Ahern, G. Schwartz, R. Davidson, B. Axelrod, and L. Yun. Anterior cingulate cortex participates in the conscious experience of emotion. *Psychosomatic Medicine*, 58:73, 1996.

[13] C. Herbert, J. Kissler, and M. Junghofer. Event related potentials to emotional adjectives during reading. *Psychophysiology*, 45:487–98, 2008.

[14] A. Kübler, S. Winter, A. C. Ludolph, M. Hautzinger, and N. Birbaumer. Severity of depressive symptoms and quality of life in patients with amyotrophic lateral sclerosis. *Neurorehabilitation and Neural Repair*, 19(3):182–193, September 2005. PMID: 16093409.

[15] D. Lulé, C. Zickler, S. Häcker, M. A. Bruno, A. Demertzi, F. Pellas, S. Laureys, and A. Kübler. Life can be worth living in locked-in syndrome. *Progress in Brain Research*, 177:339–351, 2009. PMID: 19818912.

Novel P300 BCI Interfaces to Directly Select Physical and Virtual Objects

B. F. Yuksel[1], M. Donnerer[1], J. Tompkin[1], A. Steed[1]

[1]Department of Computer Science, University College London, UK

b.yuksel@cs.ucl.ac.uk

Abstract

We discuss two novel integrations of a brain-computer interface (BCI) to inform the wider BCI community of the possibilities presented by these applications. BCIs based on the P300 paradigm often use a flashing character or picture visual stimulus to elicit an event-related potential in the brain's EEG signal. Traditionally, P300-based BCI paradigms use a grid layout of visual targets (commonly an alphabet) and allow users to select representations of objects using their thoughts. First, we present a P300 BCI application that allows users to directly select 3D objects in a fully immersive virtual environment. Second, we discuss an application that allows users to select physical objects in the real world directly. We use a multi-touch table that senses and highlights objects placed upon its surface by flashing an area of light around them. Both of these systems allows us to construct a P300-based BCI that uses a collection of objects as targets, rather than a pre-determined grid layout of representations of targets. Results show that our new paradigm works just as well as the traditional paradigm, highlighting the potential for BCIs to be integrated into a broader range of situations. This opens up the field of possibilities for the future of novel and integrated real-world P300 BCIs.

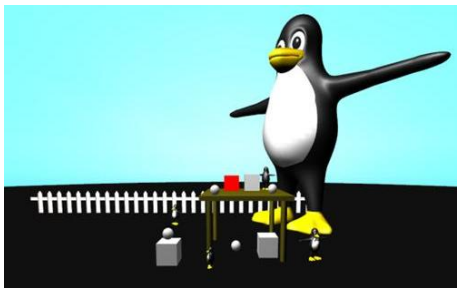
1 Introduction

A brain-computer interface (BCI) is a communication system where the user's commands "do not depend on the brain's normal output pathways of peripheral nerves and muscles" [1]. Thus, a BCI makes it possible to control a computer using only your thoughts. The most common form of BCI uses electroencephalography (EEG) as it is generally considered to be the least expensive and complicated method. BCIs can be used as communication channels for people with severe motor impairments such as amyotrophic lateral sclerosis (ALS). However, there is a growing interest in their use in more general applications [2].

We discuss two novel BCIs that highlight virtual or real objects directly. The opportunities that these two studies present are interesting and we feel they deserve to be widely read by the BCI community. The two studies employ EEG-based BCIs that use the P300 event-related potential (ERP). The P300 brain waveform is an ERP which denotes an increase in voltage of approximately 10 μV, peaking around 300 ms after the stimulus. It is triggered by an auditory, visual or somatosensory stimulus which is infrequent or particularly significant among other more routine stimuli. Its use for BCI was pioneered by Farwell and Donchin (1988) [3].

Previous P300-based BCIs have typically used a grid-based spelling task where a grid of flashing characters or symbols is displayed on a monitor. For example, even when navigating a virtual environment, the P300 grid has been used on a separate screen to allow interaction (e.g., [4, 5]). This means that users had to turn their heads away from the virtual space towards the P300 screen whenever they wished to interact with the virtual environment.

We present two of our own systems which allow users to interact directly with the target objects. The first study integrates the P300 BCI into a fully immersive virtual environment so



Standard Speller System	Multi-Touch System
36 characters flashing.	36 object-blobs & non-objects flashing.
Select one character at a time.	Select one object at a time.
Spell 3-letter words × 4.	Select 3 objects × 4.

Figure 1: *Left*: A scene containing virtual object choices from [6]. A cube on the table (in red) is currently being flashed. *Right*: MTS and SSS experiment methodology comparison.

that 3D virtual objects can be selected directly [6]. The second study takes this one step further and demonstrates how physical objects can be used as P300 targets themselves [7].

The common thread running through both of these studies opens up the space of possible applications and consequent implications of P300-based BCIs. The future of the P300 BCI could be in the real and/or virtual world with direct, integrated user interaction.

2 Selecting Virtual Objects

Donnerer and Steed (2010) [6] demonstrated that it is possible to select virtual objects from a fully immersive virtual environment in a CAVE by using 3D objects as targets. Figure 1 shows the scene with different sized objects placed irregularly. Each object was randomly flashed one at a time, 8 times each. Instead of selecting a character or image from a regular grid, users selected the object itself by keeping a running mental count[1] when their target object flashed. Electrodes were placed at: Fz, Cz, P3, Pz, P4, PO7, Oz, and PO8 based on the international 10-20 system and were connected to a g.MOBIlab + 8-channel system. The P300 classifier was trained with g.tec software (16 flashes/character) using linear discriminant analysis.

Users were asked to select three different objects of varying difficulty based on their size and their proximity to neighbouring objects. Results were promising, with users being able to select objects with a total mean of 50 % of the time. This mean includes the virtual object in the scene that was most difficult to select. Although there was no significant difference in accuracy between objects in the main trial, there were significant differences in the pilot studies.

3 Selecting Physical Objects

Our next step following these promising results [6] was the ability to select real objects directly using a P300-based BCI. Therefore, we replaced the normal P300 grid of characters with physical objects [7].

We created a system whereby objects could be placed on a multi-touch surface which recognized object outlines by a simple computer vision system. Image processing algorithms generated areas of light (object blobs) around these objects. We connected this multi-touch system (MTS) to the g.tec P300-based standard speller system (SSS). We customised the code in the g.tec software to intercept the control of the SSS and relay this over a UDP socket to the multi-touch table. Objects could then be flashed by surrounding the area underneath them with an area of light (Figure 2). Users were then able to select any object on the table by keeping a running mental count of the number of times that object was flashed.

We compared the accuracy of the MTS with the industry standard P300 speller (SSS). We asked 20 participants to select four 3-letter words from the SSS and four sets of 3 objects from

[1] A running mental count is not necessary to trigger a P300 but is a convenient methodology.

Figure 2: A participant interacting with the multi-touch table P300-based BCI from [7]. Six objects have been placed on the table. Left: The spoon is flashed by surrounding it with light. Centre: A non-object flash. Right: The area around the CD is being flashed.

the MTS (Figure 1 Right). We used the same EEG parameters as [6]. Our participants had a mean accuracy rate of 96.2 % using the SSS and 98.7 % for the MTS. Our MTS maximum bit-rate was 15.65 bits/min. We were also able to compare the overall success rates to another recent study by Edlinger et al. (2009) [4] who also used the same EEG parameters except with 15 flashes/character during trials. Table 1 shows the comparative accuracy results (Edlinger et al. do not publish bit-rate for the SSS and so we cannot compare speed results).

4 Discussion

The two studies demonstrate that P300-based spellers can be used elsewhere in real and virtual worlds. For more general users, our main contribution is the first demonstration of a P300 BCI that does not use the standard speller or a simple graphic icon interface.

For the MTS, our results suggest that interfacing the P300 BCI with real-world objects works just as well as traditional paradigms and may even increase accuracy rates of target selection. Further studies would be required to isolate whether this increase in accuracy is due to the participant sample or some aspect of our system. However, we suspect a key difference is that we ran the experiment on a very fast multi-processor PC so that the 300 ms delay in the brain response was precisely measured by the software. In addition, [4] used 15 flashes per object, whereas we used 8, thus fatigue may have been a factor in their study.

Another reason for our high classification accuracy of using real world objects may be due to the form of the interface (e.g., the multi-touch table being wide screen, or the object flashes being large areas of light). Participants did comment that it was easier to select larger cues, and this is also supported by [6]'s findings, suggesting that the size and shape of the cues are important. This would be an excellent topic for a follow-on study. Thus, these could be potential implications for the design of other P300-based BCI systems.

There are several direct applications for this new paradigm such as allowing "locked-in" persons to interact with real objects rather than a screen. For example, we could use more sophisticated computer vision techniques to recognize and label target objects in more general environments, breaking the limitations of a simple P300 screen with pictures or characters. The two studies together hint at a future scenario where real environments could be overlaid with augmented reality so that the physical objects could act as their own interfaces. In a smart home, a projector or array of lights could highlight objects to be used with the BCI. Alternatively, for the physical objects themselves to emit light rather than using a projector, e.g., with LED lights integrated into the object. The active lighting can then be co-ordinated wirelessly to synchronize the flashes.

A more advanced option could be to wear a head-mounted or pendant-like mobile device with a camera and a small projector that augments the physical world and allows the user to interact with the world through BCI. This is not a far-fetched fantasy as a very similar device has already been created [8]. This device uses a tiny projector and camera to visually augment surfaces, walls and physical objects. For example, they show a newspaper overlaid with live video news. Users do not have to wear goggles or glasses resulting in a direct and integrated user experience.

Classification Accuracy %	Edlinger et al. (2009) % of Sessions	Standard P300 Speller % of Sessions	Multi-Touch BCI % of Sessions
100	55.3	70.0	90.0
80–100	76.3	95.0	100.0
60–79	10.6	5.0	0.0
40–59	7.9	0.0	0.0
20–39	2.6	0.0	0.0
0–19	2.6	0.0	0.0
Average Accuracy of All Subjects	82	96	99

Table 1: Comparison of classification accuracies of the P300 BCI in Edlinger et al. (2009) [4], the standard industry P300 speller, and the novel multi-touch table P300-based BCI in [7].

The interface of the Mistry et al. (2009) [8] device is based on hand gesture recognition. We suggest instead to use the P300 BCI as the interface: Object recognition algorithms could highlight potential objects in the scene from which to select, and additional virtual objects could be added to the scene to provide sufficent objects to trigger the P300 response. This future scenario is an integration of the two studies that we have presented here, whereby users can directly interact with physical and virtual objects together. Thus, as was highlighted as an important need in [2], our work opens up the space of opportunities for BCI.

References

[1] J. R. Wolpaw, N. Birbaumer, W. J. Heetderks, D. J. McFarland, P. H. Peckham, G. Schalk, E. Donchin, L. A. Quatrano, C. J. Robinson, and T. M. Vaughan. Brain-computer interface technology: a review of the first international meeting. *IEEE Transactions on Rehabilitation Engineering*, 8:164–173, 2000.

[2] A. Nijholt, D. Tan, B. Allison, J. Del R. Milan, and B. Graimann. Brain-computer interfaces for HCI and games. In *CHI 2008 Extended Abstracts on Human Factors in Computing Systems.*, pages 3925–3928. ACM, New York, NY, 2008.

[3] L. A. Farwell and E. Donchin. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalogr Clin Neurophysiol*, 70:510–523, 1988.

[4] G. Edlinger, C. Holzner, C. Grönegress, C. Guger, and M. Slater. *Foundations of Augmented Cognition. Neuroergonomics and Operational Neuroscience*, chapter Goal-orientated control with brain-computer interface., pages 732–740. Springer, 2009.

[5] C. Grönegress, C. Holzner, C. Guger, and M. Slater. Effects of P300-based BCI use on reported presence in a virtual environment. *Presence*, 19:1–11, 2010.

[6] M. Donnerer and A. Steed. Using a P300 brain-computer interface in an immersive virtual environment. *Presence*, 19:12–24, 2010.

[7] B. Yuksel, M. Donnerer, J. Tompkin, and A. Steed. Using a P300 brain-computer interface in an immersive virtual environment. In *Proceedings CHI 2010*, pages 855–858. ACM, Atlanta, Georgia, April 10-15 2010.

[8] P. Mistry, P. Maes, and L. Chang. WUW - Wear Ur World - a wearable gestural interface. In *CHI 2009 Extended Abstracts on Human Factors in Computing Systems.*, pages 4111–4116. ACM, Boston, USA, 2009.

Framework for Real-Time Decoding of ECoG Signals for Controlling the Anatomically Correct Test Bed (ACT) Hand

R. Scherer[1,2], T. Blakely[3], M. Malhotra[1], J. G. Ojemann[4], Y. Matsuoka[1], R. P. N. Rao[1]

[1]Department of Computer Science and Engineering, University of Washington, Seattle, USA
[2]Institute of Knowledge Discovery, Graz University of Technology, Graz, Austria
[3]Department of Bioengineering, University of Washington, Seattle, USA
[4]Department of Neurological Surgery, University of Washington, Seattle, USA

reinhold.scherer@tugraz.at

Abstract

Functional restoration of the upper extremity has emerged as a major aim of invasive brain-computer interface (BCI) research. To achieve this aim, not only must movement related neural signatures be investigated, but the contributions of biomechanics and multimodal sensory feedback to movement must also be understood. Since BCIs are designed for intentional, goal-directed communication, only real-time feedback experiments can confirm the validity of computational models. In this paper we describe a framework for real-time decoding of electrocorticographic (ECoG) signals for controlling the anatomically correct test bed (ACT) hand, an anthropomorphic robotic hand/arm with biological features. To demonstrate the viability of the proposed approach, we performed real-time ECoG-based control of the robotic hand by means of overt tongue movement.

1 Introduction

Decoding brain activity for the restoration of functional movements emerged as a major goal of brain-computer interfacing (BCI). Others and we have shown that simple, individual finger, hand, wrist or arm movements can be decoded from electrocorticographic (ECoG) signals (e.g. [1, 2]). Common to these studies are the ECoG features used to decode motor activity: focal oscillations in the high frequency band (> 30 Hz). Such oscillatory ECoG features have been recently used for the first time to control an anthropomorphic prosthetic hand in real-time [3].

Human hand dexterity, however, is achieved only partly by direct neural control of movements. The biomechanics of the system also play a very important role. Consequently, understanding the biomechanics of the human hand and its particular contribution to movement may significantly advance our ability to interpret and decode the measured neural activity. To better understand the biomechanics of the human hand we have worked on the development of the anatomically correct test bed (ACT) hand, an anthropomorphic robotic hand/arm with biological features such as human shaped bones and tendons [4].

We believe that the use of an anatomically realistic prosthetic hand as the controlled device in an ECoG BCI and the incorporation of multisensory feedback [5] will lead to the development of enhanced decoders for ECoG-based functional control of neuroprostheses and artificial limbs. As first step in this direction, we need a framework that allows decoding of ECoG during (c)overt hand movements and generation of control sequences for the ACT hand in real-time. In this paper, we briefly describe the different components of our framework for on-line decoding of brain activity and closed-loop control of the ACT hand. We report for the first time ECoG-based online control of the ACT hand via overt movement.

2 Methods

2.1 Components of the Framework

Figure 1 (a) shows the components of our system. Real-time signal processing is performed using Matlab (MathWorks, Inc., Natick, MA, USA) and the related data acquisition toolbox (DAQ). Due to safety issues and due to the limited space, it is usually not possible to bring bigger artificial devices such as the ACT hand into hospital rooms and close to the patient. Thus, for communication, the Internet protocol suite TCP/IP and for visual feedback the popular video call and conference system Skype (Skype technologies S.A., Luxembourg, Luxembourg) was used.

2.2 The Anatomically Correct Test Bed (ACT) Hand

The ACT hand mimics human biomechanics as 20 motor-driven tendons control the thumb, index, and middle fingers through unidirectional force application similar to that of muscles. Tendons are connected to human-shaped bones through an anatomically inspired tendon hood which matches range-of-motion, moment-arm variability, and mechanical coupling in order to produce human-like dexterous hand motion with equal control complexity to the human (Figure 1 (a)) [4, 6]. The motors are controlled by a computer system that allows accurate recording and closed-loop control of the excursion length for each tendon at 200 Hz; motor position is measured by an encoder with 40,000 ticks/revolution, equating to a resolution of $\sim$ 240 nm in excursion length. Functionally, control of individual joints, fingers, or grasps can be achieved by mapping from the desired motion into the tendon space. Mappings were built for analog control of several grasps with motion range scaled from 0 (flexion) to 1 (extension) so that complexity of tendon motions would be abstracted away for low-dimensional task control. Commands sent to the ACT hand over TCP/IP specify the desired analog level for each of the thumb, index, and middle finger's flexion amount given the specified grasp type (i.e. pinch, key, power).

2.3 Participant

The patient in this study was a 23-year-old man who was being treated at Harborview Hospital (HH) at the University of Washington (UW) for intractable epilepsy. The patient underwent implantation of a subdural electrode array above the right primary motor and somatosensory cortex and numerous subdural surface strips above parietal and temporal cortex to localize the seizure focus during a 7-day monitoring period (Figure 1 (a)). Placement of the recording arrays were dictated by clinical necessity. A preoperative MRI scan 3D reconstruction was combined with postoperative radiograph to determine the electrode grid locations and the underlying anatomical landmarks [7]. Informed consent was given by the patient in accordance with University of Washington Institutional Review Board protocol.

The patient, prior to this experiment, underwent functional motor screening. The screening showed that while there was coverage of primary motor cortex, cortical areas associated with hand movement were not covered. This would be expected given the electrode grid placement. The patient was, however, able to voluntarily modulate oscillatory activity in the high frequency range recorded from electrode #55 of the grid (Figure 1 (a)), located over primary motor cortex typically associated with homunculuar facial motor function. Since the patient was able to modulate high gamma activity by overt tongue movement, this simple control strategy was mapped to initiate ACT hand pinch movements in order to assess the functionality of the proposed tools.

2.4 Data Recording and Signal Processing

ECoG was recorded from sixteen electrodes (Figure 1 (b)). The data was filtered (0.5–250 Hz, Notch at 60 Hz), amplified and digitized at 2400 Hz. Each channel was re-referenced (common average reference) and the band power in the 76–160 Hz band was estimated by band pass filtering (5^{th} order Butterworth), squaring and averaging the squared samples over the past 200 ms. Control

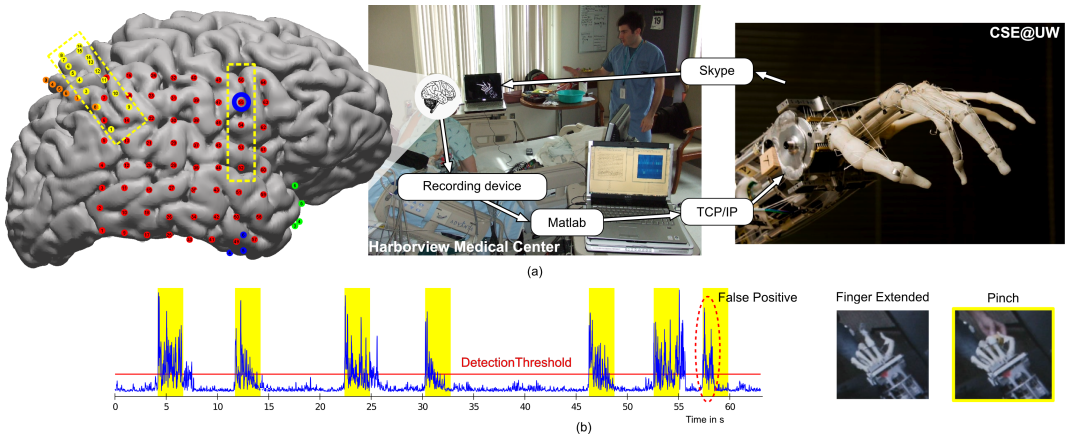


Figure 1: (a) The middle picture shows the setup in a patient room at the Harborview Medical Center, Seattle. The foreground laptop shows curves and color-coded traces of ECoG band power. The laptop in the background was used to provide visual feedback of the ACT hand that was located at the University of Washington, Seattle. The picture on the right shows the ACT hand with bone shaped fingers and tendons. Grid and strip placement on the patients' brain are depicted on the left. Only marked electrodes were recorded. Control signals were generated by electrode #55 on the grid (circled). (b) The curve shows the high frequency band power over the 1 min evaluation experiment. The horizontal line represents TH. Highlighted areas show pinch periods.

commands for the ACT hand were generated by using a simple threshold detector. Each time the focal band power value exceeded a predefined threshod TH a "pinch" control command was sent to the ACT hand. The starting position of the ACT hand was the extend finger position (Figure 1 (b)). Two seconds after exceeding the TH the ACT hand automatically switched back into the finger extended position.

The detection threshold TH was selected on-the-fly by visual examination of the band power time course during periods of rest and tongue movement. Subsequently, to get used to the system the patient was asked to autonomously switch between rest and tongue movement. After several minutes of independent self-paced training, a cue-guided experiment of 60 seconds duration was performed to assess the system performance. According to the verbal instructions of the experimenter (6 cues), the patient was asked to execute brisk tongue movements.

3 Results

By using the proposed framework, the patient succeeded in controlling the ACT hand by inducing spatially localized high frequency band oscillations modulated by overt tongue movements. During the 60 s evaluation run the patient was able to successfully activate the pinch command for the ACT hand six times as required by the paradigm. One false positive activation occurred towards the end of the test (Figure 1 (b)).

The performance of the available network connection was sufficient to send a control command from HH to UW and provide visual feedback within reasonable time delay. The round-trip time for TCP/IP packets sent from the laptop running Matlab at HV to the computer controlling the ACT hand at UW was less then 100 ms.

4 Discussion

In this paper we presented a framework for real-time ECoG translation and control of the ACT hand, and reported for the first time results from a patient controlling the ACT hand in real-time

by ECoG. Our framework includes a complete set of technologies and experimental paradigms for viable cortically controlled prostheses and will provide a solid platform for further studies.

The flexibility and rapid prototyping capabilities of the developed tools enabled us to screen the patient and provide real-time control of the ACT hand within minutes. The on-line experiment took place under real conditions, i.e., we used the available Internet access at HH and the UW. Although a lot of traffic can be expected within each and between the units, the time lag between the translation of ECoG signatures and the visual feedback of the ACT hand on the screen was reasonable, and neither troublesome nor annoying for the patient.

The patient in this study had no hand area motor coverage and we therefore used facial areas and tongue movement instead. We have shown that even with basic signal analysis and pattern recognition, volitional control was both viable and achievable. In previous studies, we started investigating the long-term stability of the high frequency band [8] and inferred computational models for predicting individual finger movements [1, 2]. In future studies with patients who have electrode grids over cortical hand motor areas, we intend to further examine the stability of the high frequency band, to evaluate and enhance the decoder performance and to investigate more complex motor tasks involving multi-finger synergies and compare these to isolated well-defined movements of individual digits or joints.

Acknowledgments

This research was supported by the National Science Foundation (0622252 & 0642848), NIH NINDS NS065186, the Microsoft External Research program, and the Packard Foundation.

References

[1] K. J. Miller, S. Zanos, E. E. Fetz, M. den Nijs, and J. G. Ojemann. Decoupling the cortical power spectrum reveals individual finger representation in humans. *J Neurosci* 29(10), 2009, pp. 3132-37.

[2] R. Scherer, S. Zanos, K. J. Miller, R. P. N. Rao, and J. G. Ojemann. Classification of contralateral and ipsilateral finger movements for electrocorticographic brain-computer interfaces. *Neurosurg Focus* 27(1), 2009, E12.

[3] T. Yanagisawa, M. Hirata, Y. Saitoh, T. Goto, H. Kishima, R. Fukuma, H. Yokoi, Y. Kamitani and T. Yoshimine. Real-time control of a prosthetic hand using human electrocorticography signals. *J Neurosurg* 114(6):1715-22.

[4] M. V. Weghe, M. Rogers, M. Weissert, and Y. Matsuoka. The ACT hand: Design of the skeletal structure. *IEEE ICRA*, 2004, pp. 3375-3379.

[5] A. J. Suminski, D. C. Tkach, A. H. Fagg and N. G. Hatsopoulos. Incorporating feedback from multiple sensory modalities enhances brain-machine interface control. *J Neuroscie* 30(50), 2010, pp. 16777-16787.

[6] A. Deshpande, J. Ko, D. Fox, and Y. Matsuoka. Anatomically correct testbed hand control: muscle and joint control strategies. *IEEE ICRA* 2009, pp. 4416-4422.

[7] D. Hermes, K. J. Miller, H. J. Noordmans, M. J. Vansteensel, and N. F. Ramsey. Automated electrocorticographic electrode localization on individually rendered brain surfaces. *J Neurosci Methods* 185(2), 2010, pp. 293-298

[8] T. Blakely, K. J. Miller, S. P. Zanos, R. P. N. Rao and J. G. Ojemann. Robust, long-term control of an electrocorticographic brain-computer interface with fixed parameters. *Neurosurg Focus* 27(1), 2009, E13.

Augmenting Gaze Control with a Brain-Computer Interface

B. F. Yuksel[1], A. Steed[1]

[1]Computer Science Department, University College London, UK

b.yuksel@cs.ucl.ac.uk

Abstract

We present a hybrid brain-computer interface (hBCI) composed of a motor-imagery-based brain switch and a head tracking device. Normal gaze-only (either head or eye gaze) interfaces suffer from the "Midas Touch" problem where unwanted selection of commands is triggered by subjects gazing at objects for too long. We use a BCI to provide a nontouch communication channel. Subjects were able to select and move objects in a fully immersive virtual environment. Object pick up and drop off was carried out by looking at the target object while using the BCI-based brain switch; object movement was controlled by the head tracker. The hBCI was compared with a control condition where gaze dwell time (DT) was used to pick up and drop off objects. Overall, the hBCI was just as fast and accurate as the DT condition, which highlights the potential for a hBCI to be used as an interaction device in a variety of user interface situations.

