The 2009 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 11-15, 2009 St. Louis, USA

A Brain Computer Interface Methodology Based on a Visual P300 Paradigm

Gabriel Pires and Urbano Nunes

Abstract— **Brain Computer Interface (BCI) systems based on electroencephalography (EEG) open a new communication channel for people with severe motor disabilities, without recurring to the conventional motor output pathways. The very low signal-to-noise ratio and low spatial resolution still limits severely BCIs communication bandwidth. This paper presents the ongoing work toward the development of a BCI system for wheelchair steering. A full system based on a visual P300 oddball paradigm is proposed. The signal processing algorithms are computationally efficient and require a short phase training. Temporal features and EEG channels are selected through a Fisher criteria. For enhancement of signal-to-noise ratio and data dimensionality reduction, a spatial filter named Common Spatial Patterns is applied. This method is widely used for classification of motor imagery events, however it is not very often used for classification of event related potentials such as P300. In this paper we show that Common Spatial Patterns is an effective approach to improve P300 classification rates. In our approach, the input features for classification are the projections of the filtered data instead of the variance of the projections as typically used in motor imagery. Offline classification results, obtained with a Bayesian classifier, are presented showing the effectiveness of the overall methodology.**

I. Introduction

For people suffering from severe motor disabilities such as amyotrophic lateral sclerosis and locked-in syndrome, and certain types of cerebral palsy, Brain Computer Interfaces (BCI) emerge as a feasible type of human-computer and human-machine interfaces that can allow these patients to interact with the world. Standard interfaces such as language processing, eye tracking and head or teeth switches are not suitable for people with total lack of motor movements or with very low motor dexterity and with unperceivable language.

Current non-invasive BCI systems based on electroencephalographic (EEG) data are divided in three main classes according to the type of neuromechanisms: 1) event related synchronization and desynchronization (ERD/ERS) of sensorimotor rhythms μ (8-12 Hz) and β (18-25 Hz). This rhythms typically decrease (ERD) during motor imagery and increase (ERS) during motor relaxation [1], [2]; 2) P300 peak elicited by a visual oddball paradigm [3], [4], [5]; and 3) steady-state visual evoked potentials (SSVEP) elicited by a constant flicker at a given frequency [6]. These approaches, already tested in our Lab, have quite different characteristics presenting weak and strong points with relevant practical implementation issues, as it will be described in the following.

The first approach requires that the subjects learn to control their brain rhythms. This is often a long and difficult task and it can happen that users are unable to learn how to control them. Control of μ and β rhythms is usually reached through mental tasks such as motor imagery, for instance, imagining that a left hand task is being performed. After some training with visual feedback, users usually can create and refine their own mental mechanisms. Knowing the map of the motor cortex (motor homunculus) it is possible to select different motor tasks with known spatial distribution so that different motor cortex areas be activated. Motor imagery requires a high degree of concentration and some mental effort. The number of discriminative patterns is usually limited due to the low spatial resolution of EEG. The number of classes proposed in current research works almost never goes beyond four classes. See for example the work presented in [7] where the imagination of left hand, right hand, foot and tongue tasks were used to discriminate four different patterns.

The second approach is based on the P300 neuromechanism which is a peak that typically occurs 300 ms after an expected, but infrequent, random event occurs. Each stimulus event corresponds to a symbol/picture with a particular meaning for the interface (e.g. letters, high level commands). The stimulus must be perceptible on the user field of view without gazing the specific stimulus. One major disadvantage of P300 arrives from the fact that the user has to wait for the occurrence of the desired (target) stimulus which randomly appears. It is not the user who decides when to provide an intention but rather the emergence of the stimulus. Moreover, processing algorithms have to run synchronously with the start of the stimuli. In terms of machine learning, the oddball paradigm reduces a n-symbol detection to a 2-class discrimination problem, i.e., the discrimination between target events (desired command - one of the n symbols) and non-target events (remaining $(n-1)$ symbols). This way, several user intentions correspond to a unique brain pattern (P300 peak signal), representing a high volume of information. However, increasing the number of possible commands (events) decreases the transfer rate because each stimuli is flashed less frequently.