1 Introduction

A brain-computer interface (BCI) is defined as a control system where "commands do not depend on the brain's normal output pathways of peripheral nerves and muscles" [1]. BCIs are primarily used by people with severe motor disabilities, such as amyotrophic lateral sclerosis (ALS), to interact with a computer to improve their quality of life. A BCI can also be used as an additional communication channel to other signals and so enhance other assistive technologies. Such a BCI has been termed a hybrid BCI (hBCI) [2] and composes one BCI and one other system.

The hBCI that we present is composed of a head tracker and a ERD-based BCI. We demonstrate our hBCI in a CAVE-like immersive virtual environment (VE) (UCL CAVE), but this is just an example: the system could equally be used with desktop displays or other forms of interactive environment such as smart homes. ERD-based BCIs have been successfully used in CAVE VEs, for example, to navigate through a virtual street [3].

Our work is targeted at users with limited motor control (such as users with only free head or eye motion). Thus, the main users would be disabled users, but we can envisage a BCI addding an additional channel of communication in situations where the hands and voice might be occupied. Eye-gaze- or head-gaze-based interfaces are good examples of interfaces that can benefit from an additional communication channel such as a BCI. Normally, the gaze has to dwell on the interface object of interest for a period of time before selection or activation. This suffers from the "Midas Touch" problem where the user selects items unintentionally, which creates user frustration. We compare the performance of our hBCI with that of eye gaze dwell time (DT) to investigate whether an hBCI can have similar efficiency to a gaze-dwell-time interface.

2 Methods

2.1 Spatial Reasoning Task

To compare the performance of the hBCI and DT conditions we created a spatial reasoning task based on the rotation of 3D objects. The aim of the task was to select and manipulate an object

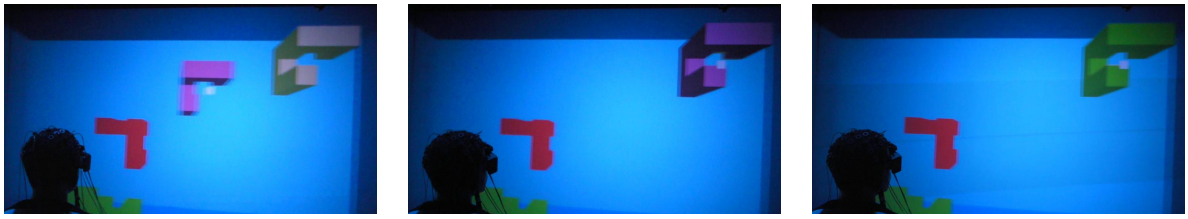


Figure 1: Left: A participant has selected the correct object and is moving it towards the target object. Centre: The object is placed on top of the target object. Right: The object has now been dropped by the participant and has turned back to green.

to match a target object. There were four objects to from which to choose (Figure 1). The correct object was set to a different rotation angle than the target object. Participants had 45 seconds in which to carry out the trial, after which there was a timeout whether they had completed the task or not. They received a 5-second rest between trials and completed 9 trials in total.

The aim of the task was to pick up the correct object, move it, and drop it on top of the target object. Participants were able to select an object by placing a white cursor on the bounding box of the object they required. When the correct object was selected, its colour turned to purple and it rotated to the correct rotation angle automatically. Once dropped off, it changed back to its original colour. If an incorrect object was selected it turned red and did not move.

2.2 Dwell Time Condition

Participants selected and de-selected an object by looking at it for 3.5 seconds. The DT was determined by pilot studies to find the best match to the level of difficulty of the task. Our DT was longer than those used in Sibert and Jacob [4] (150 ms) and Ware and Mikaelian [5] (450 ms) because we wanted to investigate tasks that had a cognitive load rather than the simple selection of a basic target. This was supported by Vilimek and Zander's [6] findings which is the closest study to ours. Their work was based on a hybrid eye-gaze BCI to perform selection of words on a screen. They found that a short DT reduced accuracy rates for their more difficult tasks to 51.1 %, whereas a longer DT provided accuracy rates of 75.6 %. Interestingly, they also found that the highest accuracy rates were in the longer DT condition for the easy task, and in the BCI condition for the difficult task. This may suggest that BCIs may be of assistance in more difficult tasks.

2.3 hBCI Condition

Participants selected and de-selected an object by carrying out hand motor imagery by imagining squeezing a ball. We used the wireless and portable g.tec g.MOBIlab+ 8 channel EEG system. Electrodes were placed at: Fz, Cz, P3, Pz, P4, PO7, Oz, and PO8 based on the international 10-20 system. The reference was placed on the right ear-lobe and the ground on the forehead. We used the BCI2000 software [7] to carry out EEG signal detection and filtering. We relayed the motor imagery data to the CAVE application over a UDP socket.

2.4 Overall Procedure

We examined the time taken to pick up and drop objects, the error rate, and the overall time taken, and we compared this to the performance of the DT system. Our six participants (all male; age range 21–23 years) were compensated. They were all students at UCL. We used a within-subjects design, which allowed us to both compare the performance with previous literature and also compare the performance of each participant with the hBCI and DT systems.

Participants were individually trained with the BCI2000 software using hand motor imagery for one hour. They were instructed to imagine squeezing a ball with their writing hand. The training

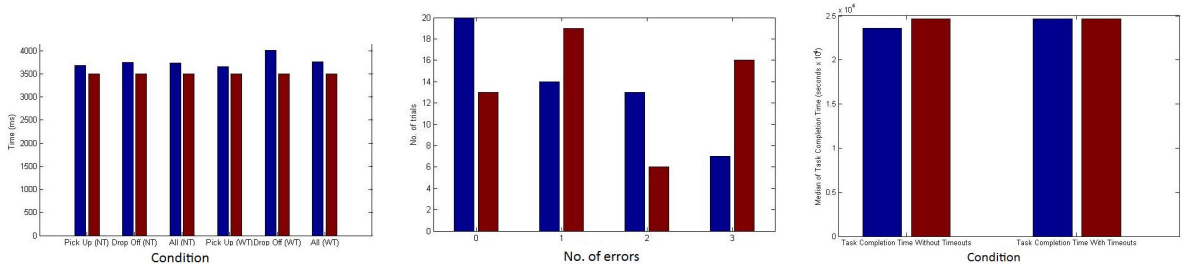


Figure 2: Blue denotes the hBCI system, red denotes the DT system. Left: Median time to pick up or drop objects. Centre: Error rates out of 54 trials. Right: Task completion times.

session both allowed for the training of the classifier and for the participants to improve their motor imagery. Two participants could not use the motor imagery BCI and did not take any further part in the study. On a different day, the six participants took part in the main experiment in the UCL CAVE using the classifier created during the training session. Each participant performed a series of 9 trials in the UCL CAVE in both the hBCI and DT systems. Half the participants first used the hBCI system, while the other half first used the DT system. Overall, the experiment lasted one hour per participant.

3 Results

In total, there were 54 trials for both the hBCI and DT systems. Five of the trials for the hBCI condition timed out. A timeout is defined as an unfinished task from the given 45 second timeslot. Four of these timeouts resulted in extremely long pick up/drop times which were outliers in the 50 trials. Therefore, statistical analysis between the two systems was carried out both with and without these timeouts. We used the Wilcoxon matched pairs test to compare the median of the three measurements: pick up and drop times, error rates, and task completion times. An error occurs when a participant selects the wrong object. As there are four objects from which to choose, one of which is the correct object, the maximum number of errors is three. We used the median instead of the mean because the data was not normally distributed.

Figure 2 shows the results of the two conditions both with and without the outlier timeout data. There were no significant differences between the medians of the hBCI and DT systems in error rates ($p = 0.1063$) and task completion times ($p = 0.8741$) (Figure 2). For the pick up/drop rates, there were no significant differences without the timed-out trials; however, the hBCI condition was slower with the timed-out data (Figure 2). Table 1 shows the results of the pick up and drop rates with the two significantly different values emboldened.

	Pick Up	Drop	Pick Up & Drop Rate
Without Timeouts	0.4544	0.0610	0.0613
With Timeouts	0.5611	**0.0095**	**0.0200**

Table 1: p-values of Wilcoxon matched pairs tests comparing the hBCI and DT systems.

4 Discussion

There were no significant differences between the hBCI and DT conditions for the time taken to pick up and drop objects (excluding two timeouts) in error rate or in task completion times. There

were 4 timeouts in the hBCI system which did not affect the overall significance of the results, but these timeouts suggest that some time-critical tasks may need slightly different mechanisms.

The only comparable study in the literature of which we know is that of Vilimek and Zander [6]. They created an eye-gaze hBCI with which 10 participants performed a search-and-select task to find words formed of consonants only. Vilimek and Zander compared the performance of shorter (1000 ms) and longer (1500 ms) DTs with their hBCI for "easy" and "difficult" tasks. Interestingly, their task completion times showed that the BCI condition was statistically slower than the DT conditions. In contrast, we found no significant difference in task completion times - in fact, the medians of our hBCI system were slightly faster; however, we did have longer dwell times.

Vilimek and Zander [6] also found the highest accuracy rates in the longer DT for the easy task, and in the BCI condition for the difficult task. In our experiment, we found no significant difference between accuracy rates. We had expected the DT system to produce more errors as participants selected unintended objects. However, some participants commented that the level of difficulty of our spatial reasoning task was too high. This could have led them to select each object randomly in both systems until they selected the correct one. This would correlate with the results found in the error rates (Figure 2).

We believe that it is possible to make our hBCI even better and possibly surpass the performance of a pure DT technique. Firstly, our BCI could be made asynchronous: with the cycle of the BCI2000 software, a participant could wait up to a maximum of 6 seconds before their motor imagery was picked up by the software. This is a considerable disadvantage especially since the DT was 3.5 seconds. Secondly, the participants could be given real-time feedback regarding the strength of their motor imagery in the form of an visual or audio bar in the CAVE.

In conclusion, we present the beginnings of an alternative to dwell time interfaces for situations where the user has limited motor control. We feel that there is much potential in building and developing hBCIs within and outside immersive virtual displays.

References

[1] J. Wolpaw, N. Birbaumer, W. J. Heetderks, D. J. McFarland, P. H. Peckham, G. Schalk, E. Donchin, L. A. Quatrano, C. J. Robinson, and T. M. Vaughan. Brain-computer interface technology: A review of the first international meeting. *IEEE Transactions on Rehabilitation Engineering*, 8(2):164–173, 2000.

[2] J. d. R. Millán, R. Rupp, G. R. Müller-Putz, R. Murray-Smith, C. Giugliemma, M. Tangermann, C. Vidaurre, F. Cincotti, A. Kübler, R. Leeb, C. Neuper, K. R. Müller, and D. Mattia. Combining brain-computer interfaces and assistive technologies: state-of-the-art and challenges. *Frontiers in Neuroscience*, 4:1–15, 2010.

[3] R. Leeb, C. Keinrath, D. Friedman, C. Guger, R. Scherer, N. C. M. Garau, A. Antley, A. Steed, M. Slater, and G. Pfurtscheller. Walking by thinking: The brainwaves are crucial, not the muscles! *Presence*, 15:500–514, 2006.

[4] L. E. Sibert and R. J. K. Jacob. Evaluation of Eye Gaze Interaction. *Proceedings ACM CHI'00*, pages 281–288, 2000.

[5] C. Ware and H. H. Mikaelian. An evaluation of an eye tracker as a device for computer input. *Proceedings ACM CHI'87*, pages 183–188, 1987.

[6] R. Vilimek and T. O. Zander. *Lecture Notes In Computer Science, Col. 5615: Proceedings of the 5th international conference on Universal Access in Human-computer Interaction. Part II: Intelligent and Ubiquitous Interaction Environments*, chapter BC(eye): Combining Eye-Gaze Input with Brain-Computer Interaction. Berlin/Heidelberg: Springer-Verlag, 2009.

[7] G. Schalk and J. Mellinger. *A Practical Guide to Brain-Computer Interfacing with BCI2000*. Springer-Verlag London Limited, UK, 2010.

Detection of Attempted Movement During Anesthesia as a Monitor of Intraoperative Awareness: Paradigm Development

Y. M. Blokland[1,2], J. Farquhar[2], J. Bruhn[1,2]

[1]Radboud University Nijmegen Medical Centre, Department of Anesthesiology, Pain and Palliative Care, Nijmegen, The Netherlands
[2]Donders Centre for Brain, Cognition and Behaviour, Nijmegen, The Netherlands

y.blokland@anes.umcn.nl

Abstract

The issue of intraoperative awareness has as yet not been resolved; during 0.1 to 0.2 % of all surgeries involving general anesthesia patients experience becoming aware, resulting in minor or major discomfort and possible development of serious psychological sequelae. This study explores the possibilities of using the event-related desynchronization (ERD) found in EEG during motor execution and imagery to detect attempted movement in aware patients. A 'brain switch' design using standard techniques from Brain-Computer Interfacing - but specifically adapted for use in the operating room - is proposed to serve as a monitor of awareness. Here, an overview of the requirements and considerations for such a system is presented. Results from a short pilot study indicate the feasibility of the proposed paradigm and provide some initial information for further development of brain switch paradigms in general and the proposed awareness detection paradigm in particular.

1 Introduction

Awareness during anesthesia is defined as consciousness during surgery and recall of intraoperative events when general anesthesia was intended, possibly leading to intra-operative pain, panic and anxiety. Whereas the main cause for awareness is an insufficient depth of anesthesia (DOA) [1], possible undesired consequences also exist for the opposite case of too deep anesthesia, e.g. hypotension and longer durations of recovery [2]. This means there is only a narrow adequate range of DOA, hence administration should be carefully monitored. However, anesthetic depth is not a single measurable variable, rather a complex reflection of the state of the central nervous system. In addition to more traditional ways of monitoring DOA, such as measurement of changes in blood pressure, heart rate, sweating and tear production, several types of EEG-based monitoring devices have been introduced, e.g. the Bispectral Index (BIS, Aspect Medical Systems, Massachusetts) and the Entropy Module (GE Healthcare, Helsinki). Despite these developments, unintended awareness during general anesthesia is still an unresolved issue with a current incidence of 0.1 to 0.2 %, amounting to approximately 26,000 cases annually in the United States alone [3].

General anesthesia involves the simultaneous administration of different components including neuromuscular blocking agents for immobilization. As a consequence of this induced paralysis, patients trying to move in order to alert the surgeon or anesthetist when awareness occurs fail repeatedly [1, 4]. Obvious parallels can be drawn with patients who are (partly) paralyzed by disease and for whom new methods of communication are currently under development. In certain Brain-Computer Interface (BCI) paradigms, frequency information in the EEG signal over the motor cortex during (imagined) movement or planning of (imagined) movement is used to create interaction between the user and a computer or other device. Based on this principle, we propose the development of a monitor of intraoperative awareness by means of detecting attempts to move. The current paper will shortly discuss important considerations and requirements for development

of this idea into a reliable, well working system. Also, the setup and results from an initial pilot study will be presented, serving as a starting position for further development.

2 Paradigm Development

In the proposed paradigm the event-related desynchronization (ERD), a feature commonly used in movement-based BCI's, will be exploited to detect a patient's urge to move and therefore function as input to a 'brain switch' type monitor. Intended movements lacking a measurable motor response (so-called 'quasi-movements') have been shown to generate a stronger ERD than the more commonly used imagined movements [5]. We hypothesize a similar effect for attempted but pharmacologically blocked movement. In the default situation, patients are under deep anesthesia. In the alternative case however, we expect attempted movement, thus eliciting a clear desynchronization in certain frequency bands. Ideally, the brain switch should detect these disruptions in the ongoing EEG, at any given time, unsynchronized to any specific time lock. Unfortunately, current asynchronous BCIs are not reliable enough for this clinical setting. Qian et al. [6] recently described a synchronous brain switch design in which subjects were instructed to turn on a virtual switch by performing an imagined movement task in response to an external cue. Such a design might be suitable for our situation as it processes ongoing EEG but is nevertheless time-locked. Despite the tradeoff of the need for patient instruction, this synchronous monitor would at least be a solution for patients with a high risk and/or fear of awareness. To arrive at a fully implementable system however, improvements in classification algorithms and BCI designs will need to be investigated, allowing for an asynchronous setting.

A second issue for not only the current paradigm but for many brain switch designs in general, is the importance of an extremely low False Positive Rate (FPR). Whereas in certain BCI's all types of errors are equally important and/or a certain rate of errors is allowed in order to improve the system's reaction time, in our application the amount of false positives needs to be kept close to zero as they will distract the anesthetist who is to remain focused at all times. In our opinion, to make it a meaningful addition to current monitors, it is unacceptable for the system to detect a 'hit' when attempted movement was in fact absent more than once per two to three hours of operating time. However, maintaining such a strict rate will increase the detection time for *true positives*. The tradeoff between the FPR and TPR (True Positive Rate, i.e. system speed) should be determined to set a feasible system sensitivity.

Thirdly, a fast system setup is clinically important. One way of achieving this is using only a small set of electrodes. Since movement ERD is highly localized, we hypothesize that a set of six channels over the motor cortex should be sufficient to obtain reliable information. In a previous brain switch study, promising results were even obtained for a single Laplacian channel only [7]. A second approach for a swift setup is to determine a fixed threshold power value beforehand above which a 'hit' is detected, rather than running a preoperative calibration session for the system to select the most useful features for classification. Experiments will be conducted to reveal whether a non-personalized system would be reliable enough or whether a user-specific calibration session is required, thereby increasing setup time.

Finally, low doses of anesthetic drugs may interfere with the EEG signal, possibly imposing difficulty on detection of ERD. The system setup and classification algorithm should be tested against this 'chemical noise'.

3 Pilot Study

A small pilot experiment was conducted to determine how reliably ERD during single hand movement as well as a more gross type movement of both arms/hands could be detected from EEG. Whereas the first type of movement is common in current BCI paradigms we expect the latter to be more representative of a patient trying to draw attention. While the ERD response of attempted movement remains to be tested in a future study, the current pilot study focused on the effect

of different system settings. Actual movement was chosen, as a stronger signal is more likely to clearly reveal performance differences. We examined whether using six Laplacian channels would be sufficient to obtain a reasonable classification rate, thereby allowing for a fast system setup. To remain representative of the clinical situation, participants kept their eyes closed and auditory rather than visual instructions were used. Furthermore, classification results using different time windows were compared. Assuming the strongest response occurs at the onset of a movement trial, combining several short trials might yield higher performances than using one continuous movement, thus making short movements preferable as a task in brain switch paradigms.

3.1 Methods

Ten participants were measured (24–44 years, 4 males). Subjects were presented with sequences of nine consecutive movement trials, each trial consisting of an auditory 3-second cue preceded by a random silence interval. Each sequence started with an auditory instruction explaining the task for the upcoming trials, i.e. the type of movement subjects had to perform during the entire duration of the cues (no movement, right hand movement, both arms movement). In total, 144 trials were collected per condition, equally divided over four experiment blocks. EEG was measured with a Biosemi Active-2 system using 64 electrodes placed according to the 10/20 system. Six channels were used for analysis (C3, C4, Cz, Cp3, Cp4, Cpz) with a surface Laplacian reference and the power spectral density computed for 8–28 Hz with 4 Hz bins using Welch's method. A linear logistic regression classifier was trained on the first block of 36 trials, representing a preoperative calibration session, and tested on the remaining data, comparing the 'no movement' condition with either the 'right hand movement' condition or the 'both arms movement' condition. Single trial classification results were calculated for each subject using time windows 0.0–1.0 s, 0.5–1.5 s, 1.0–2.0 s, 1.5–2.5 s and 2.0–3.0 s to determine the overall best one-second time window, as well as using data from the entire movement period (0.0–3.0 s). Sequence classification results using the combined outputs of three consecutive 1-second trials were calculated using the window with the average highest single trial results and compared to the single trial results for 0.0–3.0 s.

3.2 Results

Although for most subjects classification results were better in the 'both arms' condition than in the 'right hand' condition, average performances between the two conditions were rather similar. Using a 3-second time window, results reached up to 99 % for the best subjects, with an average of 87 % for right hand movement and 90 % for both arms movement (Table 1). Using the data from 0.5–1.5 seconds yielded slightly higher performance than using data from any other 1-second time window, hence this window was used for the sequence classification. Although in three subjects 100 % correct classification was reached for the 'both arms' condition when using three consecutive 1-second trials, average results for using either one 3-second trial or three 1-second trials were similar. Figure 1 shows the average classification rates for all three time window configurations.

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	avg
Right hand	92	81	81	91	88	90	81	92	88	87	87
Both arms	99	88	92	94	81	99	78	93	82	90	90

Table 1: Single trial classification results per subject (S1–S10) and averaged over all subjects (avg) for the 3-second window. In the 'right hand' condition subjects made a tapping movement with their right hand, in the 'both arms' condition subjects made a gross movement with both arms.

4 Discussion

We described the possibility of developing a brain switch design to detect attempted movement during general anesthesia. We expressed the need for an extremely reliable and stable system with

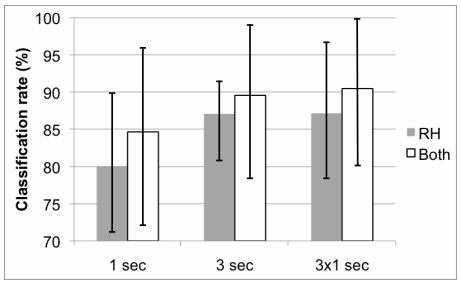


Figure 1: Classification results using data from 0.5–1.5 s, 0.0–3.0 s and the combined classifier outputs using 0.5–1.5 s of 3 consecutive trials, averaged over all subjects. 'RH' denotes right hand movement, 'Both' is movement of both arms. The error bars show the spread between subjects.

a fast setup. Initial results from a non-clinical pilot study indicate that single trial movements can reliably be classified using six EEG channels in healthy unanesthetized subjects with their eyes closed. Overall, no major difference was found between movement of both arms or of a single hand.

Our results do not indicate general performance increase when combining several short movements rather than using one prolonged movement. This is encouraging when moving towards an asynchronous paradigm; if ongoing movement can reliably be detected there is no need for a repetitive task-related cue. Further research will be conducted to test whether this still holds for movements longer than three seconds.

Studies are currently being set up to further test the implementation issues mentioned above, i.e. comparing a generic, non-calibrated system with a personalized setup thus requiring calibration time, exploring the possibilities of building an asynchronous design and testing the effect of anesthetic drugs on system reliability. Additionally, it will be determined whether the current results can be reproduced for attempted but pharmacologically blocked movement and whether performance can be improved by optimizing the choice of classification features.

References

[1] M. Ghoneim, R. Block, M. Haffarnan, and M. Mathews. Awareness during anesthesia: risk factors, causes and sequelae: a review of reported cases in the literature. *Anesthesia & Analgesia*, 108:527–535, 2009.

[2] C. D. Kent and K. B. Domino. Depth of anesthesia. *Current Opinion in Anaesthesiology*, 22:782–787, 2009.

[3] P. S. Sebel, T. A. Bowdle, M. M. Ghoneim, I. J. Rampil, E. Padilla, T. J. Gan, and K. B. Domino. The incidence of awareness during anesthesia: a multicenter United States study. *Anesthesia & Analgesia*, 99:833–839, 2004.

[4] R. H. Sandin, G. Enlund, P. Samuelsson, and C. Lennmarken. Awareness during anaesthesia: a prospective case study. *The Lancet*, 355:707–711, 2000.

[5] V. V. Nikulin, F. U. Hohlefeld, A. M. Jacobs, and G. Curio. Quasi-movements: a novel motor-cognitive phenomenon. *Neuropsychologia*, 46:727–742, 2008.

[6] K. Qian, P. Nikolov, D. Huang, D. Fei, X. Chen, and O. Bai. A motor imagery-based online interactive brain-controlled switch: paradigm development and preliminary test. *Clinical Neurophysiology*, 121:1304–1313, 2010.

[7] G. Pfurtscheller and T. Solis-Escalante. Could the beta rebound in the EEG be suitable to realize a "brain switch"? *Clinical Neurophysiology*, 120:24–29, 2009.

Hierarchical EEG Assessment Paradigms Designed for Non- and Low-Responsive Patients

S. Veser[1], B. Kotchoubey[1]

[1]Institute of Medical Psychology and Behavioural Neurology, Tübingen, Germany

sandra.veser@uni-tuebingen.de

Abstract

In the present article we describe a systematic hierarchical approach to the investigation of remained cognitive functions in low- and non-responsive patients. We subdivided the hierarchy of cognitive functions underlying consciousness into three levels: a "basic" level (e.g., tonal discrimination), a "mid-grade" level (e.g., semantic priming), and the level of "higher" cognitive functions (e.g., directed attention). For the first two levels, five different event-related potential (ERP) paradigms in the auditory domain were described, with systematically increasing complexity in terms of stimulus material (e.g., tones versus semantic stimuli). While it is long known that, at the level of group analysis, these paradigms elicit particular cognitive processes manifested in particular components, here we show that this is true for single subjects as well. Therefore, the described paradigms can be used for testing low- and non-responsive patients building up an algorithm to complement the clinical assessment. In a larger perspective, this would reduce the presently very high rate of misdiagnoses in such patients.