In the last BCI approach, as a response to a stimulus flickering at a constant frequency, a signal of the same

This work has been in part supported by Fundação para a Ciência e Tecnologia (FCT), under Grant PTDC/EEA-ACR/72226/2006. Gabriel Pires would like also to thank the support of FCT through the research fellowship SFRH/BD/29610/2006.

Gabriel Pires is with the Institute for Systems and Robotics, University of Coimbra, Portugal, and with the Department of Electrical Engineering, Institute Polytechnic of Tomar, Portugal gpires@isr.uc.pt

Urbano Nunes is with the Institute for Systems and Robotics, and with the Department of Electrical and Computer Engineering, University of Coimbra, Portugal urbano@isr.uc.pt

TABLE I

NEUROMECHANISMS APPROACHES SUMMARY

Sensorimotor ryhthms	P300	SSVEP
- Some subjects require long period of training - Some subjects are unable to control their rhythms	- No training required - Almost all subjects have P300 reaction	- No training required - Almost all subjects have SSVEP reaction
- Requires subject focus and abstraction - Can be synchronous (based on cues) or asynchronous	- Requires subject to be attentive to stimuli - Fatigue due to sitmuli focus - Needs synchronism with visual stimuli	- No special requirements - Fatigue due to stimuli flickering - No synchronism requirements
- Needs visual (or other) feedback information - For wheelchair steering, the feedback is the wheelchair pose	- The screen has to be in the field of view (no eye movement required) - For wheelchair steering the screen can disturb the environment perception	- The user has to gaze the respective stimulus. This involves eye movement - For wheelchair steering the screen can disturb the environment perception
- Small number of decoded patterns (3-4 patterns)	- Two class approach allows the codification of a large amount of information (number of commands is inversely proportional to transfer rate)	- Each discriminated frequency is a possible decoded information
IDEAL BCI	PRACTICAL BCI	NOT A TRUE BCI

frequency (SSVEP) arises at the occipital brain region (visual cortex). The user has to gaze the stimuli positioned in some part of the screen which involves the movement of the eyes. Because it depends on the brain's normal output pathway of peripheral nerves and eye muscles it can not be called a true BCI. Notwithstanding this, the interface can be suitable for people with severe motor disabilities but still able to perform small eye movements.

An advantage of motor imagery over the two other event related potential approaches is that it does not depend on visual focus or gazing. Moreover, it can be used without synchronism cues (except during the training phase). The P300 and SSVEP rely on visual stimuli which cause fatigue due to continuous user focus and flickering, respectively. If these visual based approaches are for instance to be used to steer a wheelchair, the focus on the stimuli display can affect the perception of the surrounding environment. See Table I for a summary comparing the main features of these three approaches.

In this paper we present a BCI system that is being developed as a human machine interface suitable to steer a wheelchair [8]. The user should be able to provide sparse commands indicating low-level steering directions or high-level commands or tasks, being supported at the same time by an intelligent navigation module for a safe navigation. Navigation issues are however beyond the scope of this paper. The same goal is being pursued by Millan's research group with interesting results. In their work [9], a motor imagery based BCI is used to discriminate 3 different commands to steer a wheelchair with navigation assistance in indoor environment.

The proposed BCI system is based on a visual P300 paradigm. Balancing the three BCI approaches, P300 presents two appealing features, the first one is the possibility of discriminating a large number of commands through a 2-class classifier, and the second one is that almost all users react to the P300 oddball paradigm without training. A BCI system should present the following features:

1) Pre-processing and normalization steps for robustness to inter-trial variability;

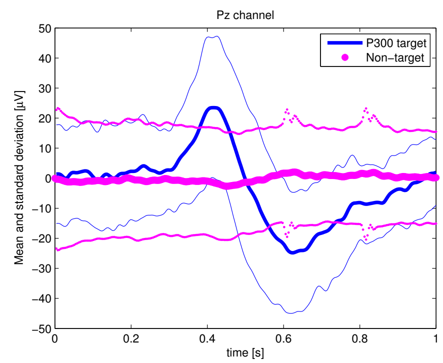

Fig. 1. Grand average of raw P300 target signal (80 epochs) and non-target signal (1120 epochs). The thick lines represent the mean, and thin lines represent the mean plus and minus the standard deviation.