1 Introduction

The clinical assessment of low-and non-responsive patients is commonly based on the patients' explicit overt behaviour. This behavioural clinical assessment is error-prone due to a number of factors (e.g., the patients' dependency on the ability to move etc.), therefore resulting in an alarmingly high number of misdiagnoses [1]. Recently the neurophysiological methods like the electroencephalography (EEG) and the functional brain imaging (fMRI) gained more and more significance in testing the remained cognitive functions independently from the behavioural responses of such patients. Earlier works indicated that the most effective way to examine these functions is to follow a hierarchical approach [2,3]. In addition, this approach helps to find out behaviourally low-or non- responsive patients able to use a BCI by detecting their residual conscious awareness.

The approach presented here is to subdivide the hierarchy of cognitive functions underlying consciousness into three levels. The first level includes basic cognitions (e.g., tonal discrimination) which can be completely unconscious and subpersonal, but without which conscious perception is impossible. This level is crucial if more difficult tasks should be accomplished. The second level includes deeper information processing (e.g., semantic processing) that indicates, but not proves, conscious intentional states; and the third level includes higher cognitive function like attentional shifts, in which patients are able to follow instructions and self-instructions. Here, we concentrate on the paradigms of the first two levels. However, a systematical approach to establish a useful clinical assessment will involve all three levels tested in a randomly set order.

Some of the ERP paradigms discussed here have already been tested with non-responsive patients [2]. However, because the paradigms were originally selected on the basis of their efficiency in group studies, considerable work remained for methodological adjustment of electrophysiological procedures to the principal task of the assessment of cognitive functions in single subjects. This

concerns the features of stimuli (e.g., chords versus sine tones; or the selection of words and sentences), their number (to optimize the signal-to-noise-ratio while avoiding very long and tiring examinations), etc. In particular, the issue of using a multifeature oddball paradigm (see below, the methods section) instead of several one-feature oddball paradigms had to be investigated.

2 Methods

2.1 Participants

Six healthy right-handed healthy native german speakers (2 males, mean age: 27 years) without any neurological disorder or hearing problems participated in this study. They received €8 per hour. The volunteers gave written informed consent after they were informed about the nature of the study.

2.2 Stimulus and Procedure

Five auditory paradigms affiliate to the first and second level of cognitive function were tested. Two of the paradigm belong to the first level of cognitive functions and were oddball paradigms. In the simple oddball paradigm (1) complex harmonic tones ($440 + 880 + 1760$ Hz) with a duration of either 50 ms (Standard, presented 90 % of the time) or 20 ms (Deviant, presented 10 % of the time) and a Stimulus-Onset Asynchrony (ISI) of 450 ms (c.f., [2]). In the multifeature oddball (2) principally adapted by [4] five deviants were used, which differed either in their frequency, duration, location, intensity and complexity. The interesting point by using this design is that we can gain the information of processing ability of five different features in one coup instead of implementing five different classical oddball paradigms in the hierarchical assessment. The difficulty of this design is that the brain needs to be able to process each feature separately because the other deviants play the part of standards, together with standards, against a given deviant. This is especially critical because the mismatch negativity (MMN) system is supposed to be operating rather on the bases of the individual features than on gestalt principals [5]. In both paradigms the ERP marker is the early MMN, which reflects an automatic pre-attentive process.

Three other paradigms belong to the second level of cognitive functions. In the simple oddball paradigm (3) a frequent complex tone (Standard: $440 + 880 + 1760$ Hz) and a rare complex tone (Deviant: $247 + 494 + 988$ Hz) were used with an ISI of 850 ms. In the word-prime paradigm (4) semantic processing of the word level is tested. Two hundred 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. The expected ERP-component is an N400 that should differ between primed and unprimed second words in pairs [6]. In the sentence understanding (5) paradigm semantic processing at the sentence level is tested. 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). To avoid the simple word priming effect in this paradigm, the key word of the sentence (e.g., eel in the examples above) was separated by a least two words from the ending. ISI between sentences were 900 ms. The expected ERP-component is here also an N400 difference between congruent and incongruent ending [7].

2.3 EEG and Analysis

The EEG was recorded according to the international 10-20 electrode system with active electrodes at the F3, Fz, F4, C3 Cz, C4, P3, Pz and P4 using gtec as a portable system on a notebook computer for EEG recording at the patient's bedside. The vertical (VEOG) and horizontal electrooculogram (HEOG) were recorded by using two pairs of electrodes attached at the outer canthi of the two eyes and below and above one eye. The EEG was digitized at 512 Hz and filtered with a bandpass filter between 0.1 Hz and 100 Hz.

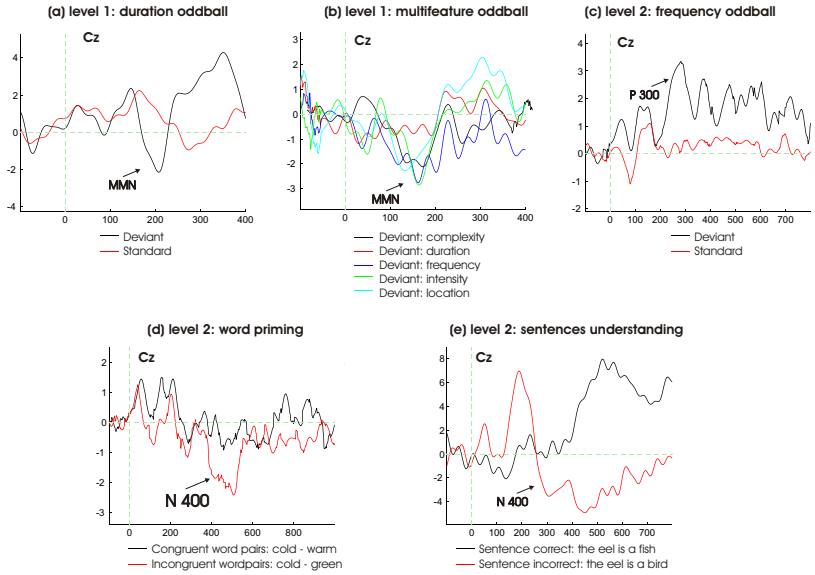


Figure 1: Cognitive hierarchical assessment of the first two levels shown for a single subject.(a) depicts the Standards and Deviants which differ in the time-window of the MMN, (b) shows the difference waves (Deviants-Standard) for all five different features which reflected an MMN, except for the feature duration, (c) shows a P300 to the rare stimulus, (d) depicts the N400 response to word-pairs when the pairs were unrelated and (e) shows the N400 response to sentences with an incongruent ending.

Offline, the EEG was filtered with Kaiser low-pass filter of 25 Hz, and ocular artifacts were corrected using the blind source algorithm sobi (window length of 2 s, shift of 4 s) using the Automatic Artifact Removal (AAR) toolbox (http://www.cs.tut.fi/ gomezher/index.htm). The continuous EEG data were split into epochs which lasted between 500 ms and 1000 ms (dependent on the respective paradigm) including a 100 ms baseline. Trials for each paradigm were averaged to obtain an ERP. The components of interest for the respective paradigm were analyzed by visual inspection. Mean amplitudes were measured in each single trial within a time window appropriate for the respective component dependent on analysed paradigm.

3 Results

For all paradigms, we found the expected ERP components (see Figure 1 for an exemplified single subject). To define the ERP components, we visually inspected the ERPs, we computed running t-tests of the target condition to the baseline and running t-tests between the two conditions. Particularly, in the MMN duration oddball the expected MMN component was found in all five participants. The results of the multifeature paradigm were less consistent. The MMN component for frequency, location and intensity deviants were found in all participants. However, only one participant exhibited an MMN response to the complexity deviant, and none of the participants had an MMN response to the duration deviant. In the frequency oddball, all five participants demonstrated a larger P300 response to rare stimuli than to frequent stimuli. In the word-prime paradigm, four of five healthy volunteers showed a clear detectable N400 response to unrelated word-pairs compared to related word-pairs. In the sentences paradigm, the brain responses of four from five participants showed a larger N400 response for semantically incorrect sentences compared to correct sentences.

4 Discussion

The results show that the developed paradigms result in consistent ERP in individual participants. The multifeature MMN paradigm should be used in patients along with the single-feature paradigms until enough data is accumulated that one of them works better than the others. The multifeature MMN paradigm would be favourable because it would help us to gather more information about the pre-attentive discrimination ability of the patients in a shorter time interval (5 deviants are examined in around 15.30 minutes) compared to the classical oddball paradigm (1 deviant is investigated in 6.30 minutes). In parallel to these clinical tests, novel paradigms should be developed in the next future aimed at the examination of the third (upper) level of the supposed hierarchy.

5 Conclusion

In conclusion, the tested paradigm can be used to assess systematically the remaining cognitive functions in non- and low-responsive patients to build up an algorithm to complement clinical assessment.

Acknowledgments

This work was supported by the European ICT Programme Project FP7-247919. The text reflects solely the views of the authors. The European commission is not liable for any use that may be made of the information contained therein.

References

[1] C. Schnakers, A. Vanhaudenhuyse, J. Giacino, M. Ventura, M. Boly, S. Majerus, G. Moonen, and S. Laureys. Diagnostic accuracy of the vegetative and minimally conscious state: clinical consensus versus standardized neurobehavioral assessment. *BMC Neurology*, 9:35, 2009.

[2] B. Kotchoubey, S. Lang, G. Mezger, D. Schmalohr, M. Schneck, A. Semmler, V. Bostanov, and N. Birbaumer. Information processing in severe disorders of consciousness: vegetative state and minimally conscious state. *Clinical Neurophysiology*, 116(10):2441–2453, Oct 2005.

[3] A. M. Owen, M. R. Coleman, D. K. Menon, E. L. Berry, I. S. Johnsrude, J. M. Rodd, M. H. Davis, and J. D. Pickard. Using a hierarchical approach to investigate residual auditory cognition in persistent vegetative state. *Progress in Brain Research*, 150:457–471, 2005.

[4] R. Näätänen, S. Pakarinen, T. Rinne, and R. Takegata. The mismatch negativity (MMN): towards the optimal paradigm. *Clinical Neurophysiology*, 115(1):140–144, Jan 2004.

[5] D. Deacon, J. M. Nousak, M. Pilotti, W. Ritter, and C. M. Yang. Automatic change detection: does the auditory system use representations of individual stimulus features or gestalts? *Psychophysiology*, 35(4):413–419, Jul 1998.

[6] P. Hagoort, C. M. Brown, and T. Y. Swaab. Lexical-semantic event-related potential effects in patients with left hemisphere lesions and aphasia, and patients with right hemisphere lesions without aphasia. *Brain*, 119 (Pt 2):627–649, Apr 1996.

[7] J. F. Connolly, N. A. Phillips, S. H. Stewart, and W. G. Brake. Event-related potential sensitivity to acoustic and semantic properties of terminal words in sentences. *Brain and Language*, 43(1):1–18, Jul 1992.

A P300 BCI with Stimuli Presented on Moving Objects

I. P. Ganin[1,2], S. L. Shishkin[1,2], A. Y. Kaplan[1,2]

[1]Lomonosov Moscow State University, Faculty of Biology, Laboratory for Neurophysiology and Neuro-Computer Interfaces, Moscow, Russia
[2]National Research Nuclear University MEPhI, Moscow, Russia

ipganin@mail.ru

Abstract

In the P300 brain-computer interface, visual stimuli are presented at spatially fixed locations. Can this BCI work if the stimuli positions are allowed to move, e.g., when attached to different moving parts of robotic devices or to virtual objects in a video game? We designed a simple P300 BCI game with moving locations of stimuli and tested it in a four session experiment. Able-bodied participants played this game in either a single-trial ($n = 6$) or a triple-trial ($n = 6$) mode through all sessions. All of them performed better than randomly, and most of them maintained a high level of interest to the task up to the last sessions. Our study demonstrates that the P300 BCI can be extended to a version with moving stimuli positions.

1 Introduction

The P300 brain-computer interface (P300 BCI) provides a relatively high information transfer rate, while requiring no special training of the user and little time for classifier calibration. It currently appears to be the most commonly used BCI [1]. However, its static visual design pose limitations for possible applications. For example, in highly engaging video games moving elements usually play important roles. Creating an attractive game on the basis of a static "control panel" of the standard P300 BCI is a difficult task. Attaching the stimuli positions to the moving objects on which the player focuses his/her attention seems to be a much more prospective way to develop an engaging BCI controlled game. Control of prosthetic or robotic devices could be also more flexible if the BCI stimuli would be allowed to be presented at freely moving positions.

The use of moving objects as stimuli have been proposed in [2]. However, in this study the stimuli were presented at fixed positions. In our recent study [3] we demonstrated that brain event-related potentials (ERP) remain highly sensitive to the difference between the attended and not attended items when the P300 BCI stimuli matrix moves. A more radical solution is to make moving each stimulus position separately. We already found that the distance between stimuli positions in the P300 BCI matrix has little effect on the ERP difference between target and non-target stimuli [4], thus, the change of the distances during the motion unlikely can impair the BCI performance. Nevertheless, many other factors can, in principle, influence the P300 wave and the other ERP components important for discrimination of the brain responses [3], and an experimental proof is needed to confirm that the motion of the items on which the stimuli are presented would not prevent recognition of the user's commands.

In this paper, we describe preliminary results of testing a P300 BCI based game with stimuli presented on objects which *move on individual trajectories*.

2 Methods

Similarly to our previous (static) P300 BCI game, the BCI Puzzle [5], the goal of the player in the new game is to assemble a full picture from its fragments. The main difference is that in the new game each fragment is put into a ball moving on the computer screen (Figure 1). For research

Figure 1: Example of the BCI game display. Each moving "ball" contains a fragment of the picture being assembled in the right. The correct order of targets is cued by letters (of Russian alphabet). A ball with letter "B" is "highlighted". In the bottom right, a counter shows the number of errors.

purposes, it was reasonable to simplify finding the current target (the ball to be pursued and attended): first, the balls were marked with the letters of Russian alphabet (all the participants were native Russian speakers) and the target order was always alphabetical; secondly, the current target was highlighted in a result panel where the puzzle was assembled.

12 healthy volunteers participated in four sessions on different days after signing the informed consent. Each session started from a calibration, and then the participants played, with short breaks, 10 games, each one with a new picture (a color photograph of an animal, plant or a car cut into 9 fragments). The fourth session consisted of only 5 games, as additional tests were done after the games (their results will be described elsewhere). In each run, after finding the target and preparing to attend the target stimuli, they pressed a mouse button, initiating start of the stimulation in 3 s.

For six participants (the single-trial group), each ball always flashed once, and the classifier was applied to single-trial EEG epochs related to each stimulus. For another six participants (the triple-trial group), three random "sequences" of flashes were given (each ball flashed three times in total), and the averages of three responses to stimuli related to each ball were used for classification. The player had to mentally note the flashes of the target item and pay no attention to flashing of all the other items.

If the ball classified as attended was the target ball, the related item in the picture in the result panel was filled, and the next to it became the new target. Otherwise, the error counter showed an increase by one and the target remained the same as in the previous run. "Winning" a game meant that all 9 fragments were identified as attended and the full picture was assembled. A game terminated before picture was assembled ("lost" game) if 10 errors was made in this game.

The ball size was 2 °, the movement trajectories were linear (with natural change of direction after collisions), and the speed was 5 °/s. They were presented on a CRT monitor at a distance about 85 cm from the eyes. A stimulus was an increase of brightness of a ball (see an example in Figure 1) for 125 ms. Balls flashed in a random order without pauses between the flashes. EEG was recorded at Cz, Pz, PO7, PO8, O1, O2 against a reference at the right earlobe, bandpass filtered in the range 1–10 Hz and decimated down to 20 Hz. A single EOG channel was also recorded. Each data epoch started from the stimulus onset and its length was 1 s. Channel data in each epoch were concatenated and formed a feature vector. Classifier weights were obtained with Fisher Linear Discriminant Analysis.

The participants rated their average interest to the task during the experiment by putting a mark on a Visual Analog Scales (VAS) in the end of the session (0 corresponded to "not interesting" and 100 to "extremely interesting"). The scales for different sessions were positioned on the same sheet one below another. This was done to allow and to encourage the participants to use the previous estimates as a basis for finding a position for the new estimate in relation to them.

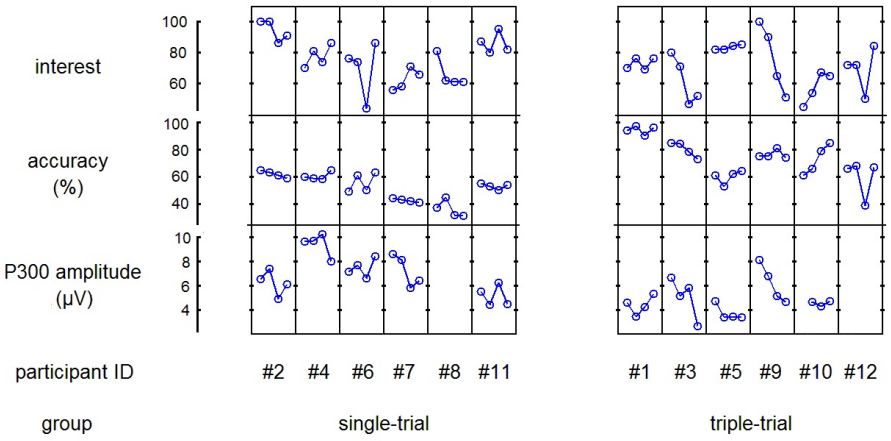


Figure 2: Self-estimated interest to task during the session, BCI accuracy and the P300 amplitude at Pz in each participant.

In the offline analysis, the P300 amplitude was measured at Pz after filtering the signal in 0.5–20 Hz band and ERP averaging, as the maximum value in the 250–500 ms interval relative to the stimulus onset. In a few cases, the P300 amplitude could not be estimated: two participants did not have a clear P300 (instead, a positive wave with latency around 200 ms was present in responses to targets); same applied to one more participant but only in his first session. Only few epochs (2 % of all epochs, on average) contained EOG artifacts; these epochs were excluded from analysis. It seemed that strong artifacts from saccades were not common in our data, as attending the moving object required only smooth pursuit eye movement and small saccades. Blinking artifacts were also rarely found, probably because the stimulation periods were short.

3 Results

Interest to task, accuracy and P300 amplitude data per participant and session are shown in Figure 2.

Interest to task in single-trial group was 78(15) (M(SD)) in the first session and 79(12) in the last one. In triple-trial group, group averaged interest values failed from 75(18) in the first session to 69(15). 2-way MANOVA showed no significant effects for factors Session, Group and for their interaction.

In triple-trial group, participant's *online accuracy* per session ranged from 39 % to 97 %, and for single-trial, from 31 % to 65 %. Thus, all subjects in all sessions performed better than at the random level, which was 11 % (as one of 9 items should be chosen in each run). According to 2-way MANOVA (Group x Session), the triple-trial group demonstrated higher accuracy than the single-trial group ($F(1, 10) = 10.8$, $p = 0.008$). However, no dependence on session was revealed in both groups.

P300 amplitude at Pz was higher in the single-trial group than in the triple-trial group and showed a tendency to decline across the sessions: effect of Group was significant ($F(1, 7) = 6.0$, $p = 0.04$), effect of Session was marginally significant (Wilk's lambda $= 0.25$, $F(3, 5) = 4.9$, $p = 0.06$) and the interaction between Session and Group factors was not significant (Wilk's lambda $= 0.33$, $F(3, 5) = 3.4$, $p = 0.11$).

Small group sizes and variations in individual dynamics across sessions (Figure 2) do not allow a detailed analysis of the relations between the studied variables. In fact, positive intra-individual correlations, especially between interest to task and the P300 amplitude, were high (in participants #9 and #11, Pearson correlation coefficient was 0.98), but negative correlations were also observed

(especially high for #7, for whom it was -0.99). The highest correlations could be, however, a result of trends existing in both variables (participants #7, #9) and appear just by chance in the rest of cases (e.g., in #11).

4 Discussion

In this paper, a P300 BCI modification with stimuli presented on moving object was described for the first time, to the best of our knowledge. Such modification could, in principle, negatively affect the performance of the BCI, for example, because of attention distracting effects from stimuli presented on objects moving around the target object and even colliding with it. However, the accuracy obtained in this study with a simple classifier was already sufficient for playing a game and for maintaining high interest to task during all four sessions even in a group using the single-trial mode of BCI operating.

The results obtained in this preliminary study are already encouraging for the application of the P300 BCI to games for healthy players, who can easily pursue the moving targets by gaze. In heavily paralyzed persons, gaze control is often impaired, and additional studies are needed to determine if it is possible to operate the P300 BCI with moving positions of stimuli without pursuing the stimuli by gaze.

5 Conclusion

This study demonstrated that the P300 BCI can operate efficiently when stimuli are presented on moving positions.

6 Acknowledgments

Authors thank the anonymous reviewers for useful comments and suggestions. Authors also thank Alexei Preobrazhensky for programming. This study was partly supported by the Federal Targeted Program "Scientific and Scientific-Pedagogical Personnel of Innovative Russia in 2009-2013" (contract P1087) and by grant from the Foundation for Assistance to Small Innovative Enterprises (UMNIK program, project 10228, theme 3; Start program, contract 7606r/10342).

References

[1] J. N. Mak, Y. Arbel, J. W. Minett, L. M. McCane, B. Yuksel, D. Ryan, D. Thompson, L. Bianchi and D. Erdogmus. Optimizing the P300-based brain–computer interface: current status, limitations and future directions. *Journal of Neural Engineering*, 8:025003, 2011.

[2] F. Guo, B. Hong, X. Gao, and S. Gao. A brain–computer interface using motion-onset visual evoked potential. *Journal of Neural Engineering*, 5:477-485, 2008.

[3] S. L. Shishkin, I. P. Ganin, and A. Y. Kaplan. Event-related potentials in a moving matrix modification of the P300 brain-computer interface paradigm. *Neuroscience Letters*, 469:95-99, 2011.

[4] I. P. Ganin. The N1 component of brain potentials and the spatial factors in the P300 brain-computer interface. *Proc. of the XIV young scientist conference on the physiology of higher nervous activity and neurophysiology. 21-22 Oct. 2010, IHNA RAS*, p. 37, 2010 (in Russian).

[5] A. J. Kaplan and S. V. Logachev. *Patent RU 2406554 C1*, 2009.

Simultaneous EEG Recordings with Dry and Wet Electrodes in Motor-Imagery

J. Saab[1,2], B. Battes[1], M. Grosse-Wentrup[1]

[1]Max Planck Institute for Intelligent Systems, Dept. Empirical Inference, Tübingen, Germany
[2]Graduate School of Neural and Behavioural Sciences - International Max Planck Research School, Tübingen, Germany

jad.saab@tuebingen.mpg.de; bernd.battes@tuebingen.mpg.de;
moritz.grosse-wentrup@tuebingen.mpg.de

Abstract

Robust dry EEG electrodes are arguably the key to making EEG Brain-Computer Interfaces (BCIs) a practical technology. Existing studies on dry EEG electrodes can be characterized by the recording method (stand-alone dry electrodes or simultaneous recording with wet electrodes), the dry electrode technology (e.g. active or passive), the paradigm used for testing (e.g. event-related potentials), and the measure of performance (e.g. comparing dry and wet electrode frequency spectra). In this study, an active-dry electrode prototype is tested, during a motor-imagery task, with EEG-BCI in mind. It is used simultaneously with passive-wet electrodes and assessed using offline classification accuracy. Our results indicate that the two types of electrodes are comparable in their performance but there are improvements to be made, particularly in finding ways to reduce motion-related artifacts.

1 Introduction

Wet electrode preparation presents a major obstacle to widespread, day-to-day use of EEG-BCI, especially for patients with impaired mobility. It is a time-consuming process, due in large part to scrubbing of the scalp to improve signal quality; it can be an unpleasant experience for subjects, especially after frequent sessions that heighten skin-sensitivity; and, over hours of use, it requires regular maintenance as the conductive gel dries and degrades signal quality [1]. Dry electrodes are a promising solution to these problems, potentially reducing set-up times, subject discomfort, and the need for maintenance.

Dry electrode performance is often compared to that of existing wet electrodes. However, for a given experimental paradigm and subject, if dry and wet electrodes are recorded in separate trials, variation between trials can distort electrode comparisons. Simultaneous recording with the two types of electrodes avoids this problem. Comparisons of performance in simultaneous active-dry and wet electrode studies have included visual inspection of spontaneous EEG waveforms [2–4] or power spectral densities (PSDs) [3, 4], hypothesis testing of evoked potential (EP) properties such as amplitude [3,5], hypothesis testing of band-power values for an alpha-rhythm concentration task [5], and classification accuracies in EP and alpha-rhythm concentration tasks [5]. Comparisons of performance in simultaneous passive-dry and wet electrode studies have included correlation of electrodes during reactive alpha and motor-imagery tasks [6] and the coherence of electrodes [7]. In general, performance levels of dry electrode systems are found to be comparable to traditional gel-based electrodes. However, the small number of subjects ($\leq$ 12) employed in these studies limits the conclusiveness of findings.

In this article, we present a 20-subject study on motor-imagery (the most frequently used paradigm in BCI-research [8]) in which we record simultaneously with active-dry and passive-wet electrodes, and compare their performance using classification accuracy.

2 Methods

2.1 Experimental Design

Subjects were instructed to perform kinesthetic motor-imagery of the right hand [9]. It was emphasized that they try to feel, rather than visualize, the imagined motion. A computer monitor approximately 1.5 meters away from the seated subjects provided visual prompts to start and stop motor-imagery. The display was black except for a colored square containing a black fixation cross, both centered on the screen. The color of the square represented the two experimental conditions, rest and imagery. For the rest condition trials, the square was gray and subjects were asked to relax and maintain fixation on the cross. For the imagery condition trials, the square was green and subjects were asked to perform motor-imagery. Each trial lasted between 6.25 and 7.25 seconds, selected randomly. The sequence of trials was pseudo-randomized but the number of consecutive rest or imagery trials was limited to 3 so as to avoid subject fatigue during prolonged periods of the same condition. Two runs, each consisting of 30 trials per condition, were recorded for every subject, with a short pause in between. This resulted in a total of 120 trials per subject. Each subject provided informed consent in accordance with guidelines set by the Max Planck Society.