2) Feature selection and channel selection. These are critical issues that can lead to data reduction dimension needed to increase classification performance.
3) Computational efficient methods for real-time application;
4) Short time training phase. A training phase for each subject and for each session is always needed because there is a inter-session variability and high inter-subject variability.

The P300 signal presents a very low signal-to-noise ratio (SNR) as can be seen in Fig. 1, evidenced by the overlapping of the averages and standard deviations of the P300 signal (target) vs. standard signal (non-target). To reach a single trial classification, robust signal processing and learning algorithms have to be applied. This paper proposes a methodology that combines a feature selection, spatial filtering, and Bayesian classification scheme that aims the improvement of single trial classification. The relevance of this paper to the robotics community is two-folded: 1) The conceptual architecture, and signal processing and classification techniques presented in this paper are not limited to BCI applications, but also to other human-machine interfaces/interaction in general, and for signal processing of sensor data used in

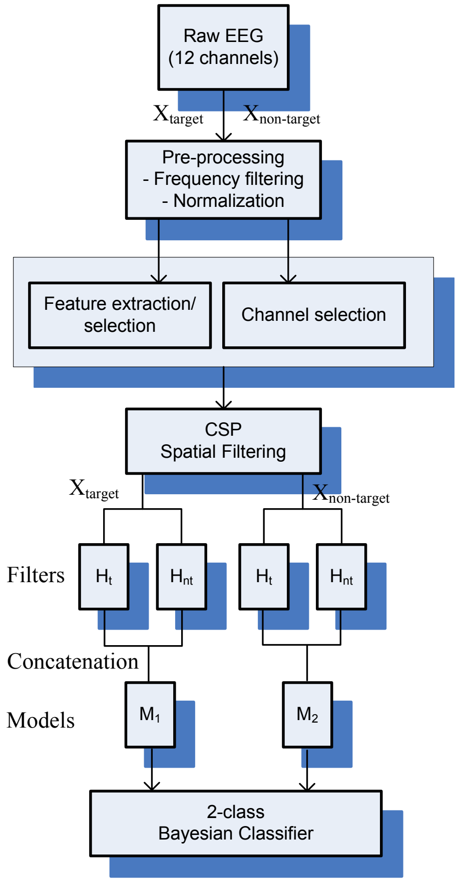

Fig. 2. Overall BCI system.

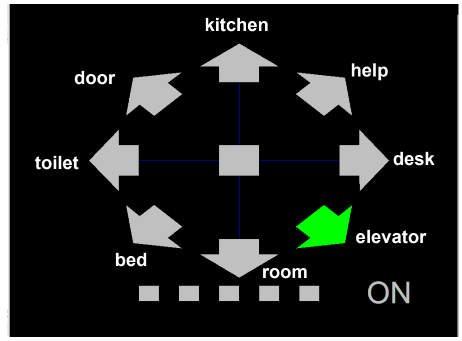

Fig. 3. P300 arrow paradigm. Each symbol is flashed during 100 ms and the time between flashes is 200 ms. Each arrow can be associated to a high level command.

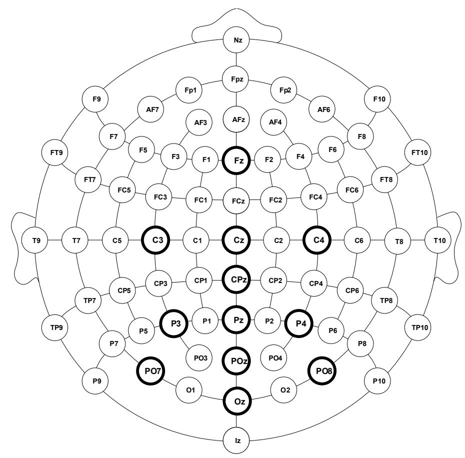

Fig. 4. Electrodes of the extended 10-20 international standard system. The EEG electrodes used for data acquisition are marked with a solid circle.

the robotic fields; 2) BCIs are reaching real applications in the rehabilitation field, which requires from robotics the development of suitable devices for BCI integration in robotic systems, in aspects that goes from ergonomics and haptic devices to specific navigation assistance.