2.2 Hardware Setup

Standard off-the-shelf passive-wet electrodes were used. The active-dry electrodes, Figure 1 (A), were designed and manufactured at the Max Planck Institute. They were made to attach to an EEG cap using modified disc-electrode connectors. The connectors were reinforced inside the cap, increasing the area of contact with the head and improving electrode stability. Pressured air, applied to a chamber within the electrodes' housing, is used to protrude an array of 19 gold-plated pins and regulate their contact force. The signal conducted by the pins is then amplified by an op-amp circuit contained within the electrodes.

2.3 Experimental Data & Data Analysis

EEG was recorded using two pairs of electrodes, one wet and one dry, simultaneously. To capture modulation of the sensorimotor-rhythm (SMR) resulting from motor-imagery of the right hand, one electrode pair was positioned at C1 and C5 and the other at Fc3 and Cp3, according to the 10-20 system. The position of each electrode pair was counter-balanced across subjects, i.e., half the subjects had the wet pair positioned at C1-C5 and the dry pair at Fc3-Cp3 and the other half had the dry pair positioned at C1-C5 and the wet pair at Fc3-Cp3. Signals were sampled at 500 Hz using a QuickAmp amplifier (BrainProducts GmbH, Gilching, Germany) with a built-in common-average reference. A wet ground electrode was attached to subjects' right ear lobes. Twenty-three subjects (S1-S23) participated, 9 of whom were female, with a mean age of 28.70 years and a standard deviation of 5.90 years. All subjects were right-handed and none had known neurological disorders. Two subjects were excluded from further analysis since their recordings showed millivolt-scale amplitudes well beyond those of typical EEG. In both cases, hair thickness seemed to prevent the dry electrode pins from contacting the scalp effectively. A third subject was excluded after observing that only 50 Hz mains noise was being measured. The remaining subjects' recordings showed EEG-like amplitudes and typical EEG artifacts associated with muscle movements (e.g. clenching of the jaw) and eye movements.

For each subject, the amplifier's built-in common-average was removed from the data by calculating two bipolar recordings, one between the two wet electrodes and the other the two dry electrodes. A Fast Fourier Transform (FFT) with a Hanning window was applied to the first 6.25 seconds of each trial, per bipolar recording. The log-bandpower, commonly used in EEG-BCI analysis [9], was then averaged across 2 Hz frequency bands, from 7 Hz to 39 Hz. This 16-dimensional feature space (16 log-bandpower values per bipolar recording) was used offline to train two linear ν-Support Vector Machines (ν-SVM) [10], one for each bipolar recording, to discrimi-

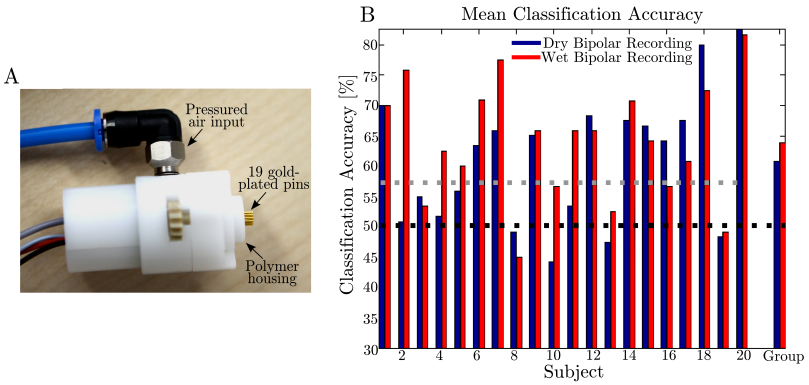


Figure 1: (A) Dry electrode prototype. (B) Mean classification accuracies of the dry and wet bipolar recordings. The dotted black line represents chance-level accuracy (50 %). The dotted gray line represents the classification accuracy required to reject the null-hypothesis of chance-level accuracy at $\alpha = 0.05$ (57.50 % for individual subjects and 51.67 % for the group, not shown).

nate between trials of the rest and imagery conditions. To avoid over-fitting, ν was selected using 10-fold cross-validation for each of the 10 outer folds.

3 Results

The performances of the active-dry and passive-wet electrodes are presented in Figure 1 (B). The group mean classification accuracies of the dry and wet bipolar recordings are 60.83 % and 63.88 %, respectively. These relatively low means are not surprising given that only one bipolar recording is used for classification in each case [11]. Based on a permutation test at significance level $\alpha = 0.05$, permuting the bipolar recording label, i.e., wet or dry, of the subject-specific classification results 10,000 times, the mean of classification differences is not large enough to reject the null-hypothesis of equal classification accuracy ($p = 0.0610$).

4 Discussion

Figure 1 (B) demonstrates that neither dry nor wet classification accuracies are consistently higher across all subjects. However, there are outliers in the single-subject results that best explain reasons for lower classification accuracies, on average, with the dry bipolar recording. In S2's case, wet classification accuracy is higher than that of the dry by 25 %. After examining S2's raw recordings, it was apparent that the second run was dominated by 50 Hz mains noise. Retraining the ν-SVM without the second run, the classification accuracies were identical for both dry and wet at 70 %. It is likely that the dry electrode pins lost contact with the subject's scalp between runs. Both dry and wet classification accuracies are low for S13 (47.5 % and 52.5 %, respectively) and S19 (48.33 % and 49.17 %, respectively). S13's raw data shows large artifacts at the same points in time for both the dry and wet recordings. It is likely that the dry electrodes, which protrude 3.5 cm above the EEG cap, were moving and subsequently shifting the wet electrodes. In S19's raw data, a heart-beat is visible in one dry electrode recording. Such low-frequency, periodic oscillations can be removed by high or band pass filtering. However, the skin-stretch artifacts caused by blood vessels and muscles underneath the scalp may not always display low-frequency periodicity, in which case filtering may be ineffective and classification accuracies are consequently influenced.

It should be noted that some subjects in this study found the electrode pins unpleasant. This brings attention to a trade-off in direct-contact dry electrodes - excessive force may cause the subject discomfort but insufficient force can lead to noisy signals or prevent electrode pins from penetrating thick hair [12].

5 Conclusion

This 20-subject study provides evidence that, for a motor-imagery task during which active-dry and passive-wet electrodes are used simultaneously, the two types of electrodes produce comparable classification accuracies. However, motion, whether that of the dry electrode, cap, or subject, seems to adversely affect classification accuracies. To reduce motion-related artifacts, further research into positioning the dry electrodes on the head (and maintaining their contact with the scalp) is particularly important. In addition, increasing the number of dry electrodes would allow for source separation techniques such as Independent Components Analysis, making it easier to examine physiological artifacts and design improved dry electrodes accordingly. Future designs should also consider the discomfort reported by some subjects by redesigning the pins, adjusting the forces applied to the scalp, or considering pin-free alternatives. The ultimate hope is to make new applications of EEG, such as non-invasive BCI, more viable for day-to-day use.

References

[1] C. Falco, F. Sebastiano, L. Cacciola, F. Orabona, R. Ponticelli, P. Stripe, and G. Di Gennaro. Scalp electrode placement by EC2 adhesive paste in long-term video-EEG monitoring. *Clinical Neurophysiology*, 116(8):1771–1773, 2005.

[2] H. Iguchi, K. Watanabe, A. Kozato, and N. Ishii. Wearable electroencephalograph system with preamplified electrodes. *Medical and Biological Eng. and Comp.*, 32(4):459–461, 1994.

[3] B. A. Taheri, R. T. Knight, and R. L. Smith. A dry electrode for EEG recording. *Electroencephalography and Clinical Neurophysiology*, 90(5):376–383, 1994.

[4] C. Fonseca, J. P. Silva Cunha, R. E. Martins, V. Ferreira, M. A. Barbosa J. P. Marques de Sá, and A. Martins Silva. A novel dry active electrode for EEG recording. *IEEE Transactions on Biomedical Engineering*, 54(1):162–165, 2007.

[5] T. O. Zander, M. Lehne, K. Ihme, S. Jatzev, J. Correia, C. Kothe, B. Picht, and F. Nijboer. A dry EEG-system for scientific research and brain-computer interfaces. *Frontiers in Neuroscience*, 5:53, 2011.

[6] G. Gargiulo, R. A. Calvo, P. Bifulco, M. Cesarelli, C. Jin, A. Mohamed, and A. van Schaik. A new EEG recording system for passive dry electrodes. *Clinical Neurophysiology*, 121(5):686–693, 2010.

[7] C. Grozea, C. D. Voinescu, and S. Fazli. Bristle-sensors—low-cost flexible passive dry EEG electrodes for neurofeedback and BCI applications. *Journal of Neural Engineering*, 8(025008), 2011.

[8] S. Mason, A. Bashashati, M. Fatourechi, K. Navarro, and G. Birch. A comprehensive survey of brain interface technology designs. *Annals Biomedical Engineering*, 35(2):137–169, 2007.

[9] G. Pfurtscheller and C. Neuper. Motor imagery and direct brain-computer communication. *Proceedings of the IEEE*, 89(7):1123–1134, 2001.

[10] B. Schölkopf, A. Smola, R. Williamson, and P. Bartlett. New support vector algorithms. *Neural Computation*, 12(5):1207–1245, 2000.

[11] M. Kamrunnahar, N. S. Dias, and S. J. Schiff. Optimization of electrode channels in brain computer interfaces. In *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2009)*, pages 6477–6480, 2009.

[12] F. Popescu, S. Fazli, Y. Badower, B. Blankertz, and K. R. Müller. Single trial classification of motor imagination using 6 dry EEG electrodes. *PLoS ONE*, 2(7):e637, 2007.

Brain-Computer Interface Control with Dry EEG Electrodes

C. Guger[1], G. Krausz[1], G. Edlinger[1]

[1]g.tec medical engineering GmbH/Guger Technologies OG, Graz, Austria

guger@gtec.at

Abstract

Brain-computer interfaces (BCI) are mostly realized using the P300, motor imagery or steady-state visually evoked potentials (SSVEP) measured with the electroencephalogram (EEG) to control external devices. The EEG is measured non-invasively with electrodes mounted on the human scalp using conductive electrode gel for optimal impedance and data quality. But the usage of gel has also some disadvantages: long montage time, abrasion of the skin, need to clean the skin after the recording,... The gel based EEG acquisition limits also the frequent usage of BCI systems on a daily basis. Therefore a dry active electrode system was developed and compared to gel based active electrodes. Three subjects performed P300, motor imagery and SSVEP based BCI experiments. Evoked potentials, event-related desynchronization, power spectrum and accuracies were calculated for dry and gel based electrodes to compare them. The study showed that the new dry electrodes are able to pick up the corresponding frequency ranges of the EEG data for all three BCI approaches. The major advantages are the fast montage, no abrasion and no need of cleaning the skin. Nevertheless dry electrodes are more sensitive to noise and therefore a careful montage is necessary.

1 Introduction

For the realization of a brain-computer interface (BCI) system 4 different principles can be used: slow cortical potentials, oscillations in alpha and beta range, steady-state visually evoked potentials (SSVEP) and the P300 event-related potentials. Recently BCI technology has been utilized not only for supporting subjects with special needs but also to possibly enrich control options in robotics or gaming areas. However, some subjects report about discomfort when participating in EEG experiments or even rejected participation as hair washing after the experiments is necessary. In order to improve the subjects' acceptance of BCI technology many research groups are now working on the practical usability of dry electrodes to completely avoid the usage of electrode gel. Dry electrodes use either micro-needles to penetrate the first layer of the skin and to get in conduct with the conducting layers, use capacitive sensors or are penetrating the skin with mechanical springs that press the electrodes into the skin [1–3]. Early work focused on the usage of active and dry electrodes for the recording of electrocardiogram signals which is easier to do because of the larger signal to noise ratio and easier montage on the thorax.

Single trial classification of motor imagination using 6 dry electrodes was already shown by the Berlin BCI group [2] and resulted in about 30 % lower information transfer rate than with gel electrodes. Gargiulo constructed a dry electrode system with conductive rubber showing a high correlation between gel based and dry electrodes [4]. A stainless steel disk with 3 mm was used to prove the usefulness of it for spontaneous EEG and evoked potentials (EP) [3].

In this publication it is demonstrated that a recently developed dry electrode system can be used for motor imagery, SSVEP and P300 based BCI systems. Therefore the power spectrum, the time course of evoked potentials, event-related desynchronization/synchronization (ERD/ERS) values and BCI accuracy are compared for three different BCI setups.

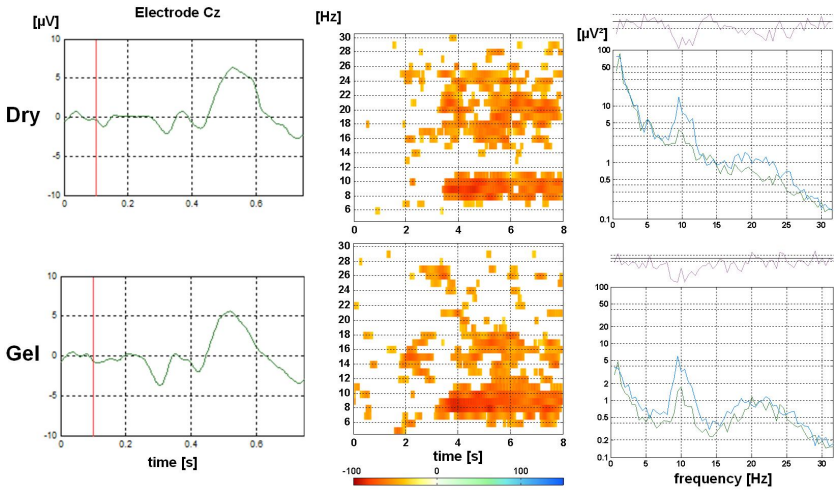


Figure 1: Left: P300 reponse for dry and gel electrodes in the copy spelling run of 1 subject. Each run had 5 characters and each character flashed 30 times (15 rowsm 15 columns). This gives in total about 5 minutes per run. The y-axis is scaled with ±10 µV, the x-axis in seconds. Middle: ERDmaps of electrode positions C3 during right hand movement imagination for dry (top) and gel electrodes (bottom). Both show a strong ERD in the alpha range from second 3.5 until 8 over C3. The dry electrodes show a broader beta ERD. Only pixels with significant ERD/ERS values are displayed (bootstrap, $p < 0.05$). Right: Reactive frequency components of the reference interval (0–2 s, blue) and active interval (6–8 s, green) of C3 of dry (top) and gel (bottom) electrodes. The graph above each power spectrum shows significant changes if the line crosses the dashed line (sign test, $p < 0.05$).

2 Methods

Three subjects performed the P300, motor imagery and SSVEP experiments with the electrode montage Fz, Cz, P3, Pz, P4, PO7, Oz and PO8 for P300, FC3, CP3, FC4 and CP4 for motor imagery and PO3, PO, PO4, PO7, O1, Oz, O2 and PO8 for SSVEP. For P300 and SSVEP experiments subjects performed one run with dry electrodes (g.SAHARA) and one run with gel electrodes (g.BUTTERFLY). In the case of the motor imagery BCI the dry and gel electrodes were mounted close together with 1.5 cm distance. The EEG was amplified with a 24 Bit high resolution biosignal amplifier (g.USBamp). Subjects were seated about 1 m in front of the computer monitor and were instructed about the experimental procedure.

P300 experiments were performed with intendiX. The intendiX speller shows 50 characters (A, B, ... Z; 0, 1, ... 9; and special characters) on the computer screen and highlights a whole column or row for 100 ms. Between the flashes there is a short time while only the grey matrix items are visible (60 ms). The BCI system must be calibrated in a first step on individual EEG data. Therefore the subject was asked to "select" (or attend to) the word WATER, one letter at a time. This training procedure took about 5 minutes. After training the BCI system using the calibration data, the subject was asked to write the word LUCAS, one character at a time, taking about 5 more minutes. The system uses a linear discriminant analysis for classification [5]. Figure 1 shows the evoked potential calculated from the EEG data from the copy spelling run for dry and gel electrodes. The P300 response is very similar in amplitude and latency.

For the motor imagery experiment gel based and dry electrodes were mounted beside each other to record EEG data (40 trials of left and 40 trials of right hand movement imagination) almost from the same region (1.5 cm apart). The motor imagery experiment started with the display of a fixation cross in the center of a screen. After 2 s, a warning stimulus was given in the form of a "beep". After 3 s, an arrow (cue stimulus) pointing to the left or right was shown for 1.25 s. The subject was instructed to imagine a right-hand movement or left-hand movement until

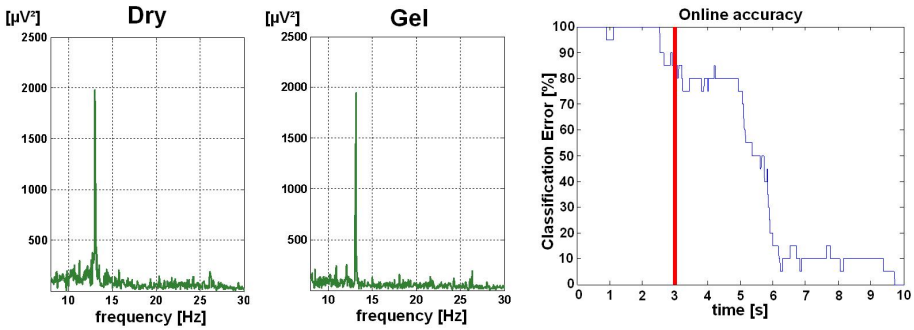


Figure 2: Power spectrum of EEG data of electrode Oz during 13 Hz LED stimulation. Right: Error rate of the SSVEP based BCI system with dry electrodes and 4 classes.

the end of the trial, depending on the direction of the arrow. One trial lasted 8 s and the time between two trials was randomized in a range of 0.5–2.5 s to avoid adaptation.

The motor imagery BCI estimated the bandpower in two different frequency bands of the EEG data. The bandpower features were classified with a linear discriminant analysis resulting in a subject specific weight vector [6]. The reactive frequency bands in the alpha and beta range were identified from the power spectrum and a time-frequency evaluation of the ERD/ERS activity (ERDmaps) as shown in Figure 1. First the EEG data was visually inspected and about 5 % of the trials containing artifacts were removed. In both cases an ERD in the alpha and beta ranges can be found. EEG measured with dry electrode recordings shows a broader activity in the beta frequency range. The power spectrum allows to identify the reactive frequency components in the EEG data. In the baseline period (without movement imagination) two alpha peaks can be found for this subject in both derivations (dry and gel). It is known from previous experiments from this subject that the higher alpha activity is more suppressed during the hand movement imagination and therefore this frequency band is used for the BCI control. The significant difference between baseline and imagination is proven by the sign test. EEG power spectra for the dry electrode show a higher difference in the beta region than for the gel based electrodes. However both measurements were done at nearby but still distinct locations. A clear difference comparing the two power spectra is the higher power found below 3 Hz for the dry electrode signal. However comparing power levels in alpha and beta ranges it can be stated that the ERDmaps and power spectra show very similar results for both types of electrodes.

One subject performed the SSVEP experiment with a training run to setup a subject specific weight vector and a testing run. This was done for dry and also for gel electrodes. The task of the subject was to attend for 7 seconds to one of 4 LEDS flickering with a certain frequency (10, 11, 12, 13 Hz) and then to rest for 3 seconds. The task was repeated for the remaining three LEDS and the whole loop was repeated 4 times (16 trials). The 4 LEDS were arranged in a 12×12 cm box and were controlled by a microcontroller resulting in a frequency error < 0.025 Hz.

The SSVEP based BCI system is controlled with discrete frequency peaks showing if the subject is looking at a certain LED. Figure 2 displays the power spectrum of an important electrode positions Oz for dry and gel based electrodes when the subject is looking at the 13 Hz LED computed from the complete four 7 seconds segments. The peak at 13 Hz is very similar and also the 1^{st} harmonic components at 26 Hz can be seen. The real-time classification was done with a linear discriminant analysis of minimum energy parameters [7].

3 Results

The P300 subject reached 100 % accuracy with gel and dry electrodes when LUCAS was spelled with 5 minutes of training data only. The motor imagery BCI accuracy was compared using a 10 times 10 fold cross validation technique that mixes the data randomly to have separate training

and testing data. The error with gel based and dry electrodes mounted beside each other is 18 % versus 15 %. Dry electrodes performed in this case slightly better and had an earlier best classification time point (7.5 s versus 8 s). For the SSVEP BCI the accuracy for dry electrodes is shown in Figure 2, right. The red line indicates the time point when the subject started to look at one of the 4 LEDs. In the reference interval the error is around 100 % and in the action interval the error drops down to finally 0 % at the end of the trial. A similar behavior was observed for the gel based electrodes reaching also an error rate of 0 %.

4 Discussion

We could show that the used dry electrode sensor concept can be used for motor imagery, SSVEP and P300 based BCI systems. For dry electrodes no conductive gel is used and therefore a much higher skin-electrode impedance than for gel based electrodes can be expected. Electrodes with higher impedance can pick up more artifacts and are mostly sensitive for movements of the electrodes and cable swings which results in signal amplitudes much higher than for normal EEG. Electrodes with high impedance can also pick up electrostatic voltages in the surrounding and electro-magnetic noise. To solve these problems we reduced the impedance with multiple gold coated pins per electrode being in contact with the skin. Secondly we integrated an amplifier unit into the electrode itself to make it resistant against artifacts and to be able to record EEG with a high electrode impedance. Dry electrodes also show a higher polarization voltage than gel based electrodes and therefore the recording equipment must be able to accept DC voltages up to several mV. This was solved with an amplification unit with high input range in combination with a 24 Bit ADC.

It was shown that dry and gel electrodes reach similar accuracies and are able to pick up similar physiological responses for P300, SSVEP and motor imagery BCI experiments. Nevertheless group studies are required to show the usefulness in real life situations as needed for home applications.

References

[1] M. Matteucci, R. Carabalona, M. Casella, E. di Fabrizio, F. Gramatica, M. di Rienzo, E. Snidero, L. Gavioli, and M. Sancrotti. Micropatterned dry electrodes for brain-computer interface. *J. Microelectronic Engineering*, 84(5-8):1737–1740, 2007.

[2] F. Propescu, S. Fazli, Y. Badower, B. Blankertz, and K. R. Müller. Single trial classification of motor imagination using 6 dry EEG electrodes. *PLoS One*, 7(e637):1–5, 2007.

[3] B. A. Taheri, R. T. Knight, and R. L. Smith. A dry electrode for EEG recording. *EEG Clinical Neurophysiology*, 90:376–383, 1994.

[4] G. Gargiulo, P. Bifulco, R. A. Calvo, M. Cesarelli, C. Jin, and A. van Schaik. A mobile EEG system with dry electrodes. *IEEE Biomedical Circuits and Systems Conference*, pages 273–276, 2008.

[5] C. Guger, S. Daban, E. Sellers, C. Holzner, R. Carabalona, F. Gramatica, and G. Edlinger. How many people are able to control a P300-based brain-computer interface (BCI)? *Neuroscience Letters*, 462(1):94–98, 2009.

[6] C. Guger, G. Edlinger, W. Harkam, I. Niedermayer, and G. Pfurtscheller. How many people are able to operate an EEG-based brain-computer interface. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11:145–147, 2003.

[7] O. Friman, I. Volosyak, and A. Gräser. Multiple channel detection of steady-state visual evoked potentials for brain-computer interfaces. *IEEE Transactions on Biomedical Engineering*, 54:742–750, 2007.

Low Cost Brain-Computer Interface First Results

A. J. Portelli[1], I. Daly[1], M. Spencer[1], S. J. Nasuto[1]

[1]School of Systems Engineering, Brain Embodiment Lab,
University of Reading, Reading, England

http://www.reading.ac.uk/cirg/

Abstract

Brain-Computer Interfacing (BCI) has been previously demonstrated to restore patient communication, meeting with varying degrees of success. Due to the nature of the equipment traditionally used in BCI experimentation (the electroencephalograph) it is mostly confined to clinical and research environments. The required medical safety standards, subsequent cost of equipment and its application/training times are all issues that need to be resolved if BCIs are to be taken out of the lab/clinic and delivered to the home market. The results in this paper demonstrate a system developed with a low cost medical grade EEG amplifier unit in conjunction with the open source BCI2000 software suite thus constructing the cheapest per electrode system available, meeting rigorous clinical safety standards. Discussion of the future of this technology and future work concerning this platform are also introduced.

1 Introduction

Interfacing between brain and machine holds great promise in restoring communication channels to patients of long term debilitating conditions such as amyotropic lateral sclerosis (ALS) or motor neuron disease, but is limited by many practical considerations. In the case of EEG preparation time is lengthy and is confounded by the necessity of applying a layer of conductive gel to each electrode prior to application. Additionally, due to inter-subject and inter-trial variability, reliably and significantly identifying goal oriented behaviour from background EEG activity generally requires a lengthy training process. Nevertheless, EEG has been the *de facto* standard for bio-potential monitoring for more than ninety years. In this time, important advances have been made in the identification of specific brain activity and in turn this knowledge has been incorporated into a number of BCI paradigms. Some of these include the P300 speller [1], alpha band cursor control [2] and beta band motor imagery experiments [3].

Although alternative imaging technologies have been utilised for BCI, e.g. Near Infrared Spectroscopy (NIRS) [4] or functional Magnetic Resonance Imaging (fMRI) [5], their costs are prohibitively high. Moreover, EEG exhibits excellent temporal resolution and is comparitively easy to ease.

Manufacturer	Name	No. of Channels	≈ cost per channel
DeyMed Diagnostic	TruScan32	32	$ 150.50
Bio-Semi	ActiveTwo Analog Input Box	32	$ 200.00
Cadwell	Easy II EEG PSG Machine	32	$ 280.00
Brain Master	2EB Clinical System	2	$ 650.00

Table 1: Costs associated with various EEG systems (given per channel).

Despite its significantly lower cost, in comparison to other imaging modalities, typical EEG systems are still too expensive to be widely used outside of clinical and laboratory settings. EEG

units can vary in cost based on the number of channels and the supplier of the equipment, ranging from a few hundred dollars for two channel systems up to hundreds of thousands of dollars for 128/256 channel systems.

In order to construct a digital EEG recording unit specific elements are required, i.e. the electrodes, the analog to digital converters, the amplifiers, optical isolation and the transmission into a computer for processing. The obvious and perhaps simplest solution would be to construct a bespoke EEG amplification system such as [6]. However, this would introduce several problems including safety issues as well as prototyping and testing costs/time. These concerns make the task of an independent construction less desirable as a solution.

After exploring a number of EEG amplifiers available on the market, we have identified the DeyMed Diagnostic TruScan32 system as the cheapest (per channel) commercially available system which meets European medical safety standards and retails at approximately $ 5000 for 32 channels (approx. $ 150 per channel). Other commonly used systems are shown in Table 1 listed with approximate cost per electrode channel.

Cost per channel is an appropriate metric for the assessment of EEG system costs due to the lack of a priori knowledge often associated with the channel selection in BCI paradigms. During development of a new BCI paradigm it is favourable to have access to as many channels and thus as much EEG data as possible. It is not until after validation that it becomes appropriate to reduce the number of recording channels and subsequently preparation/computation time.

Figure 1: TruScan32 EEG Amplifier Box.

Figure 2: Demonstration of experimental on-line BCI setup.

In this paper the use of a low cost, medical quality EEG amplifier is utilised as an alternative to other well known systems. The DeyMed Truscan32, has been documented as a possible solution to this problem previously used in an off-line configuration [7]. In this work the system is coupled with the open source BCI2000 software suite and other bespoke software to obtain closed loop, online BCI recordings. The primary distinction between this and other low cost proposed BCI solutions is the low cost of the unit combined with its medical grade (under CE 93/42 UE Council directive 93/42) and expandable number of electrodes (up to 128).

2 Apparatus/Equipment

The TruScan32 is comprised of two stages, the headbox and adaptor. The headbox is battery operated and responsible for the primary acquisition, amplification and analog to digital conversion. The digital data is then passed via optical cable providing total isolation from mains onto the adaptor stage which connects to the computer using a 25-pin parallel port (later versions of the unit use USB). The design of the hardware allows it to acquire signals at sampling rates ranging from 128 Hz to 1024 Hz providing (at 256 Hz and above) full spectrum EEG signals. The expand-

able interface allows anywhere from 1 to 128 (not including reference and ground) electrodes to be attached to the headstage and the referencing montage then configured in software.

The software provided by Deymed is intended primarily for clinical applications and is mostly unsuitable for EEG experimentation and online processing as required for BCI applications. However, the software is capable of providing a UDP packet stream. Software was developed for integration into the BCI2000 framework to connect to this stream and interpret data in the UDP packets for online processing in a BCI application. As far as the authors are aware this is the first application of the Truscan32 amplifier system to be used in an online BCI. For the purposes of signal processing a MATLAB script was written to utilise specific frequency components of the subject's EEG to allow one dimensional cursor control.

3 Experimental Procedure

The objective of this proof of concept closed loop BCI experiment was for the subject to control the movement of a computer simulated ball which was moving across the screen from the left to right at a constant speed. The subject was tasked to alter the vertical position of the ball either up or down by a fixed increment by respectively engaging into mental operations increasing cognitive load (e.g mental arithmetic) or relaxing. The trial was deemed successful if the subject managed to manipulate his/her cognitive load (and thus their alpha power) so that the ball could hit a target, a vertical bar of length equal to half the screen height displayed on the far right side of the screen either at the upper or lower half, selected at random. The position of the target was fixed for the duration of the trial and changed randomly with equal probability between trials. The experimental setup is illustrated in Figure 2. Seven electrodes were selected from the international 10-20 (not including ground and reference); Pz, P3, P4, T5, T6, O2 and O1. These were used to record EEG signals from each subject in a common reference montage at a sampling frequency of 256 Hz. The alpha power for each electrode was calculated individually using MATLAB and then averaged to obtain a single value. The alpha band power was defined as the power of the EEG data filtered between 8–13 Hz. This value was calculated once per second. The mean of the range of alpha power values was determined after a training session from background EEG activity. Subsequently when the alpha power was higher (lower) than this calculated mid-point the ball would be displaced downward (upward).

The experiments to achieve the results outlined in the next section were conducted with student volunteers at the University of Reading (1 female and 4 male) between the ages 23 and 49.

4 Results

The results are presented in Table 2 where the window size refers to the amount of time in seconds that the ball took to travel across the screen in each trial, each session being ten trials long. The average classification rate is shown for each subject, with four out of five subjects achieving significant classification rates ($p < 0.01$).

Subject	Session Numbers	Window Size(s)	Mean Classification	Significance
1	5	10	78 %	$p < 0.01$
2	5	10	90 %	$p < 0.01$
3	5	10	72 %	$p < 0.01$
4	5	10	60 %	$p > 0.05$
5	5	10	76 %	$p < 0.01$

Table 2: Results obtained from BCI experimentation.

These results were achieved with a maximum training time of fifteen minutes. Experimental conditions(environment, lighting, setup and number of electrodes) were kept as similar as possible for each subject.

5 Discussion

The results shown are only preliminary but demonstrate that cursor control with this low cost system is possible to a statistically significant degree ($\alpha = 0.01$) of accuracy with minimal subject training. Only one subject out of the five tested subjects was unable to reach a significant classification rate. This is in keeping with the approximate twenty percent of subjects who do not meet the required classification rate for control of BCI based systems, the so-called 'BCI illiteracy problem' [8].

6 Conclusion

BCI experimentation has already proven itself to be a powerful tool, enabling the provision of extra communication channels for users with debilitating or chronic conditions. At present however, its use is commonly limited to laboratory or clinical environments. The design of systems that utilise low cost EEG solutions will provide the advance necessary for this technology to be removed from the lab and made accessible. With an increase of accuracy, information transfer and improvements of human computer interactions in BCIs, the costs of the proposed platform may be sufficiently low for health care systems to offer a viable solution to the severely disabled. Further development of these systems can significantly increase the independence of the patients and hence reduce the workload of carers. The results of the experiments demonstrate that a low cost brain-computer interface platform meeting medical grade safety standards is viable, practical and expandable.

Future work will explore further experiments with the platform in order to assess its versatility and applicability to other BCI control paradigms. We are currently working on validation of this platform for SSVEP, P300, and multimodal paradigms. It is hoped that such development will contribute to increase the momentum of BCI related research whereby more labs will be able to conduct BCI experiments.

References

[1] J. D. Bayliss and D. H. Ballard. A flexible brain-computer interface. *New York*, 2001.

[2] G. R. McMillan. The technology and applications of biopotential-based control. *RTO Lecture series on "Alternative control technologies: Human Factor Issues"*, (October):1–11, 1998.

[3] F. Popescu, S. Fazli, Y. Badower, B. Blankertz, and K. R. Müller. Single trial classification of motor imagination using 6 dry EEG electrodes. *PLOS ONE*, 7:1–5, 2007.

[4] R. Sitaram, H. Zhang, C. Guan, M. Thulasidas, Y. Hoshi, A. Ishikawa, K. Shimizu, and N. Birbaumer. Temporal classification of multichannel near-infrared spectroscopy signals of motor imagery for developing a brain-computer interface. *NeuroImage*, 34(4):1416–27, February 2007.

[5] N. Weiskopf, K. Mathiak, S. W. Bock, F. Scharnowski, R. Veit, W. Grodd, R. Goebel, and N. Birbaumer. Principles of a brain-computer interface (BCI) based on real-time functional magnetic resonance imaging (fMRI). *IEEE transactions on bio-medical engineering*, 51(6):966–70, June 2004.

[6] Aguazul. The open EEG project, http://openeeg.sourceforge.net/doc/, 2005.

[7] A. J. Portelli and S. J. Nasuto. Toward construction of an inexpensive brain computer interface for goal oriented applications. In *AISB 2008 Convention Communication , Interaction and Social Intelligence*, pages 7–12, Aberdeen, 2008.

[8] B. Z. Allison, C. Brunner, V. Kaiser, G. R. Müller-Putz, C. Neuper, and G. Pfurtscheller. Toward a hybrid brain–computer interface based on imagined movement and visual attention. *Journal of Neural Engineering*, 7(2):026007, April 2010.

Brain-Computer Interface and ERP Recordings: a Close Look on Trigger Signal

S. Silvoni[1], J. Mellinger[2]

[1]Dep. of Neurophysiology, S.Camillo Hospital Foundation I.R.R.C.S., Venice, Italy
[2]Institute of Medical Psychology and Behavioral Neurobiology - Eberhard-Karls-University, Tübingen, Germany

stefano.silvoni@ospedalesancamillo.net

Abstract

Timing precision is one of the relevant issues concerning the recording of event-related potentials (ERPs). ERP-based studies, including P300-based brain-computer interfaces (BCIs), have been conducted to obtain clinical information about cognitive processing. To ensure clinical interpretability of recorded ERP traces, timing accuracy of stimulus onset within BCI frameworks needs to be described. We focused attention on visual stimuli and two BCI frameworks: BCI++ and BCI2000. Experimental tests that prove the existence of a delay between stimulus onset triggering and actual stimulus appearance are reported. A close examination of factors that might influence the observed delays could lead to software and hardware solutions that improve the reported BCI frameworks' behaviour.

1 Introduction

ERPs may be recorded and analyzed for many applications. Possible uses of ERPs range from cognitive and psycho-physiological evaluation [1, 2], prognosis in vegetative states [3], to BCI-based communication [4]. Offline and online processing of ERP traces have been used to estimate parameters such as latencies and features to be decoded for peak recognition. These measures rely on two principles: i) a random stimulus sequence presentation, where only one stimulus represents the target for a specific task [1], and ii) a good timing accuracy of the stimulus onset in order to correctly detect relative peak latencies [5–7]. The first principle has been used to develop many types of ERP-based BCI paradigms; some examples are: visual matrix-speller [4], symbolic communication [8], and auditory presentation. The possibility to extract clinical, cognitive and neuro-psychological information from ERP traces, such as attentive and executive function related indices [1, 2], suggests that in this type of recordings a good timing accuracy should be ensured. Since most studies on P300-based BCIs do not report detailed information about the timing confidence of the stimulus onset in EEG recordings, the present contribution investigates the delay between stimulus onset in a trigger channel, and actual appearance of a stimulus, focusing on visual stimuli and two BCI frameworks: BCI++ [9] and BCI2000 [10].

During recording, BCI2000 acquires blocks of data, and processes each, introducing a data processing delay. After processing, it generates visual output by instantly refreshing the content of video memory. As soon as video memory has been updated, a stimulus time stamp is recorded. In addition, a trigger called *StimulusCode* is set, which is aligned with the beginning of the previously acquired data block. Thus, visual presentation is always delayed with regard to the trigger by a data and video processing latency (which is reflected in the stimulus time stamp), and an additional video-memory-to-display latency (which is not reflected in the time stamp). In the present paper, we do not consider the stimulus time stamp for analysis, and concentrate on the delay existing between trigger, and stimulus appearance.

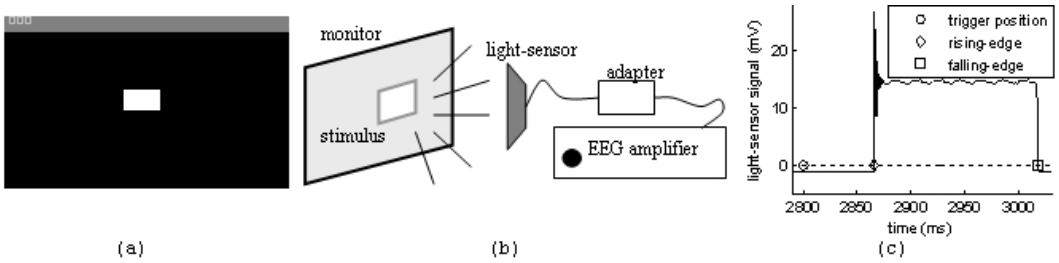


Figure 1: (a) P3 speller with a white squared picture, (b) LS coupling, and (c) LS signal.

BCI++ operates similarly to BCI2000 with the difference that the stimulus presentation is carried out by an application (AEnima), while a second application (HIM) provides data acquisition and processing. Because of these two applications are connected via a TCP-IP protocol, an additional delay may be present between the updating of the *AEnima-trigger* state variable (i.e. stimulus onset, provided by HIM application) and the actual time-point of visual output.

2 Methods

2.1 Experimental Setup

To measure the delay between visual stimulus onset triggering and actual stimulus appearance on the screen, one needs to detect the time-instant of the pixel colour/intensity change of the associated screen area. To this purpose we used a customized light-intensity sensor (LS, with a transient response lower than 100 µs) whose circuitry output was opportunely linked to an input of the EEG amplifier. This device was utilized with both BCI++ and BCI2000 frameworks in a visual paradigm. BCI++ provides a customizable graphic user interface (AEnima) to implement a visual ERP paradigm, while BCI2000 allows to quickly set up a P3 speller paradigm. In both frameworks an automated software mechanism allows stimuli presentation onset tracking on EEG recordings (*AEnima-trigger* and *StimulusCode* states).

In our modified version of the visual paradigm a white squared picture flashed over a black background, while the LS was located over this picture with a suction cup to capture pixel intensity changes (see Figure 1), recording them in one of the EEG channels. Actual stimulus onset and termination corresponded to the LS signal's rising and falling edges, as detected by a threshold condition on the signal's first derivative (±1 mV/ms); this was compared to the position of the trigger in the recorded data file (see Figure 1 (c)).

2.2 BCI++

This framework [9] was used to carry out a series of 4 runs using a four-choice paradigm [8], where a white squared picture flashed every 2 s for 200 ms in the top-right corner of the screen. The acquisition system consisted of two PCs connected with a 10/100 MB dedicated local network and one EEG amplifier; its specifications are as follows: desktop PC for HIM software (v7.0): HP Pentium Dual CPU E2220 2.4 GHz, 2 GB RAM, Windows XP 2002 SP3; desktop PC for AEnima software (v8.7): Intel Pentium 4 CPU 3 GHz, 512 MB RAM, Windows 2000 SP4, LCD HPL1906 1024 × 768 60 Hz, video-card SiS 661FZ (this PC was used to run the graphic interface); EEG Amplifier: SynAmps (NeuroSoft, Inc.), 6 channels (one for LS output), 2500 Hz sampling rate, low-pass filter at 500 Hz (a block of 100 samples for each channel was transferred at a time via a TCP-IP protocol). The stimuli dataset consisted of 234 picture flashes.

Framework/PC/Amp.	N.of stimuli	Sampling rate (Hz)	Stimulus delay (ms) mean (std)	Flash duration (ms) mean (std)
BCI++/DT/SA	234	2500	132.8 (26.8)	208.4 (12.0)
BCI2000/DT/SA	240	2500	86.7 (10.4)	191.7 (14.6)
BCI2000/DT/GA	240	2400	55.0 (8.3)	163.8 (8.3)
BCI2000/LT/GA	240	2400	36.3 (5.4)	186.3 (6.3)

Table 1: Stimulus delay and duration results. DT: desktop PC for BCI++ and BCI2000 tests. LT: laptop PC for BCI2000 test. SA: SynAmps amplifier. GA: g.USBamp amplifier.

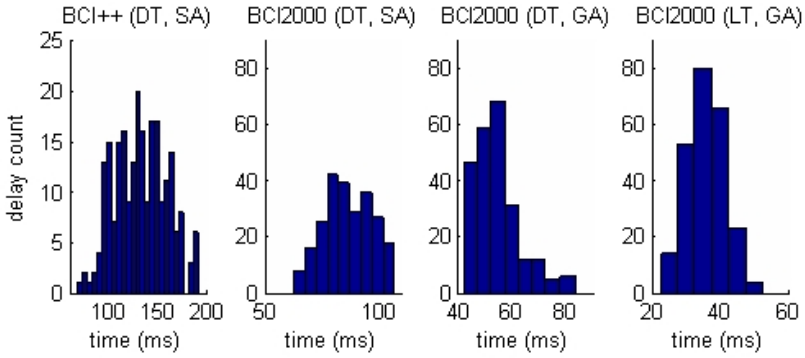


Figure 2: Delay distributions.

2.3 BCI2000

The P3 speller paradigm was used in this framework (downloaded at May 27, 2010, [10]) to carry out a series of 8 runs, where the white squared picture was located in one element of a 6×6 matrix; other elements were set to blank. Stimulus duration, as well as the inter-stimulus interval, were set to 200 ms. The acquisition system consisted of one PC and one EEG amplifier. Three tests were carried out: In the first two tests we used the same PC as with BCI++, while in the third test a laptop PC was used. The combinations are as follows : i) using the desktop PC and the SynAmps amplifier, ii) using the desktop PC and the g.USBamp amplifier, and iii) using a laptop PC and the g.USBamp amplifier. The characteristics of the laptop PC and the amplifiers are as follows: laptop PC for BCI2000: HP CPU Intel Core2 Duo P8600 2.4G Hz, 2 GB RAM, Windows XP 2002 SP3; standard integrated LCD 1400 × 1050 60 Hz, video-card ATI Mobility FireGL V5700; first EEG Amplifier: SynAmps (with the same setting used for BCI++ test); second EEG amplifier: g.USBamp (g.tec GmbH), 16 channels (one for LS output), 2400 Hz sampling rate, low-pass filter at 500 Hz, SampleBlockSize 32. For each PC the stimuli dataset comprised 240 picture flashes.

3 Results and Discussion

As shown in Table 1, and Figure 2, a variable delay between stimulus onset triggering and actual stimulus appearance exists. Depending on framework and PC characteristics, it may be of moderate size (BCI2000 laptop PC). Furthermore, flash duration was slightly corrupted by a small jittering or a small pre-shift. These results suggest that peak latencies and graphs reported in many ERP-based studies may be affected by this phenomenon. Stimulus presentation behaviour may be explained by some factors, such as the time needed to transfer graphical information from the video card into the LCD, and the time required for the CPU to serve graphics commands (according to the processes' priority). For the BCI++ framework, a further delay-factor may be attributed to the transfer of the *AEnima-trigger* value via a TCP-IP-based protocol [9]; this is due to the uncertainty of the *AEnima-trigger* synchronization within a single acquired block of 100 samples corresponding to 40 ms. The results for BCI2000 agree with the values for stimulus

timing jitter for LCD displays found in [7], but significantly differ in overall stimulus delay which is larger for the present study.

Adopting specific hardware and/or software interventions, such as specific low-level drivers [5,6], or a hardware trigger input, timing accuracy could be improved. For BCI2000, using the recorded stimulus time stamp in data analysis or online classification would considerably improve accuracy. An enhanced stimulus presentation accuracy could influence two aspects of BCI systems: i) a greater clinical interpretability of the recorded ERP traces combined with the paradigm specification, by controlling the time delay between trigger and stimulus; and ii) the possibility to increase BCI classifier performance due to a higher stability of the ERP recordings, by reducing timing jitter, especially with BCI++ framework or when it becomes relevant due to more sensitive classifier technology.

4 Conclusion

Both investigated BCI frameworks present a variable delay between visual stimulus onset triggering and actual stimulus appearance. We suggest a close examination of system software and hardware characteristics to improve the reported BCI frameworks' behaviour.

Acknowledgments. We are grateful to the contribution of L.Criveller and C.Genna (University of Padova, Italy) in developing and testing the LS, and to the BCI++ support of L.Maggi, P.Perego, S.Parini (Politecnico di Milano University, Italy).

References

[1] J. Polich. Updating P300: an integrative theory of P3a and P3b. *Clin Neurophysiol.*, 118:2128–2148, 2007.

[2] B. Kotchoubey, S. Lang, S. Winter, and N. Birbaumer. Cognitive processing in completely paralyzed patients with amyotrophic lateral sclerosis. *Eur J Neurol.*, 10:551–558, 2003.

[3] M. Cavinato, U. Freo, C. Ori, M. Zorzi, P. Tonin, F. Piccione, and A. Merico. Post-acute P300 predicts recovery of consciousness from traumatic vegetative state. *Brain Inj.*, 23:973–980, 2009.

[4] L. A. Farwell and E. Donchin. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Clin Neurophysiol.*, 70:510–523, 1988.

[5] K. I. Forster and J. C. Forster. DMDX: a windows display program with millisecond accuracy. *Behav Res Methods Instrum Comput.*, 35:116–124, 2003.

[6] S. Xie, Y. Yang, Z. Yang, and J. He. Millisecond-accurate synchronization of visual stimulus displays for cognitive research. *Behav Res Methods.*, 37:373–378, 2005.

[7] J. A. Wilson, J. Mellinger, G. Schalk, and J. Williams. A procedure for measuring latencies in brain-computer interfaces. *IEEE Trans Biomed Eng.*, 57:1785–1797, 2010.

[8] S. Silvoni, C. Volpato, M. Cavinato, M. Marchetti, K. Priftis, A. Merico, P. Tonin, K. Koutsikos, F. Beverina, and F. Piccione. P300-based brain-computer interface communication: evaluation and follow-up in amyotrophic lateral sclerosis. *Front Neurosci.*, 3, 2009.

[9] L. Maggi, S. Parini, P. Perego, and G. Andreoni. BCI++: an object-oriented BCI prototyping framework. *Proceedings of 4th International Brain-Computer Interface Workshop and Training Course, Graz*, pages 379–384, 2008.

[10] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw. BCI2000: a general-purpose brain-computer interface (BCI) system. *IEEE Trans Biomed Eng.*, 51:1034–1043, 2004.

A Unified XML Based Description of the Contents of Brain-Computer Interfaces

V. Putz[1], C. Guger[1], C. Holzner[1], S. Torrellas[2], F. Miralles[2]

[1]g.tec medical engineering, Schiedlberg, Austria
[2]Barcelona Digital Centre Tecnologic, Barcelona, Spain

bci2011@tugraz.at

Abstract

In the past decades, a variety of applications and devices were interfaced with EEG based Brain-Computer Interfaces (BCIs) with the aim to offer assistive technology to severely handicapped users. Nevertheless, up to now no standardized description of the possible interaction options (i.e. the tasks a user can perform with the aid of the BCI) was available. Each function provided by an application or device connected to the BCI had to be hard–coded. In this contribution, we propose a new platform–independent XML based description of the interaction options. The scheme is interpreted by a middleware layer which connects applications and devices to the BCI. In its current version, scheme and middleware layer are designed for a BCI which provides graphical feedback to the user.

1 Introduction

An EEG-based Brain-Computer Interface (BCI) can be implemented using various strategies like slow cortical potentials, the P300 response, steady-state visual evoked potentials (SSVEP) or cortical oscillations (ERD/ERS) [1]. In the past decade, several examples using one of these methods to control applications including a Virtual Reality environment [2], a speller [3] and physical devices such as small robots [4] and smart home devices [5] were published. Although a very general description of the BCI and its control interfaces to devices was published by Mason and Birch [6], to the best of our knowledge, no approach to achieve a standardized coding description of the interaction options provided to the users has been published yet. The requirements for such an interface description are obvious: At first, it should be able to serve BCIs based upon different strategies like P300, SSVEP and ERD/ERS. Thus, each user could benefit from an optimal BCI strategy for controlling his/her individual applications. Furthermore, the connection between the BCI and a variety of external devices and applications should be possible without constantly reprogramming the BCI. Finally, the interface should enable engineers without BCI experience to connect new applications and devices to the BCI by simply following a provided coding scheme.

2 XML Description of the BCI Contents

In this section, a standardized XML scheme to describe the contents of a BCI and its interaction with connected devices and applications is introduced. Three levels of control options are proposed:

- **Level 1 (ControlGroup):** All applications and devices which are connected to the BCI with the same controller (i.e. which are related to one IP address) are combined to one ControlGroup. *Example: A smart home controller.*
- **Level 2 (Application)**: A ControlGroup as described above may manage one or more subsequent devices. All UI commands required to control a device or application are grouped as Application. *Examples: devices connected to a smart home controller (TV, Radio, ...).*

- **Level 3 (Actions)**: Actions define the functions related to each interaction option which shall be accessible to the user. Each action includes an instruction to be forwarded to the connected controller. *Examples: Switching the TV channels.*

With these levels, a BCI offering several control options grouped to a hierarchical menu structure can be described. All XML tags which are required to yield a graphical representation of Level 1 and Level 2 control structures are enlisted in Table 1. Please note that enclosed within Level 1, several Level 2 Applications can be listed.

ControlGroup	**Description**
⟨ControlGroup⟩	Name of the ControlGroup
⟨CGAddress⟩	Network address (IP address and port number) of the controller
⟨MaskSize⟩	Number of rows and columns required for a matrix shaped GUI
⟨DispSymbol⟩	Icon or alternative text information to be displayed within the GUI
⟨Application⟩	Listing of Level 2 (Application) commands
Application	**Description**
⟨AppName⟩	Name of the Application
⟨AppID⟩	Unique Application ID
⟨Commands⟩	Three different options for Level 3 Commands: ⟨SingleCommand⟩, ⟨ToggleCommand⟩, ⟨MaskCommand⟩

Table 1: XML tags required for describing Level 1 (ControlGroup) and Level 2 (Application).