II. BCI System Based on P300 Paradigm

The overall training BCI system is described in Fig. 2. Data are collected and separated according to target and non-target events. A pre-processing module is used to filter and normalize data. Then, best features and channels are selected. The CSP spatial filtering increases the separability between target and non-target classes. CSP projected data is used to create models to be used on a Bayesian classifier. The classifiers are user-dependent. Each block of the system is described in the following sections.

A. *P300 visual stimuli paradigm*

The paradigm is shown in Fig. 3 and was already presented in [10]. It is composed by 8 direction arrows, a stop square, a ON/OFF switch and 5 small squares. The paradigm was specifically designed to steer a robotic device though the detection of the desired steering direction arrow. Each arrow and square can be associated to any other symbol. In the example illustrated in Fig. 3 they were associated to high level commands such as room, toilet, desk, door, etc. Each symbol is randomly flashed with a uniform distribution, therefore the target event occurs only once on each 15 flashes, providing this way an oddball paradigm (each round of 15 flashes is called a trial). The subject is asked to focus its desired command and to mentally count the number of target flashes.

B. *EEG acquisition and pre-processing*

The dataset experiments were collected from our previous work [11]. Three healthy subjects participated at the experiments. The subjects were seated in front of a computer screen at about 60 cm. The EEG activity was recorded from 12 Ag/Cl electrodes at positions Fz, Cz, C3, C4, CPz, Pz, P3, P4, PO7, PO8, POz and Oz according to the internacional extended 10-20 standard system using a g.tec cap (Fig. 4). The electrodes were referenced to the right mastoid and the ground was placed at AFz. The EEG channels were amplified with a gUSBamp (g.tec, Inc.) amplifier, bandpass filtered at 0.1-30 Hz and notch filtered at 50 Hz and sampled at 256 Hz. All electrodes were kept with impedances under $5K\Omega$.

A training phase session occurs before the testing session. Usually the training consists on approximately 80 target epochs and 1120 non-target epochs, which takes about 4 minutes (the epochs are overlapped in time). Each epoch has a duration of 1 second and is synchronized with the start

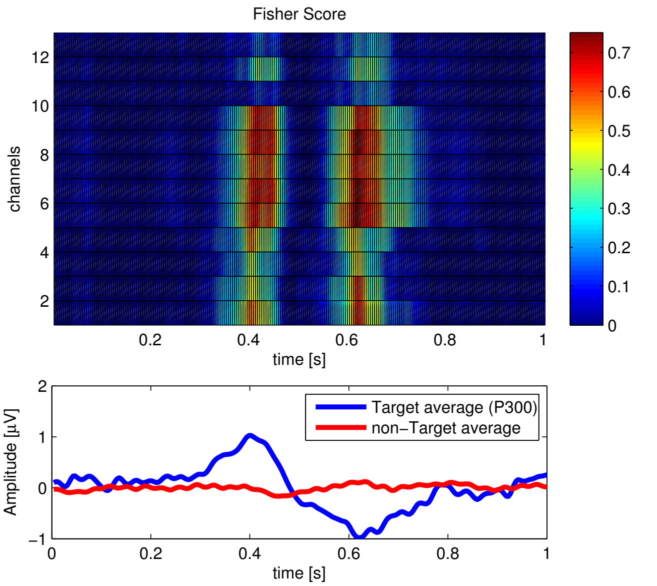

Fig. 5. Top: color map indicating level of discrimination between target and non-target classes (channels are ordered according to Fz, Cz, C3, C4, CPz, Pz, P3, P4, PO7, PO8, POz and Oz); bottom: Average of target events (P300) and non-target events for Fz channel.

of the event stimulus. The EEG signal is low-pass filtered by a 4th order Butterworth filter with 7 Hz cut frequency. Each epoch is normalized to zero mean and unit standard deviation.