The actual set of interaction options offered to the user is defined at Level 3 of the XML scheme: A Level 3 Action can be either a *SingleCommand*, a *ToggleCommand*, or a *MaskCommand*. A SingleCommand (see Table 2) causes an instruction to be sent to connected applications or devices (e.g. increasing the volume of a TV set). A ToggleCommand (see Table 3) memorizes a binary state of a device (e.g. light on/off). When it is selected, the current state is toggled, and an instruction is sent to the connected device. Finally, a MaskCommand (see Table 3) is used to build a hierarchy of masks with parent- and submenus. Coding examples for Level 1, Level 2 and Level 3 controls are provided in Figure 1.

SingleCommand	**Description**
⟨CommandName⟩	Name of the Action
⟨Instruction⟩	Instruction string (sent to the connected device/application)
⟨ICONPosition⟩	Position of the action button within a matrix interface, e.g. $[1, 1]$
⟨DispSymbol⟩	Data for an icon or an alternative text to be displayed
⟨CommType⟩	**single**: After the selected instruction is sent, the interface switches to the disabled mode. **multi**: The selected instruction is sent without disabling the interface afterwards. **writebuffer**: The selected instruction is appended to a buffer (ideal when spelling text). **usebuffer**: The contents of the buffer is sent to the controller. **delbuffer**: The last instruction is deleted from the buffer. **closemask**: Return from a submenu to its parent menu.

Table 2: XML tags required for a SingleCommand.

3 Handshake Procedure and Data Transfer

With the presented XML description, functions provided by applications and devices can be encoded and integrated with the BCI. To load and parse a list of XML tags and to control communication between the BCI and connected controllers, a new middleware layer is introduced (see Figure 2). Currently, BCI software and middleware layer run on the same PC, while devices

ToggleCommand	**Description**
⟨CommandName⟩	Name of the Action
⟨InstructionOn⟩	Instruction sent, when the state is toggled to “on”
⟨InstructionOff⟩	Instruction sent, when the state is toggled to “off”.
⟨ICONPosition⟩	see Table 2
⟨DispSymbolOn⟩	Icon to be presented when the state is “on”.
⟨DispSymbolOff⟩	Icon to be presented when the state is “off”.
⟨Init⟩	Inital state of the toggle command (either on or off).
⟨CommType⟩	see Table 2
MaskCommand	**Description**
⟨MaskName⟩	Name of the mask
⟨MaskSize⟩	Number of rows and columns of the submask
Level 3 Commands	Level 3 Commands of the submask

Table 3: XML tags required for a ToggleCommand and a MaskCommand.

Level 1: ControlGroup

```
<ControlGroup CGName="CG1">
    <CGAddress> ... </CGAddress>
    <MaskSize>
        <NoLines>3</NoLines>  <NoCols>3</NoCols>
    </MaskSize>
    <DispSymbol> ... </DispSymbol>
    <Application AppName="App1"> ...</Application>
</ControlGroup CGName="CG1">
```

Level2: Application

```
<Application AppName="App1">
    <AppID>APP_001</AppID>
    <ToggleCommand CommandName="T1">
...</ToggleCommand>
    <SingleCommand CommandName="S1"> ...</SingleCommand>
    <MaskCommand CommandName="M1">...</MaskCommand>
</Application>
```

Level 3: SingleCommand

```
<SingleCommand
  CommandName="S1">
  <Instruction> Vol_up</Instruction>
  <ICONPosition> ... </ICONPosition>
  <DispSymbol>
      <Text>Vol +</Text>
  </DispSymbol>
  <CommType>multi</CommType>
</SingleCommand>
```

Level3: ToggleCommand

```
<ToggleCommand CommandName="Power">
    <InstructionOn>On</InstructionOn>
    <InstructionOff>Off</InstructionOff>
    <ICONPosition>...</ICONPosition>
    <DispSymbolOn> ...</DispSymbolOn>
    <DispSymbolOff>...</DispSymbolOff>
    <Init>Off</Init>
    <CommType>multi</CommType>
</ToggleCommand>
```

Level3: MaskCommand

```
<MaskCommand  MaskName="Sub1">
    <ICONPosition>[3,1]</ICONPosition>
    <DispSymbol> ... </DispSymbol>
    <MaskSize> ...  </MaskSize>
    <SingleCommand
        CommandName="S2">
    </SingleCommand>
    ...
</MaskCommand>
```

Figure 1: Coding examples for the three levels of the proposed XML description.

and applications are controlled over the network. To achieve streaming communication without the danger of a faulty device blocking the whole interface, we decided to use UDP. However, to assure communication with connected devices at startup, we defined a handshake procedure (see Figure 2): At first, the middleware sends a broadcast message to initialize the handshake. Listening controllers can answer the call by sending an acknowledge message. Next, the middleware asks for a list of XML tags encoding commands for the BCI. Thus, a controller can initialize the state of the BCI depending upon the current state of connected devices or applications (e.g. to correctly forward the state of a light switch to the middleware). This command list must match the XML scheme described in Section 2. The middleware allows more than just one controller to respond to the broadcast call and to send a list of commands. All received XML command lists are merged by the middleware and parsed. The parsed list is forwarded to the BCI software, and the paradigm GUI is initialized. In our current implementation, the BCI software enables P300 stimulation and signal processing. Thus, the GUI fulfils two purposes: Firstly, it informs the BCI user about his/her current position within the hierarchic structure of the interface. Secondly, it serves as flashing stimulation paradigm.

After a successful handshake, the BCI is ready to use. Throughout usage, the BCI software feeds back information about items selected by the user to the middleware layer. The middleware then forwards the contents of the ⟨Instruction⟩ tag (compare to Figure 1) of this item to the connected controller.

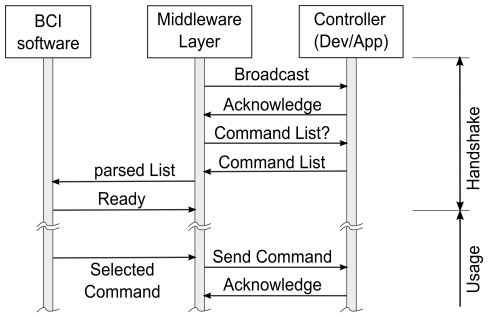


Figure 2: Communcation between the BCI software, the Middleware Layer and a connected network controller, which forwards instructions to devices or applications.

4 Conclusion

With the provided XML structure and handshake procedure, it is possible to connect a BCI to a set of devices and applications. In its current version, the XML description is interpreted by visually based P300 BCI software which enables different flashing modes. We use the middleware to control devices (a light switch and a TV set) and applications (a Virtual Reality environment and a spelling application). However, the used XML tags can easily be extended to be interpreted either visually or acoustically: By interchanging BCI software blocks (see Figure 2) the same set of devices and applications is accessible for users with varying physical abilities. A potential line of work for improving the XML scheme is the dynamic adaption of the GUI contents to a changing context: For example, an unexpected behavior from a connected device may require a device reset. An adaptive GUI is expected to provide flexibility and to enhance user's experience with the BCI.

Acknowledgments

The research leading to these results has received funding from the EC's 7^{th} Framework Programme under the projects BrainAble, Better and Vere and from the AAL project ALIAS.

References

[1] J. R. Wolpaw and C. B. Boulay. Brain signals for brain-computer interfaces. In *Brain-Computer Interfaces*. Springer, 2010.

[2] C. Guger, C. Holzner, C. Grönegress, G. Edlinger, and M. Slater. Brain-computer interface for virtual reality control. In *Proceedings of ESANN*, 2009.

[3] C. Guger, S. Daban, E. Sellers, C. Holzner, G. Krausz, R. Carabalona, F. Gramatica, and G. Edlinger. How many people are able to control a P300-based brain-computer interface (BCI)? *Neuroscience Letters*, 462:94 – 98, 2009.

[4] R. Prückl and C. Guger. A brain-computer interface based on steady-state visual evoked potentials for controlling a robot. In *10^{th} International Work-Conference on Artificial Neural Networks*, 2009.

[5] C. Guger, C. Holzner, C. Grönegress, C. Edlinger, and M. Slater. Control of a smart home with a brain-computer interface. In *4^{th} International Brain-Computer Interface Workshop and Training Course*, 2008.

[6] S. G. Mason and G. E. Birch. A general framework for brain-computer interface design. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11:70–85, 2003.

The BCIClassifier: a User-Friendly and Features Rich Tool for Brain-Computer Interfaces

L. R. Quitadamo[1,2], D. Mattia[2], L. Astolfi[2,3,4], J. Toppi[2,3], M. Risetti[2], F. Cincotti[2], L. Bianchi[2,5]

[1]Department of Electronic Engineering, University of Tor Vergata, Rome, Itay
[2]Neuroelectrical Imaging and BCI Laboratory, Fondazione Santa Lucia, Rome, Italy
[3]Department of Computer Science and Systems, University La Sapienza, Rome, Italy
[4]Department of Physiology and Pharmacology, University La Sapienza, Rome, Italy
[5]Department of Neuroscience, University of Tor Vergata, Rome, Italy

l.quitadamo@hsantalucia.it

Abstract

Brain-Computer Interface (BCI) systems can be used to pilot an external peripheral without using muscular control but just by means of cerebral activity. In order to analyze whatever BCI related aspect, tools have been developing that are versatile, powerful and user-friendly and allow to use different preprocessing methods and feature extraction strategies, to run different classification algorithms, to evaluate system performances and also to test the reliability of results. All these facilities are suitable to analyze data coming from different BCI protocols, so different types of EEG responses, and from different patients who can benefit from a BCI technology, and will provide a valid and robust mean to foster the BCI technology application in clinical environment.

1 Introduction

Brain-Computer Interface (BCI) systems are meant to allow communication without using the classical brain paths, such as nerves or muscles, but just by using cerebral signals, for example the EEG. They are mainly intended for those people that have lost the possibility to interact with the environment after a disabling trauma or pathology.
Although several EEG-based BCI systems have been developed over the past 2 decades, they have not yet reached the maturity to be translated into a widespread clinical application. One of the limits is represented by the system robustness and by necessary expert supervision. Here we describe a novel tool, the BCIClassifier, deployed within the context of DECODER project (www.decoderproject.eu), which aims at providing an easy-to-use and reliable BCI platform for clinical applications in non-responsive patients (i.e. disorder of consciousness; severely motor impairment) with a twofold aim: i) assessment of cognitive functions by means of active commands and ii) once a voluntary and robust response is detected, a channel for binary communication. This tool, that is part of the NPXLab suite (www.brainterface.com), allows to use different preprocessing methods and feature extraction strategies; to run different classification algorithms, to evaluate system performances and also to test the reliability of results. Also, it has been implemented as compatible with different file formats, so that data coming from different laboratories can be analyzed; it contains interfaces for the classification of data relative to different protocols and for the evaluation of the performances of systems and it associates statistical analyses to classification results for testing their reliability. All these features, together with its easiness to use and its versatility, make this tool a powerful mean for analyzing data coming from the different typologies of patients who can benefit from BCI technology.

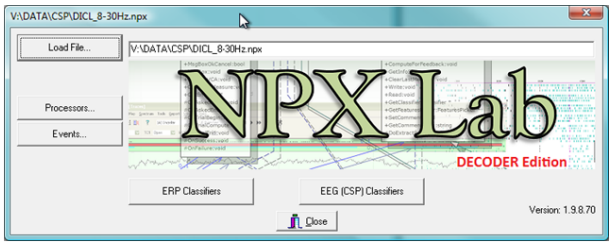


Figure 1: Main form of the BCIClassifier.

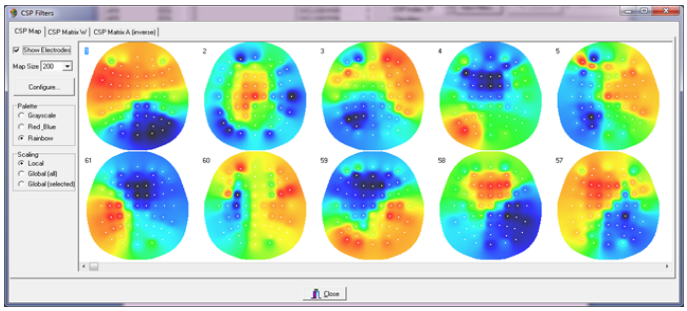


Figure 2: The form for viewing Common Spatial Patterns.

2 Methods

The BCIClassifier was written in C++ for the Windows platform and has been implemented on the basis of the functional model described in [1]. The starting form of the program allows to select two main modalities, one based on event related potentials (ERPs), by selecting "ERP Classifiers" and one on mental tasks imagery, by selecting "EEG (CSP) Classifier", see Figure 1.

2.1 Spatial Filtering: ICA and CSP

Among the preprocessing methods ("Processors" button in the same form of Figure 1), the BCI-Classifier includes some spatial filtering techniques, such as Laplacian, common average reference (CAR), Independent Component Analysis (ICA) and Common Spatial Patterns (CSP). For what concerns ICA, once the components have been computed, it is possible to remove from the signal the undesired ones, such as those relative to ocular or muscular artifacts, therefore cleaning the traces for further analyses. For what concerns CSP, instead, the method, given two different conditions, for example two different mental or motor imagery tasks, allows to compute different patterns that, premultiplied to the original signal, maximize the differences in the variances of the two tasks, so increasing the discriminability of the classes for the successive classification phase, as reported in [2]. An example of the best ten patterns relative to two different mental tasks, spatial navigation on the top maps and playing tennis on the bottom ones, is reported in Figure 2.

2.2 The Main Classifier Form

The main form for the classification process is reported in Figure 3. In the example, the form relative to ERP is shown, while for the CSP studies a similar form has been implemented.

Different actions are possible in order to set all the configurable parameters for the analyses: 1) The selection of the channels set on which to perform the analyses. 2) The selection of temporal intervals and decimation factors, in order to set the time samples to input to the classifiers. 3) The selection of the encoder, that is the map between the class detected by the classifier (logical symbols, e.g. the rows and columns in a P300 speller application) and the final output (semantic

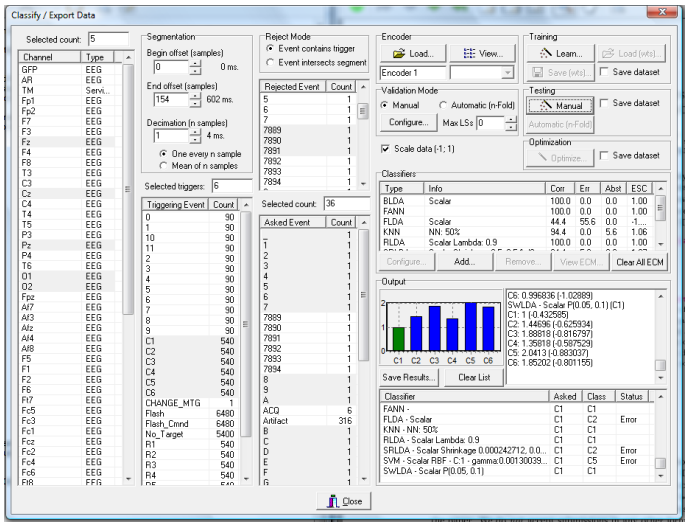


Figure 3: The main form of the ERPClassifier.

symbols, e.g. the characters on the spelling matrix). 4) The selection of the classification modality, that can be manual or automatic. The manual modality allows to select the events in the file that will be used to train the classifiers and those that will be set for the testing phase. The automatic modality, instead, allows to perform n-folded classifications, during which the entire subset is randomly divided in n sets that alternatively act as training or testing set. The operations can be iterated k times to increase the number of classifications. 5) The selection of the classification algorithms among StepWise Linear Discriminant Analysis (SWLDA), Fisher LDA, Bayesian LDA, Regularized LDA (basic or by means of shrinkage), Artificial Neural Network (ANN), k-nearest neighbor (kNN), Support Vector Machine (SVM). Each classifier is characterized by different parameters that the user can set, for example, the p-in and p-out values in the SWLDA model, the kernel functions in SVM, the number of layers and the activation functions in ANN, and so on, so creating a virtually unlimited number of different classifier configurations.

2.3 The Extended Confusion Matrix (ECM) and the Metrics Form

After running classifications, the performances of each classifier are stored into the Extended Confusion Matrix [3], that can be visualized in the ECM form, Figure 4. In an ECM the desired (rows) and actual (columns) classifications are represented, so that one can immediately realize the behavior of a transducer. For example in Figure 4, a mental task protocol was analyzed based on the imagination of playing tennis and of spatial navigation (36 trials for each task). A leave-one out crossvalidation was performed (all trials but one were used for computing CSPs and training the classifier and the remaining for testing). The ECM (top) matrix shows that the navigation task was successfully detected 28 times and misclassified as tennis for 8 times, whereas the tennis task was correctly classified 29 times and misclassified 7. In the same form, several performance indicators are shown, such as the accuracy (percentage of correct classifications, 79.16 %), the error rate (the percentage of misclassifications, 20.83 %), the Information Transfer Rate (ITR), the mean Expected Selection Cost (ESC) [3], the Cohen's Kappa Coefficient, the Mutual Information, etc.

Finally, it is possible to visualize the results of statistical tests (e.g chi-square test) on the obtained classifications in order to investigate the reliability of the results with respect to the chance level. For example, an accuracy of 60 % with a significance of the 99 % would be surely better than an accuracy of 99 % with a significance of the 60 %. In the example in Figure 4, the chi-square value is 24.5, while the p-value is $<< 0.01$, therefore the test is highly significant.

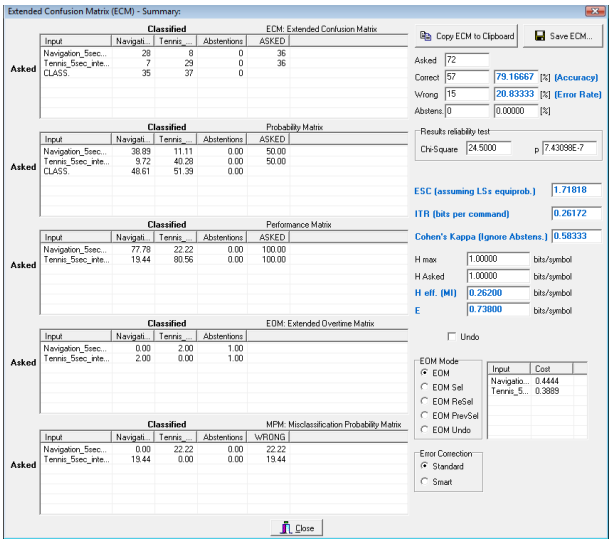


Figure 4: The form for viewing the ECM and the metrics.

3 Discussion and Conclusion

The facilities of the BCIClassifier make it a powerful mean for the typical analyses to be performed in a BCI environment. The opportunity to analyze data coming from different protocols, to store them in different file formats, to preprocess and classify data with different modalities (also by comparing several classification algorithms), to evaluate the performances of BCI systems with different metrics, represents a first step towards a large scale clinical applications of BCI technologies. Furthermore, the possibility to test the reliability of results by means of a statistical analysis, will provide information about the robustness of detected brain responses: this means that, whenever the p-value of the test is lower than a set threshold, with a reasonable certainty we are detecting differences in the responses to a specific request that are not due to chance; this is of great importance even in cases of minimal residual or apparently absent voluntary responses. This means that real-world applications will be necessary to finally assess the value of this tool.

Acknowledgments

This work is partially supported by the European ICT Programme Project FP7-247919. 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. R. Quitadamo, M. G. Marciani, G. C. Cardarilli, and L. Bianchi. Describing different brain computer interface systems through a unique model: a UML implementation. *Neuroinformatics*, 6(2):81–96, 2008.

[2] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K. R. Müller. Optimizing spatial filters for robust EEG single-trial analysis. *IEEE Signal Proc Mag*, 25(1):41–56, 2008.

[3] L. Bianchi, L. R. Quitadamo, G. Garreffa, G. C. Cardarilli, and M. G. Marciani. Performances evaluation and optimization of brain computer interface systems in a copy spelling task. *IEEE T Neur Sys Reh*, 15(2):207–16, 2007.

TiA – Standardizing Raw Biosignal Delivery in BCIs

C. Breitwieser[1], C. Neuper[1,2], G. R. Müller-Putz[1]

[1]Institute for Knowledge Discovery, BCI Lab, Graz University of Technology, Graz, Austria
[2]Department of Psychology, University of Graz, Graz, Austria

c.breitwieser@tugraz.at

Abstract

TiA is a concept to transmit raw bisoignals in a standardized way for BCI purposes. It provides the possibility for multirate and and block-oriented data transmission. Different kinds of signals, divided into so called signal types (e.g., EEG, EMG, ECG) can be transmitted at the same time, whereby a data distinction is always possible. TiA utilizes a client-server principle with one server performing data acquisition and multiple clients as data consumers. Data is divided into immutable meta information and raw biosignals. A standardized handshaking protocol and TiA data packets were introduced. Applying this concept an abstraction layer is evolved facilitating the standardization process for BCI development.

1 Introduction

Many different Brain-Computer Interface (BCI) systems like BCI2000, OpenVibe or xBCI [1–3] have been built since 1973 when Vidal mentioned his idea of a BCI the first time in literature [4]. Generally speaking, BCIs deal with brain signals and provide a user the possibility to gain control of some device just using his thoughts [5]. Thirty years after Vidals idea, Mason and Birch introduced the first common structure of a BCI [6] as shown in Figure 1. It is visible that a BCI is divided into distinct modules connected via different interfaces. Standardizing those interfaces would be a first step to provide the possibility for interchangeable BCI components and ensured compatibility within BCI systems by supporting the standardized interfaces. Within this work the first interface in Masons model ("Amplifier" - "Feature Extractor") was analyzed and a first standardization concept was evolved. Comparing mentioned BCI systems in the beginning, different commonalities regarding the raw data acquisition like recording a defined number of channels with a given sampling rate can be found. But no standardized interface is yet available although data processing is similar. We introduced a concept called TiA (Tools for BCI - Interface A) facilitating standardization of raw biosignal data acquisition. To additionally cope with BCI systems also processing other signal types than just brain signals (e.g., EOG, EMG, joysticks,...), called hybrid BCIs [7–9] a broad spectrum of potential kinds of signals for BCI usage was taken into account. With TiA it is possible to deal with various kinds of different signals in a standardized form and evolve an abstraction and standardization layer this way.

2 Requirement Analysis and Design

BCIs provide the possibility to gain control just by thought. In case of patients having residual motor abilities, using just a BCI alone might not be the ideal solution. Combining a BCI with assisitive devices (AD) like buttons and joysticks or even treating a BCI as an AD itself might be a better solution to enhance the patients quality of life. Therefore a BCI system should not be limited just processing brain signals.

Different BCI systems have already been built using pure brain control as well as hybrid combinations. EEG [9], magnetoencephalogram (MEG) [10], or the NIRS signal [11] have already

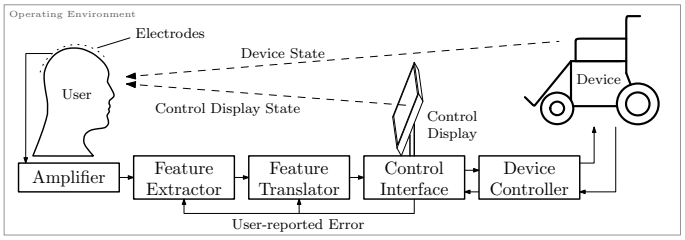
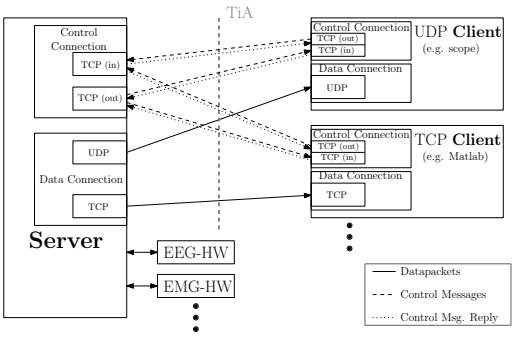


Figure 1: Functional model from Mason and Birch (modified from [6]).

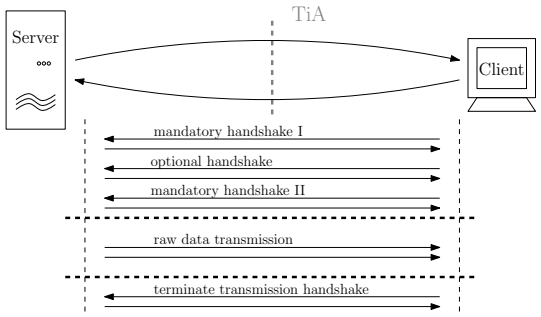
been utilized to control a BCI using mentioned brain signals. Developing hybrid BCI (hBCI) systems, the number of potential kinds of possible signals further increases. Signals like the electromyogram (EMG) [9] have already been successfully combined with EEG to control an hBCI. But also other signals like electrooculogram (EOG) or information delivered by various assistive devices (e.g. buttons or joysticks) could be used in combination with an arbitrary brain signal to form an hBCI. A concept to distinguish between different kinds of signals was introduced using so-called "Signal Types", allocating every different kind of signal a unique identifier. Within the work from Mason and Birch [6], a unidirectional transmission from the amplifier to the feature extraction is presented (see Figure 1). Such a principle can be realized using a client-server architecture. Therefore the data acquisition interacts as a server and processing modules act as clients. Using a client-server approach facilitates multiple and also distributed processing chains (within [6] only one processing chain is presented).

2.1 Design Principles

TiA transmits two types of information: (i) mutable and (ii) immutable information. For that reason, information exchange within TiA is split into initial meta information transmission (for immutable information) and a continuous data stream (for mutable data) to the client. The meta information is transmitted via xml based control messages utilizing a defined handshaking procedure. All mutable data like the voltage from an EEG channel is transmitted via a binary unidirectional data stream to the client. Thus individual data streams can be established to different clients, being able to attach to the server at runtime. The principle TiA is working is illustrated in Figure 2(a). A whitepaper describing the design and implementation of TiA (e.g., signal type flags) is available at arXiv.org [12].



(a) TiA principle showing one server and two clients.



(b) TiA - client-server handshaking principle.

Figure 2: TiA client-server scheme and handshaking principle.

3 Software Design

Data from potentially multiple hardware devices is acquired via a single data acquisition system, implementing the TiA interface. A connection is set up using HTTP-like control messages (hyper text transfer protocol) sent over two separate TCP (transmission control protocol) connections. Those messages are utilized for the handshaking process between client and server. Those control messages could optionally be extended with meta information. Mutable raw data is transmitted using either a TCP or UDP (user datagram protocol). The desired protocol has to be chosen during the handshaking process. The first TCP connection used for the handshaking process is client-server oriented, the second one server-client oriented. Establishing and supporting the client-server is a mandatory requirement of TiA. Supporting the server-client connection is optional. Messages received on one connection are always answered on the same one.

3.1 Handshaking Process

Figure 2(b) illustrates the principles of the client-server communication. Arrows represent the information exchange. Messages are utilize defined control messages and data packets and are transmitted in a standardized way. The transmission can be separated into mandatory and optional handshaking, the raw data transmission and a termination handshaking procedure. In case of an error during this handshaking procedure no connection is created. Detailed information concerning the TiA handshaking procedure is available at arXiv.org [12].

3.2 Data Transmission

A binary data stream is utilized to transmit acquired mutable raw data using TiA data packets. The TiA data packet is divided into three sections: (i) a fixed header; (ii) a variable header; and (iii) the raw data. Parsing and interpreting the content of the fixed and variable header it is possible parse and read the whole data packet. It is also with a timestamp, a packet number, and a unique identifier per connection. This facilitates proper timing and provides the possibility to detect lost packets using UDP. Signal types stored inside the data packet can be identified with a unique flag. Within TiA a constant number of channels over time is assumed, but block sizes and sampling rates can vary between signal types. One signal type however must have only one block size and one sampling rate. Raw data is stored using a 32 bit binary single precision floating point number. A single data packet can contain different signal types at once. A detailed description of TiA and the data packet is available at arXiv.org [12].
Performing the TiA handshaking process is a mandatory requirement to receive any data packet. Additionally TCP or UDP has to be chosen as a data transmission protocol. Lost packet if using UDP are not re-sent. To ensure data reception, TCP has to be chosen. In case of TCP, a separate TCP connection is established from the client to the server. Data packets are sent over this connection.

4 Discussion

With TiA we have introduced a standardized layer between the amplifier and the feature extractor module as described described by Mason and Birch [6] in their functional BCI models. For this purpose a TiA library and a data acquisition software using TiA have been implemented and can be downloaded at: `http://bci.tugraz.at/downloads.html`. The library has been successfully integrated into different programs used for BCI purposes nowadays (BCI2000, Matlab). Further improvements like extended meta information (electrode location, head geometry, hardware information,...) SSL encryption to protect patient-sensitive meta-information are currently in development. Using TiA is a first step towards Masons standardized BCI model as it becomes possible to decouple a BCI system from the used data acquisition hardware.

Acknowledgments

This work is supported by the European ICT Programme Project FP7-224631. 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] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw. BCI2000: a general-purpose brain-computer interface (BCI) system. *IEEE Trans Biomed Eng*, 51(6):1034–1043, 2004.

[2] Y. Renard, F. Lotte, G. Gibert, M. Congedo, E. Maby, V. Delannoy, O. Bertrand, and A. 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.

[3] I. P. Susila, S. Kanoh, K. Miyamoto, and T. Yoshinobu. xBCI: A generic platform for development of an online BCI system. *IEEJ Trans Elec Electron Eng*, 5(4):467–473, 2010.

[4] J. J. Vidal. Toward direct brain-computer communication. *Annu Rev Biophys Bioeng*, 2:157–180, 1973.

[5] 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.

[6] S. G. Mason and G. E. Birch. A general framework for brain-computer interface design. *IEEE Trans Neural Syst Rehabil Eng*, 11(1):70–85, 2003.

[7] J. D. R. Millán, R. Rupp, G. R. Müller-Putz, R. Murray-Smith, C. Giugliemma, M. Tangermann, C. Vidaurre, F. Cincotti, A. Kübler, R. Leeb, C. Neuper, K. R. Müller, and D. Mattia. Combining brain-computer interfaces and assistive technologies: State-of-the-art and challenges. *Front Neurosci*, 4, 2010.

[8] G. Pfurtscheller, B. Z. Allison, C. Brunner, G. Bauernfeind, T. Solis-Escalante, R. Scherer, T. O. Zander, G. R Müller-Putz, C. Neuper, and N. Birbaumer. The hybrid BCI. *Front Neurosci*, 4:42, 2010.

[9] R. Leeb, H. Sagha, R. Chavarriaga, and J. Del R. Millán. Multimodal fusion of muscle and brain signals for a hybrid-BCI. *Conf Proc IEEE Eng Med Biol Soc*, 1:4343–4346, 2010.

[10] 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.

[11] 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(3):219, 2007.

[12] C. Breitwieser and C. Eibel. TiA - Documentation of TOBI Interface A. *ArXiv e-prints*, 2011.

Slow Phase-Related Oscillations of Prefrontal (De)oxyhemoglobin and Central EEG Alpha and Beta Power in the Resting Brain

G. Pfurtscheller[1], G. Bauernfeind[1], C. Neuper[1,2]

[1]Institute for Knowledge Discovery, Laboratory of Brain-Computer Interfaces, Graz University of Technology, Graz, Austria
[2]Department of Psychology, University of Graz, Graz, Austria

pfurtscheller@tugraz.at

Abstract

The present work examines the close interaction between brain and heart and shows that slow (de)oxyhemoglobin oscillations in the prefrontal cortex and changes in blood pressure and heart rate are coupled with a cyclic EEG alpha and beta power decreases in motor areas in the resting brain.

1 General Overview

Important features of biological systems are slow and ultraslow oscillations, particularly cyclic fluctuations around 0.1 Hz. Large-scale infraslow oscillations in the human cortex ranging from 0.02 to 0.2 Hz were reported by Vanhatalo and colleagues [1] in full-band EEG recordings. In this respect the study of Monto et al. [2] is of particular interest, because they demonstrated that the phase but not the amplitude of such slow oscillations correlated strongly with the subject's psychophysical performance. These results, and the observation that combined fMRI-EEG studies have demonstrated slow (<0.1 Hz) baseline fluctuations during resting wakefulness [3], strongly support the idea that slow EEG fluctuations are associated with fluctuations of the excitability of large-scale cortical networks (resting state networks, RSNs). Slow oscillations have not only been observed in cerebral oxyhemoglobin concentrations, oxygen availability of cortical tissue, cerebral blood flow velocity, and cerebral blood flow, but also in the cardiovascular system (so-called Mayer waves around 0.1 Hz [4]).

Of interest is the mutual interaction between brain and heart. From non-human animal research, we have evidence that cortical systems, especially including the medial-prefrontal cortex, act as a network together with subcortical systems (known as the "central autonomic network") to initiate and represent cardiac autonomic adjustments (Figure 1). This cooperative behavior of the two most important organs of the body is not unexpected, the great French physiologist Claude Bernard first discovered the connection between brain and heart (for a review see [5]). It is well documented that the prefrontal cortex is linked to autonomic motor circuits responsible for sympathoexcitatory and parasympathoinhibitory control of the heart. This means cortical activity controls cardiac activity, not only during movement execution and motor imagery, but also in the resting brain.

Prior to self-paced movement, the Bereitschaftspotential is generated and the EEG desynchronized. Aditionally, the heart rate decelerates and the heart rate intervals (RRIs) increase, respectively. All these changes in brain and heart activity start at least 1-3 seconds before the motor act, are unconscious and enter consciousness around 250 ms before movement. These observations can be seen as an example of the close interaction between brain and heart and the relatively long time durations of the individual phenomena in the order of some seconds.

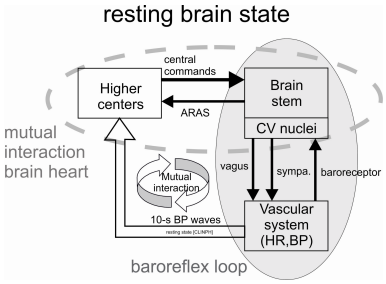


Figure 1: Simplified scheme of interactions between brain and heart during rest.

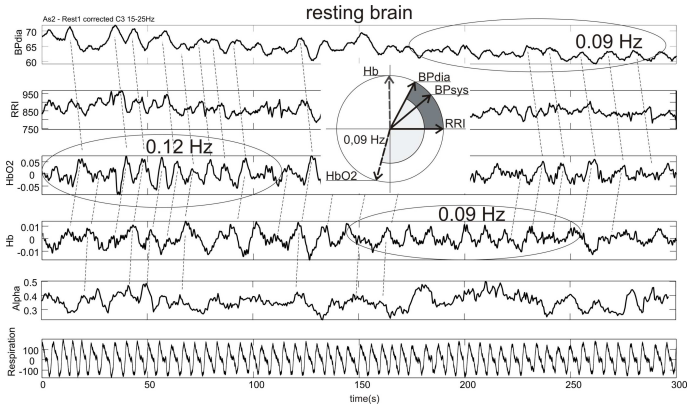


Figure 2: Time courses of diastolic BP (BPdia), RRI, HbO_2, Hb, respiration and EEG beta power recorded during 5 minutes of rest. Dominant in all traces are the varying oscillations around 0.1 Hz. The phase relationships between slow waves at 0.09 Hz are indicated by stippled lines and in the phase plot. Note the instability of slow cyclic oscillations especially in the (de)oxyhemoglobin concentration

2 Methods

We continuously measured blood pressure (BP), heart rate intervals (RRI), prefrontal oxyhemoglobin (HbO_2) and deoxyhemoglobin (Hb) with the near-infrared spectroscopy (NIRS) method [4] and central EEG signals during 5 minutes of rest. 19 healthy subjects were investigated, about 60% of whom had slow HbO_2/Hb oscillations around 0.1 Hz in the resting brain.

3 Results and Discussion

During 5 minutes of rest, the BP, RRI and HbO_2/Hb exhibited slow oscillations around 0.1 Hz, which were non-stationary. Examples of one characteristic subject are shown in Figure 2. The average phase shift between diastolic BP and RRI oscillations was 81 ± 21^o (mean±SD) with BP leading. The phase shift between BP and HbO_2 varied between 0 and 200^o with BP also leding. Some subject displayed a clear phase relationship between prefrontal HbO_2/Hb oscillations and central EEG alpha or beta power changes during rest.

In summary, slow out-of-phase HbO_2 - Hb oscillations in prefrontal cortex are coupled with a cyclic EEG alpha and beta power decrease in motor areas in the resting brain. Hence, the timing of movement execution or imagery at free will appears to be modulated by fluctuations of brain excitability in the prefrontal cortex coupled to motor cortex activity.

4 Conclusion

Three points are of interest for BCI research and new developments in the near future:

- First, oscillations around 0.1 Hz have to be considered as physiological noise when an optical BCI is realized. They can be removed, e.g. by the use of a transfer function model. The prefrontal cortex is of special interest because the optodes can be easily attached there and because the user's intent is accompanied by the activation of prefrontal network. Such optical BCIs with optodes placed over the prefrontal cortex are especially suitable for use at home or work.
- Second, rhythmic cortical excitability changes around 0.1 Hz can be used to present the cue-stimuli phase-related to slow prefrontal oxyhemoglobin, blood pressure or heart rate oscillations during e.g. ERD-based BCI training. It can be expected that such an experimental paradigm can shorten the trainings time and improve the BCI performance.
- Third, pathways originating in the prefronal cortex can activate cardiovascular neurons in the brain stem and result in cardiac changes. Such heart rate changes can be on the order of 10 beats/min during foot motor imagery, and easily be detected in the ongoing ECG signal [6]. Parallel processing of EEG and ECG signals in a hybrid BCI can improve the BCI performance.

5 Acknowledgments

The authors' BCI research has been supported by the EU project PRESENCCIA (IST-2006-27731), "Land Steiermark" (project A3-22. N-13/2009-8) and the Neuro Center Styria (NCS) in Graz, Austria.

References

[1] S. Vanhatalo, J. M. Palva, M. D. Holmes, J. W. Miller, J. Voipio, and K. Kaila. Infraslow oscillations modulate excitability and interictal epileptic activity in the human cortex during sleep. *Proc Natl Acad Sci U S A*, 101(14):5053–5057, 2004.

[2] S. Monto, S. Palva, and J. Voipioand J. M. Palva. Very slow EEG fluctuations predict the dynamics of stimulus detection and oscillation amplitudes in humans. *J Neurosci*, 28(33):8268–8272, 2008.

[3] D. Mantini, M. G. Perrucciand C. Del Gratta, G. L. Romani, and M. Corbetta. Electrophysiological signatures of resting state networks in the human brain. *Proc Natl Acad Sci U S A*, 104(32):13170–13175, 2007.

[4] G. Pfurtscheller, D. Klobassa, C. Altstätter, G. Bauernfeind, and C. Neuper. About the stability of phase-shifts between slow oscillations around 0.1 hz in cardiovascular and cerebral systems. *IEEE Trans Biomed Eng*, PP(99):1, 2011.

[5] J. F. Thayer and R. D. Lane. Claude bernard and the heart-brain connection: further elaboration of a model of neurovisceral integration. *Neurosci Biobehav Rev*, 33(2):81–88, 2008.

[6] G. Pfurtscheller, B. Z. Allison, C. Brunner, G. Bauernfeind, T. Solis-Escalante, R. Scherer, T. O. Zander, G. R. Müller-Putz, C. Neuper, and N. Birbaumer. The hybrid BCI. *Front Neurosci*, 4:30, 2010.

A Preliminary Survey on the Perception of Marketability of Brain-Computer Interfaces (BCI) and Initial Development of a Repository of BCI Companies

F. Nijboer[1], B. Z. Allison[2], S. Dunne[3], D. Plass-Oude Bos[1], A. Nijholt[1], P. Haselager[4]

[1]Human Media Interaction, University of Twente, Enschede, the Netherlands
[2]Graz University of Technology, Graz, Austria
[3]Starlab, Barcelona, Spain
[4]Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, the Netherlands

femke.nijboer@utwente.nl

Abstract

The marketability of BCI applications may greatly influence the decisions of goverments, the industry and academia. In this paper we first explored with a survey when respondents ($N = 145$), who were present at the 4th International BCI Meeting, expect that different BCI applications will become commercially available. Second, we surveyed how well existing BCI companies are known to respondents. Third, we compared the findings with our own preliminary overview of the marketability of BCIs and our repository of 28 companies. Respondents were optimistic about the marketability of BCIs for healthy users and users who need assistive technology (AT), but 72.4 % of the respondents was unaware that companies already exist which market BCI's. Based on a preliminary market overview we cautiously suggest that optimism in relation to applications for healthy users is more appropriate than in relation to BCI-based AT. In future we plan surveys among a broader range of stakeholders and more profound analyses of the market.

1 Introduction

Brain-Computer Interface (BCI) research is rapidly increasing, with considerable enthusiasm for applications for users with and without physical disabilities [1, 2]. A roadmap for BCI research and development could help address emerging markets and opportunities. The future Brain/Neuronal Computer Interaction project (fBNCI; http://future-bnci.org/) aims to develop such a roadmap by drawing on the expertise of BCI researchers, as well as many other stakeholder groups, including companies, end users, patient organizations, policy makers, and the general public. One focus of the roadmap is to evaluate the marketability of current and future BCI applications. Hence, we surveyed researchers on the marketability of BCIs at the 4th International BCI Meeting, which took place at the Asilomar conference centre in June 2010. This conference provided an excellent opportunity to engage many qualified respondents, although it has to be noted that many other qualified respondents (e.g. from the field of human-computer interaction or ambient intelligence) do not typically attend this meeting which may bias our findings (see Section 5). Other results from this survey and more demographical data of the repondents are available elsewhere ([3]) and under review ([4]).In this paper we focus on three aspects of the survey. First, we explored when respondents expect that different BCI applications will become commercially available (see Section 3.1). Second, we surveyed how well existing BCI companies are known to respondents (see Section 3.2). Third, we compared the findings with our own preliminary overview of the marketability of BCIs and our repository of companies (see Section 3.3).

2 Methods

A total of 145 (105 males, 39 females) out of 289 conference attendees responded to the questions about the marketability of BCIs. Seventy three persons were aged between 18 and 30, sixty-nine persons between 31 and 55 and two persons were aged between 56 and 70. The sample consisted of experts from various disciplines (e.g. neuro-, computer or cognitive scientists, electrical engineers, psychologists etc). One participant did not give demographical data.

Participants completed an online survey that included four questions on the expected marketability of different purpose BCIs. In the survey, for convenience, we ordered some BCI types based on their function: 1) BCIs for healthy users, 2) BCIs as AT, 3) BCI-controlled prostheses and 4) BCIs as therapy tools.

Specifically, participants were asked to indicate when they expected to see these types of BCI systems to become available on the market. Participants could choose from 5 answer options: "never", "between 0–5 years", "between 5–10 years", "more than 10 years" or "it already exists on the market". Participants who indicated that this type of BCI already exists on the market were asked which group or company offers the product.

Only companies which offered BCI products and services were included in the repository. Thus, companies supplying hardware and software needed for BCI research were not included. The repository was build through: 1) web searches, 2) written and verbal interviews of experts and 3) iterative postings on LinkedIn groups ("brain-computer interface group", "neuromarketing" and "BrainGain").

3 Results

3.1 Respondents' Expectancies of Marketability of BCI

Table 1 presents the respondents' expectancies per BCI class.

BNCI technology for:	on market	0–5 yrs	6–10 yrs	> 10 yrs	never	# of resps
healthy users	26.6 %	44.8 %	13.8 %	12.4 %	1.4 %	145
users who need AT	16.6 %	42.1 %	33.8 %	7.6 %	0.0 %	145
users who need prostheses	0.7 %	20 %	43.4 %	34.5 %	1.4 %	145
users who need therapy	9.9 %	39.4 %	39.4 %	10.6 %	0.7 %	142

Table 1: Overview of percentage of respondent which indicated when BCI applications will enter (if ever) the market for healthy users, users who need AT, prostheses and therapy. The last column shows the number of respondents (# of resps) who rated the item.

3.2 Respondents' Knowledge About Existing Companies

Thirty-two respondents reported which companies they knew that already marketed BCI's for health users. By far the best known companies are Neurosky (counted 13 times) and Emotiv (counted 14 times). Other companies that were mentioned: Hitachi, Ambient, OCZ technology, Starwars Science, Mattel, inc., g.Tec, Brain Actuated Technologies, InteraXon, Zeo, inc. and Interactive Productline. Twenty-two respondents reported which companies they knew that already marketed BCIs as AT. The best-known company is g.Tec (counted 16 times). Other groups and companies that respondents listed were: Ambient, Brain Actuated Technologies, OpenVibe and BCI2000. Nine respondents commented that they knew BCI's as therapy tools already are on the market, but only two gave concrete company names: Brain master Technologies, inc. and EEGinfo. None of the respondents identified a company that already markets BCI-based prostheses.

3.3 Preliminary Overview of Companies and Marketability

We summarized a preliminary overview of 28 companies related to Brain-Computer Interfacing (see Figure 1), which we expect to be more complete at the time of the Graz conference. Currently, neurofeedback companies are not yet listed in the repository. We invite all readers to comment on this overview and continue completing and criticizing the overview so we can offer the European Commission the most accurate summary possible. This overview currently shows that as many as 10 companies market BNCI's for entertainment and gaming. Six companies offer BNCI as AT, although two of these companies (Neural Signals, inc and BitBrain) do not yet market products. However, there is a big difference in the number of products sold between the two markets. Neurosky, for example, has sold approximately 1 million Mindsets, whereas g.Tec has sold between 30–40 Intendixes (personal communication). Our survey did not address neuromarketing. We counted six companies in this area. BCI researchers ($N = 144$) have divided opinions on whether neuromarketing technologies do (35.4 %) or do not count (47.9 %) as BCI systems (Nijboer et al, in review). Nevertheless, neuromarketing developments influence BCIs, and it is predicted that this market will grow tremendously [2, 5]. Finally, two companies (Brain Fingerprinting inc. and No Lie MRI) provide BCI services for criminal investigation.

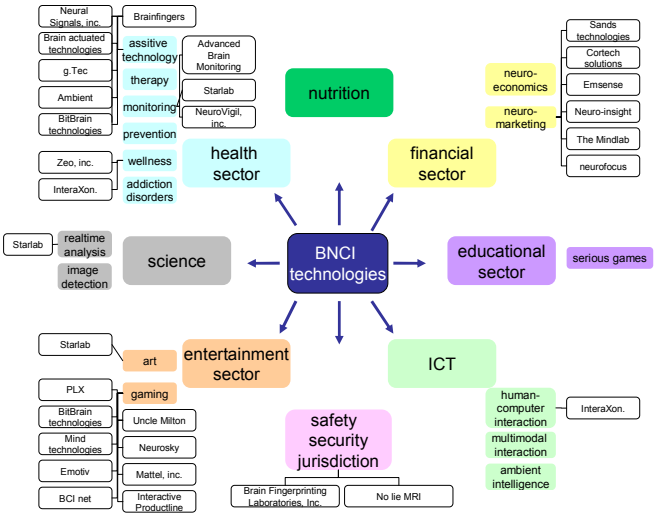


Figure 1: Preliminary overview of existing companies.

4 Discussion and Conclusion

The survey results show that respondents were generally optimistic about the near-term marketability of BCI's for healthy users and user in need of AT. Repondents were slightly less optimistic about BCIs for therapy and expect BCI-controlled neuroprostheses to require more than 10 years to enter the market. Our preliminary overview shows that companies - more than respondents realize - already market BCIs for healthy users and users who need AT. However, having a product on the market does not yet mean that a product is sold successfully. BCI research toward AT has a history of several decennia yet 30–40 products were sold by g.Tec. BCIs for healthy users are available since less than 5 years and (at least) more than a million products have been sold. It is remarkable that 72.4 % of the respondents are not aware that BCIs for healthy users are already on the market. Since 10 companies already exist in this direction we argue that optimism for this type of applications is quite appropriate. In our opinion BNCI for AT is not

likely to create large economic value, unless it becomes available for a broader range of categories of (more numerous) end-users. The societal value of AT-BCI products is enormous, since they can enable users with disabilities to participate in the society at interpersonal and professional levels. However, we predict that the research and development of BCIs targeted at the general public will more easily mobilize the industry and facilitate the tech transfer between universities and industry. This will lead to more acceptance of BNCI technology in the public and possibly to a spillover to the AT industry. In addition, even though the market for BCI-controlled prosthetics requires a more time-consuming trajectory, ultimately we expect these products to create large economic and societal value similar to Deep Brain Stimulation, because the number of potential end-users is relatively high.
The survey reported here contains some flaws. First, since we could not make the entire survey too time-consuming, some questions were generalized. For instance, we did not specify whether the BCI products would work with non-invasive or invasive brain signals nor did we consider the usability of diffent BCIs. However, these specifications would influence the cost, number of available users, training time, time to market and many other factors. Second, the respondents were mostly people involved primarily in BCI research and had the time, funds, and enthusiasm to attend a major conference, so the sample may be positively biased. fBNCI in future will survey a broader range of stakeholders, including more experts from human-computer interaction and the industry.

Although we view these data as a first crude estimation of the fields' perception of marketability of BCIs, we cautiously conclude that BCI researchers seem optimistic about how much time it will take before the first BCIs for non-medical purposes and AT enter the market. We agree with them that the market for BNCI applications for healthy users is promising, but we are not convinced that the market for AT-BCI applications is equally promising simply because the number of end-users is relatively low in comparison. Moreover, the limited knowledge of respondents about existing companies should raise awareness in the BCI community that tech transfer is already well established largely without the involvement of the scientists who consider themselves the core of the BCI field. A more pro-active role of BCI scientists in tech transfer would be desirable for high quality product development.

5 Acknowledgments

The authors gratefully acknowledge the support the Future BNCI project (Project number ICT-248320) and of the BrainGain Smart Mix Program. We thank all respondents and Thorsten Zander, Jan van Erp, Yann Renard, Fabien Lotte, Ariel Garten, Tom Sulivan, Christoph Guger and Hendri Hondorp for valuabe input.

References

[1] A. Nijholt and D. Tan. Brain-computer interfacing for intelligent systems. *Ieee Intelligent Systems*, 23(3):72–72, 2008.

[2] B. Z. Allison. Toward ubiquitous BCIs. In Bernhard Graimann, Gert Pfurtscheller, and Brendan Allison, editors, *Brain-Computer Interfaces*, The Frontiers Collection, pages 357–387. Springer Berlin Heidelberg, 2010.

[3] F. Nijboer, J. Clausen, B. Z. Allison, and P. Haselager. Researchers opinions about ethically sound dissemination of BCI research to the public media. *International Journal of Bioelectromagnetism*, in press.

[4] F. Nijboer, J. Clausen, B. Z. Allison, and P. Haselager. The Asilomar survey: researchers opinions on ethical issues related to brain-computer interfacing. *Neuroethics*, in review.

[5] Z. Lynch. *The Neuro Revolution: How Brain Science Is Changing Our World*. St. Martin's press, New York, 2009.

Evaluating User Experience with Respect to User Expectations in Brain-Computer Interface Games

H. Gürkök[1], G. Hakvoort[1], M. Poel[1]

[1]Human Media Interaction, University of Twente, Enschede, The Netherlands

h.gurkok@utwente.nl, gido.hakvoort@gmail.com, m.poel@utwente.nl

Abstract

Evaluating user experience (UX) with respect to previous experiences can provide insight into whether a product can positively affect a user's opinion about a technology. If it can, then we can say that the product provides a positive UX. In this paper we propose a method to assess the UX in BCI systems with respect to user expectations. We demonstrate the application of our method in a preliminary study. The study results showed that BCI game control was natural and enjoyable despite the low reliability of the BCI. However SSVEP based selection induced fatigue on participants. The proposed method can extract the right and wrong practices to employ BCI in applications and can suggest interaction paradigms or considerations to follow when developing BCI systems.

1 Introduction

Evaluation of user experience (UX) in brain-computer interface (BCI) systems is a relatively new practice since much of the attention has been directed to optimising performance. Nevertheless, some approaches adopted from the human-computer interaction domain have been suggested for UX evaluation in BCI systems [1]. Among the suggested possibilities, subjective evaluation through questionnaires and interviews would seem to be a promising method. The questionnaires are especially easy and comfortable to apply, suitable for extracting statistical analyses quickly, strong and reliable once validated, and applicable to the majority of BCI system users. Questionnaires make it possible to infer UX for the use of a product but they are not powerful enough to shed light on the mechanism which builds up UX.

UX is influenced by users' values, abilities, prior experiences and knowledge as well as the context of use. Every experience a user has with a product affects not only their next experience with the product but their future experience with any product using the same technology as control input. So, a product can change the conceptions or conclusions about a technology. If the change is in a positive way, then we can say that the product provides a positive UX. BCI technology is no exception for this causal chain of UX. For example, a disabled user who has previous experience with a motor imagery based speller might have an opinion that BCI control is slow but their opinion can change positively when they use a P300 speller. However, due to the stimulation in the P300 speller, their opinion can also change that BCI control is tiresome. To give another example, for a non-disabled person, motor imagery based wheelchair control might seem to be difficult and inefficient. But when they play a motor imagery based car racing game they might consider the difficulty of BCI as a challenge they enjoy tackling. The key to success is to find the right interaction design which enables the technology to satisfy users' needs.

As we explained above, evaluating UX with respect to previous experiences can provide insight into whether a product positively affects a user's opinion about a technology and thus affords positive UX. In this paper we propose a method to gather user opinions (i.e. expectations) before using a BCI system and UX after using it. We analyse subjective UX ratings baselined to user expectations to explore whether the system provided a positive, neutral or negative UX. To demonstrate the application of the method, we use it to asses the UX in a BCI computer game.

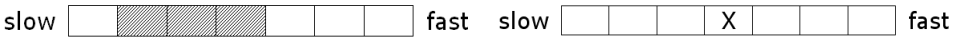
2 Method

2.1 The Initial Method

The rationale of our method is based on the SUXES [2] method which can be used to evaluate UX in service-oriented multimodal systems. With the SUXES method, through an expectations questionnaire completed before a product is used, the acceptable and desirable levels for several dimensions of UX are collected from the user and a zone of expectations (ZoE) is identified. After the product is used, through an experiences questionnaire, the experienced level is gathered for each dimension. Then it can be determined whether the experiences are within the ZoE. We conducted a pilot study with 6 participants, applying the SUXES as is in order to test its appropriateness for BCI games. We identified two problems. Firstly, despite our best efforts in perfecting the written instructions, all the participants experienced difficulty in figuring out the mechanism of the expectations questionnaire. Even after additional verbal explanation, some participants still filled in the questionnaire incorrectly (e.g. the desired level for an item was sometimes lower than the acceptable level). Secondly, we noticed that the questionnaire items were not fully fitting for BCI systems and particularly games. For example the questionnaires asked about the usefulness of the system while in a game usefulness is not a major concern. On the other hand the items lacked dimensions such as fatigue and fun which are important for entertaining BCI games.

2.2 Adaptation of the Questionnaire

In order to address the first issue we mentioned in the previous subsection, we decided to reduce the two-column expectations questionnaire design (i.e. one column for the acceptable and another for the desirable level) to a single-column one. Therefore, in our questionnaire, we use a 7-point semantic differential scale which is anchored by opposite phrase pairs at the ends (see Figure 1). The participants can then indicate their ZoE for each item by shading the box scale. Other than the phrase pairs, the scale contains no additional anchoring. We expect that if in the experiences questionnaire the user marks their experience for an item lower than the ZoE they would have negative UX and if they mark it higher they would have positive UX.



(a) Expectations with respect to speed. (b) Experience with respect to speed.

Figure 1: Interpreting expectations and experiences: (a) implies that the user will be surprised if the interface is faster than level 4 and will be disappointed if it is slower than level 2. So the zone of expectations (ZoE) is $< 2, 4 >$ (b) indicates that the user rated the speed as level 4 which is within the ZoE thus meets the expectations.

To overcome the second issue, based on our experience with BCI systems and particularly with games, we chose to include the following items and corresponding phrase pairs in parentheses (both in the given order), in our questionnaires: speed (slow–fast), pleasantness (pleasant–unpleasant), accuracy (erroneous–error-free), fatigue (tiring–effortless), learnability (easy to learn–hard to learn), naturalness (natural–unnatural) and enjoyability (boring–fun). Ordering of the phrases was consistent in expectations and experiences questionnaires.

3 Application of the Method

3.1 The Game: Mind the Sheep!

Mind the Sheep! (see Figure 2) is a multimodal computer game where the player needs to herd a flock of sheep across a field by commanding a group of dogs. The game world contains three dogs, ten sheep, a pen and some obstacles. The aim is to pen all the sheep as quickly as possible.

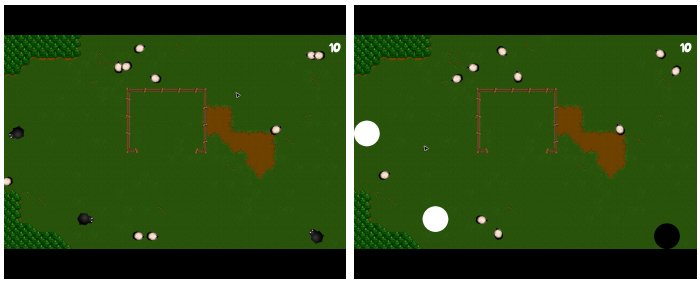


Figure 2: Screenshots from the game. On the left the SSVEP stimulation is off and on the right the stimulation is on.

To command a dog, the player positions the cursor at the point to which the dog is supposed to move and they hold the mouse button pressed. Meanwhile, the dog images are replaced by circles flickering at different frequencies. The player concentrates on the circle replacing the dog they wish to select (so as to generate a steady-state visually evoked response, shortly an SSVEP). The stimulation persists and electroencephalograph (EEG) data is accumulated as long as the mouse button is held. When the user releases the mouse button, the signal is analysed and a dog is selected based on this analysis. The selected dog immediately moves to the location where the cursor was located at the time of mouse button release.

3.2 Experimental Procedure and Tasks

Fourteen people (2 female) participated in the experiment. They had an average age of 24.5 ($\sigma = 2.88$), ranging from 19 to 28 years. Four of them had previous experience with BCIs, not with Mind the Sheep!. Another group of four indicated that they played games more than 5 hours per week.

Participants sat on a comfortable chair approximately 60 cm away from a 20″ screen with a resolution of 1280×960. They read the instructions to play the game while the experimenter was mounting the EEG. Before the game, based on their current knowledge and previous experiences, they filled in the expectations questionnaire to indicate their ZoE for selecting dogs using BCI. They were instructed to shade any number of boxes (between 1 and 7) they wished to, with respect to the devices they would need to use and tasks they would need to do to select a dog. After that, the experimenter collected the questionnaire, left the room and the game began. The participants played the game until all the sheep were penned or the play time reached 10 minutes. Finally, after the game, they filled in the experiences questionnaire.

3.3 Results and Discussion

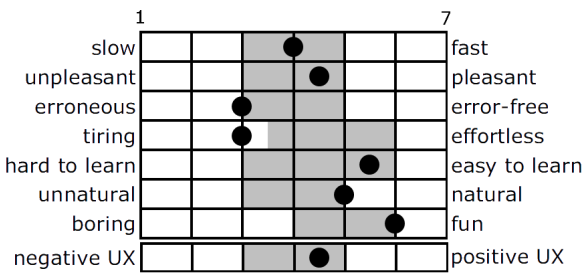


Figure 3: Median values across all participants for ZoEs (grey cells) and UXs (black circles).

For each item, we computed the medians of the experienced levels and the lowest and the

highest expected values across the participants (Figure 3). We also computed the medians across all the items (the last row in the figure) as the indicator of overall UX. Experience levels for all items, except for fatigue, and for overall UX were within the ZoE meaning that the game provided a neutral UX. An item level analysis can provide better insight into what was right and wrong in the design of the game or the interaction. Especially the items rated beyond or on the border of the ZoE are of interest. Ratings for accuracy show that the SSVEP detection mechanism was not reliable. Despite this, experience levels for naturalness and enjoyment were on the positive end of the ZoE meaning that participants found selecting dogs by concentration intuitive and they enjoyed the gameplay. However, as the ratings for fatigue imply, they were tired by the SSVEP stimulation. These findings collectively mean that participants enjoyed playing the game using a BCI but perhaps another type of BCI, such as a stimulus independent one, could have been a better alternative.

Another research direction is to cluster participants according to their previous experience with BCIs and explore how the expectations and experiences differ between the participant groups. We performed this analysis but did not find any difference between experienced and naive BCI users. Perhaps, with a well-distributed and larger sample space interesting results can be achieved.

4 Conclusion

UX is influenced by users' values, abilities, prior experiences and knowledge as well as the context of use. Evaluating UX with respect to previous experiences can provide insight into whether a product positively affects user's opinion about a technology and thus affords positive UX. In this paper we proposed a method to assess the UX in BCI systems with respect to user expectations and demonstrated its application in a preliminary study. The study results showed that BCI game control was natural and enjoyable despite the low reliability of the BCI. However SSVEP based selection induced more than expected fatigue in participants. No difference was found in UX evaluations of naive and experienced BCI users although this might be due to the unbalanced and small sample space.

The proposed method can be used to determine whether a BCI system improves upon users' previous experiences with BCIs or not. It can extract the right and wrong practices to employ BCI in applications and can suggest interaction paradigms or considerations to follow when developing BCI systems. It can also be used to study how and why UX changes for different user groups. The proposed questionnaires can be modified for use with for general BCI systems, following the method we have described. Then the method can be used for a broad range of applications (e.g. assistive, recreational BCIs) and users (e.g. patients, gamers).

Acknowledgments

The authors gratefully acknowledge the support of the BrainGain Smart Mix Programme of the Netherlands Ministry of Economic Affairs and the Netherlands Ministry of Education, Culture and Science. They also thank A. Minuto, B. van Dijk, L. Packwood and two anonymous reviewers for their help in improving the article.

References

[1] B. van de Laar, H. Gürkök, D. Plass-Oude Bos, F. Nijboer, and A. Nijholt. Perspectives on user experience evaluation of brain-computer interfaces. In *Proceedings of the 14th International Conference on Human-Computer Interaction*. Springer, 2011. To appear.

[2] M. Turunen, J. Hakulinen, A. Melto, T. Heimonen, T. Laivo, and J. Hella. SUXES - user experience evaluation method for spoken and multimodal interaction. In *Proceedings of INTERSPEECH 2009*, pages 2567–2570. ISCA, 2009.

Are We ready? Issues in Transferring BCI Technology from Experts to Users

R. Leeb[1], A. Al-Khodairy[2], A. Biasiucci[1], S. Perdikis[1], M. Tavella[1], L. Tonin[1], T. Carlson[1], J. d. R. Millán[1]

[1]Chair in Non-Invasive Brain-Machine Interface, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland
[2]Clinique Romande de Réadaptation-Suvacare, Sion, Switzerland

robert.leeb@epfl.ch

Abstract

Brain-Computer Interfaces (BCIs) are no longer only used by healthy subjects under control conditions in laboratory environments, but by patients controlling applications at their homes, without the BCI experts around. But are the technology and the field mature enough for this? In this work, we want to summarize the experiences gained and the lessons we learned while transferring BCI technologies from the lab to the user's home. These lessons range from pure BCI issues (technical and handling), to common communication problems between different people involved, and lessons encountered while controlling the applications. The points raised are very general and will be faced similarly by other groups, if they move on to bring the BCI technology to the end-user.

1 Introduction

The idea and the technology enabling Brain-Computer Interfaces (BCIs) to control machines, not by manual operation, but by mere "thinking" have been around for quite a long time. Most often the electrical activity is recorded from the brain by the electroencephalogram (EEG). Control parameters are extracted from this activity, which can be used by disabled people to establish a new communication channel between the human brain and a machine.

Several BCI prototypes have been demonstrated over the last decade [1] for applications such as (i) communication and control, such as writing on a virtual keyboard or browsing the Internet, (ii) the control of wheelchairs or robots, and (iii) computer games for healthy users. Although most of the prototypes and applications target disabled users, the vast majority of the published work is based on the analysis of data from healthy subjects. Recently the community has started to address this issue and more studies based on patient data are now available. Nevertheless, most of this work is done by BCI experts who either host patients at their research labs or go to the patients' home. Yet it is crucial for the field that we cross another frontier, by letting caregivers or therapists around the patients use the BCI themselves without (or with a minimum) interference from the BCI experts. A third possibility is to install a remote support platform that allows the BCI expert to remotely assist, while non-BCI experts are physically working at the patient's place. Such an approach has already been presented as "telemonitoring" in [2] as a method to remotely train people to achieve BCI control.

In this work, we want to report our experiences, and the problems we encountered, while transferring our BCI from the lab to the clinics and to the patient's home. We started with naive patients who first performed BCI training and then used their BCI to evaluate several BCI prototypes (either a writing application for communication or a robotic telepresence platform). Rather than present a technical description of the BCI setups and the remote support infrastructure, we will focus on the experiences and challenges that we encountered.

As mentioned above, our plan was for caregivers to undertake the whole job of BCI setup and operation, while the BCI experts provide troubleshooting advice (if needed at all), via a telephone support hot-line. To accomplish this we hid the complexity of BCI operation as much as possible.

In addition, we setup a remote support infrastructure: on one hand for exchanging data and settings, and on the other hand, for troubleshooting. This remote support infrastructure allows us to assist more than one patient at the same time.

2 Methods

In the presented studies, a BCI based on motor imagery was used, but the general strategies would be the same for other BCI paradigms. In our case, the participants had to start by imagining left hand, right hand and foot movements during calibration recordings. Afterwards the two best discriminable classes were used for a 2-class control task in case of the online experiments and with the applications.

2.1 Brain-Computer Interface Details

The brain activity was acquired via 16 EEG channels over the motor cortex. From the Laplacian filtered EEG the power spectral density was calculated. Canonical variate analysis was used to select subject-specific features that are stable over time and maximized the separability between the different tasks. These features were used to train a Gaussian classifier. Furthermore, decisions with low confidence on the probability distribution were filtered out and the temporary evidence about the executed task was accumulated until a probabilistic threshold is reached (for more detail see [3]). The task of the participant was to control a liquid cursor, either to the left or right. A number of graphical user interfaces (GUIs) and scripts were built around our BCI to hide all the complexity and to give to the operators only the necessary functionality without overloading them. Furthermore, the feature selection and classifier training were done remotely by the BCI expert, so no complex task has to be performed at the patients place.

2.2 Application Prototype Details

Since the goal of the BCI is not to control a cursor to the right or to the left on a computer screen, but is to be used by the patient to interact with the surrounding world, we tested two different applications. In the first application a commercially available assistive technology text entry system called QualiWorld (QualiLife, CH) with a standard auto-scanning mode was used to hierarchically select a letter of the alphabet. The subject used one BCI command to select the currently highlighted item (block, row, letter). The second BCI command was used to switch the scanning order. On the screen, the matrix with the alphabet and the BCI output was displayed. The task of the subject was to write pre-defined simple words (copy-spell). The second application was a telepresence platform based on the Robotino robot (Festo, CH). The subject controlled remotely the robot turning to the left or to the right to reach several targets within an office environment. Importantly, subjects had never been in such an environment. The default behavior of the robot was to move forward and to avoid obstacles with the help of shared control and its on-board sensors. The subjects saw a video-transmission from an on-board camera of the robot in parallel with the BCI output. Both applications were quite demanding for the subject, since a proper BCI performance is necessary to perform the requested BCI action on time while the subject has to focus their visual attention on the application to achieve the given task.

2.3 Remote Support Infrastructure

Following the requirements in [2], we set-up an infrastructure with state-of-the-art technologies. A synchronized data folder allowed an automatic transfer (via Unison) of the recorded files from the patient to the BCI experts and of classifiers or configuration files back to them. Communication via Skype (chat, speech or video) was possible to give instructions to the caregiver or patient. Since sometimes the support cannot help in overcoming some (mostly technical) errors with only verbal instructions, a remote takeover of the laptops was also possible. This was done via Remote

Desktop under Windows and SSH under Linux. Finally OpenVPN was used to remotely access laptops or to share certain resources (i.e. robot), even in environments with limited connectivity.

3 Results

Up to now the number of subjects trained at the various out-of-the-lab locations (either at clinics, AT support centers or users home) is very small, so no quantitative numbers will be given here. Six subjects (2 female, 4 male; 42±13 years) have participated once every 1–2 weeks for up to 3 hours/day. The BCI performance achieved by participants trained remotely is in the same range as with on-site training or training with healthy subjects. Two subjects had very good control and could test the applications, all others didn't reach that level within 10 sessions (maximum possible amount). Two subjects had to be excluded because of inherent muscular artifacts coming from their impairments. We have to mention that the subjects were mostly impaired by different levels of myopathy (4), spinal cord injury (1) and spino-cerebellar ataxia (1). Interestingly, the results with the writing applications were slightly worse than expected, whereas with the telepresence robot the patient achieved a better performance than all healthy participants (for more details see [4, 5]). The mood and motivation of the participants were conducted via questionnaires [5].

Lessons from BCI Usage

- The data, classifier and configuration transfer works very well, as long as Internet is available.
- The remote support infrastructure can solve all problems.
- The current version of our BCI is not user-friendly enough. For non-BCI experts it is still too complex and too many cables had to be connected between the 2 laptops, amplifier...
- It would be helpful if the caregiver/therapist has some technical understanding about the BCI (suggested in [2]), which could help in adjusting on-site simple parameters, e.g. thresholds.
- BCI experts and therapists do not have the same background knowledge and have a different (technical) vocabulary. Many problems are triggered because of simple misunderstandings.
- This case is even stronger if the communication is not done in their mother tongues.
- The instructions to the patients have to be given in his/her mother tongue.
- Mounting the electrodes by non-experts takes too much time (up to 1.5 h) and contains too many sources of error (floating electrodes, bad impedances, misplaced cap).
- The training phase could be made more engaging and should provide more fun for the user.
- Highest motivation is achieved, if the patient is seeing a personal future need in the BCI.

Lessons from Aplication Experiments

- Having a good BCI control does not guarantee good control over the application, because here a split attention between application and BCI is required.
- Generally BCI training does not require users to achieve 100 % classification accuracy, but most applications demand almost perfect performance. The impact of an error is critical in applications since one wrong decision needs a series of correct ones to overcome/correct the error, what imposes heavy demands on users. A better way to handle wrong decisions is required, by means of either an easy "undo" possibility or smarter application designs.
- Subjects cannot deliver all BCI classes with the same easiness. Sometimes a bias towards one of the classes exists which results in a strong performance deterioration.
- Most BCI trainings are intended to improve the intentional control performance, but for most applications intentional-non control (INC) is more important - which is not trained per se. By INC we mean the periods in which the participant is not wanting to deliver any command, e.g. the robot is moving forward, or waiting for the next text selection.
- Therefore, most applications are using the BCI incorrectly, because they are forcing long "waiting" periods, during which the subjects deliver many false positives. These "unsuccessful" periods result in frustration/stress that degrades the subject's overall performance

and leads to short periods of very poor performances.
- Patients mentioned that a "pause" mode would be beneficial, otherwise it is too tiring.
- Shared control helps the user to perform better and makes it less demanding [4].

4 Discussion and Conclusion

Besides the positive experiences and the promising results we have gained with the patients, we have to acknowledge that a lot of work is still needed. Although we tried to hide the complexity of the BCI and of the prototype applications, our system is not ready to be used alone at the patient's place (as it is the case for most other BCI systems, especially motor imagery ones). This raises the question: How mature does BCI technology have to be before we can give it away?

Furthermore, the difference between the outcome of a successful BCI training (intentional control) and the needs of the application (non-intentional control) became obvious. Therefore, we actually suggest a change of the approach and identified some possibilities: (i) Include INC in the training process. (ii) Since a normal 2-class BCI is sometimes biased towards one class, we could exploit this natural bias. For instance, we could use the "hard" task for key commands (e.g., for selection). (iii) Design an "active" BCI: in the case of text entry or web browsing, the user makes the scan progress forward or backward by delivering mental commands. To select, the user stays in the INC state (which means that he does not deliver any commands) for a short period of time (akin to dwelling in eye-tracking). (iv) Design of hybrid BCIs where key commands (e.g., error corrections, pause, selection) are delivered through other channels such as EMG.

To conclude, we want to mention explicitly that this paper aims to report the lessons learned, not possible mistakes in the operation of the BCIs which are natural as for any other new advanced technology requiring time to master it. Consequently, any limitation in the use of BCI technology remains mainly on our shoulders as researchers and developers, not on the users and caregivers. We just want to use this opportunity to wake up the community to these important issues and share our, sometimes, frustrating experiences.

Acknowledgments

This work is supported by the European ICT programme projects TOBI: Tools for Brain-Computer Interaction (FP7-224631). 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. The authors would like to thank all patients for their participation and all therapist and caregivers for their work.

References

[1] J. d. R. Millán, R. Rupp, G. R. Müller-Putz, R. Murray-Smith, C. Giugliemma, M. Tangermann, C. Vidaurre, F. Cincotti, A. Kübler, R. Leeb, C. Neuper, K. R. Müller, and D. Mattia. Combining brain-computer interfaces and assistive technologies: state-of-the-art and challenges. *Front Neurosci*, 4:161, 2010.

[2] G. R. Müller, C. Neuper, and G. Pfurtscheller. Implementation of a telemonitoring system for the control of an EEG-based brain-computer interface. *IEEE Trans Neural Syst Rehabil Eng*, 11:54–59, 2003.

[3] F. Galán, M. Nuttin, E. Lew, P. W. Ferrez, G. Vanacker, J. Philips, J. del R. Millán. A brain-actuated wheelchair: asynchronous and non-invasive brain-computer interfaces for continuous control of robots. *Clin Neurophysiol*, 119(9):2159–2169, 2008.

[4] L. Tonin, et al. Brain-controlled telepresence robot by motor-disabled people. *Proc 32nd Annual Int Conf IEEE Eng Med Biol Soc*, 2011.

[5] TOBI deliverable D10.5: "Evaluation Report I". *www.tobi-project.org*, 2010.

INDEX

VISTA-Medical
Medizintechnik - GmbH

BRAIN PRODUCTS
Solutions for neurophysiological research

BCI Package

The smart solution for your BCI research

Brain Products has put together the ultimate BCI Package: A handy system perfectly matched to the needs of BCI researchers. Your work will be faster, easier and more productive!

- a 16-channel V-Amp EEG amplifier system compatible with all major software for BCI applications (e.g. BCI2000, OpenViBE and BCI2VR)
- a set of 18 active electrodes (ground and reference channels included)
- ImpBox to easily determine the impedance values of the active electrodes
- 2 softcaps with 32 possible electrode positions
- a Starter Set
- a Software Package
  - › BrainVision Analyzer 2 interacts easily with MATLAB and BCI2000
  - › BrainVision Recorder with a Remote Data Access module
  - › Basic Module of BrainVision RecView with Online Signal tool for real time analysis
  - › Software Development Kit
- an add-on cd with samples of BCI applications originated by Brain Products

And all this at a truly unbeatable price! Just ask your local distributor for the Brain Products BCI Package.

bci package
brain computer interface

www.brainproducts.com


g.tec
LEADING
BCI TECHNOLOGY
FOR YOUR LAB
www.gtec.at

enobio

BRAINWAVES MADE SIMPLE

- **Wireless communication**
  allows for mobile data recording away from lab or controlled environments
- **Active electrodes**
  to reduce environmental noise
- **Modular design**
  allows users to customize a single system for various applications
- **Use wet or dry**
- **Easy Set-up**
  Comfortable full head system
- **Flexible electrode placement**
  Based on the standard 10-10 International system

Number of channels:
4

Output:
Simple ASCII or EDF compatible data file and/or digital streaming over TCP/IP

Brandwidth:
0 to 125 HZ

Wireless communication:
IEEE802.15.4

Dimensions:
66x 55 x 25 mm

Weight:
65 g

Operating time:
8 hours

Input:
Biopotentials (EEG, ECG, EOG)

Sampling:
250 S/s

Effective dynamic range:
22 bits

DC Offset:
Automatic offset compensation for each channel

Common mode rejection Ratio: 96 dB

NE neuroelectrics®

For more info contact:
Neuroelectrics, Teodor Roviralta 45
08022 Barcelona - Spain
Tel.+34 93 254 03 66
www.enobio.com