C. Feature and Channel Selection - Fisher Criterion

The P300 features correspond to the amplitudes of the P300 temporal pattern. The Fisher criterion (FC) [12] provides the normalized level of discrimination between target and non-target classes. It is applied to each instant time of the epoch-window using all training trials. Therefore it is possible to select the temporal features (top ranked) and the best channels presenting higher discrimination. The Fisher score FS_j of the feature j (instant time) is given by

$$FS_j(\mathbf{X}) = \frac{(\mu_j(\mathbf{X_t}) - \mu_j(\mathbf{X_{nt}}))^2}{\sigma_j^2(\mathbf{X_t}) + \sigma_j^2(\mathbf{X_{nt}})} \quad (1)$$

where $\mu_j(\mathbf{X_t})$ and $\mu_j(\mathbf{X_{nt}})$ are respectively the target and non-target averages over all training trials for each feature j, and $\sigma_j^2(\mathbf{X_t})$ and $\sigma_j^2(\mathbf{X_{nt}})$ are respectively the target and non-target variances for each feature j. The top of Fig. 5 shows a color map representing the Fisher score. High level colors denote high levels of discrimination. The bottom figure presents the averages of target (P300 resulting from oddball paradigm) and non-target classes. For that particular subject and session, the best discrimination occurs around 0.4 and 0.6 s and for channels CPz, Pz, P3, P4 and PO7.

D. Common Spatial Patterns

Spatial filtering is an effective approach that already proved to improve EEG spatial resolution and SNR. One form of spatial filtering is based on EEG referencing methods, such as Common Average Reference, Small Laplacian and Large Laplacian [13]. These filters act as high-pass spatial filters that enhance local activity and decrease the distributed activity. A different spatial filter approach, more oriented for pattern classification, was proposed by Koles [14] despite it was firstly used for neurophysiologic component decomposition in clinical electroencephalography.

The CSP method is based on the simultaneous diagonalization of two real symmetric matrices proposed by Fukunaga [12]. The simultaneous diagonalization allows the decomposition of raw EEG signals into two discriminated patterns extracted from two populations (classes) simultaneously maximizing the variance of one class and minimizing the variance of the other class. Using only the projections with larger discrimination the data dimension space is reduced which improves the classification process.

This method has been successfully applied in BCI research for the two-class motor imagery discrimination [15], [16], [17]. Some variants of CSP were also already proposed for the multiclass problem [18], [19], [20]. Despite its successful application on motor imagery based BCIs, there are very few applications of CSP for the discrimination of Event Related Potentials (ERP) such as the P300. In [21], the Common Spatio-Temporal Patterns (CSTP) is proposed, which incorpores spatio-temporal covariance matrices into CSP to extract more prominent spatio-temporal patterns.

We propose here the application of a standard CSP for feature extraction followed by a Bayesian classifier based on probabilistic models of spatial filtered data. This methodology was already presented in [11] but using different methods for feature and channel selection, and with different combination of features for classification.

Consider two spatio-temporal matrices $\mathbf{X_t}$ and $\mathbf{X_{nt}}$ with dimension $N \times T$, where N is the number of channels and T is the number of samples of the time series epoch of each channel. The matrix $\mathbf{X_t}$ represents the P300 potential evoked by the target event and $\mathbf{X_{nt}}$ represents the ongoing EEG for non-target events. The CSP method is based on the principal component decomposition of the the sum covariance $\mathbf{R}$ of the target and non-target covariances

$$\mathbf{R} = \mathbf{R_{nt}} + \mathbf{R_t} \quad (2)$$

where $\mathbf{R_t}$ and $\mathbf{R_{nt}}$ are the normalized $N \times N$ spatial covariances computed from

$$\mathbf{R_t} = \frac{\mathbf{X_t X_t'}}{tr(\mathbf{X_t X_t'})} \quad \mathbf{R_{nt}} = \frac{\mathbf{X_{nt} X_{nt}'}}{tr(\mathbf{X_{nt} X_{nt}'})} \quad (3)$$

where $'$ represents the transpose operator and $tr(A)$ represents the trace of A.

The spatial filters are estimated from the overall set of trials gathered during training. Therefore it is used the average of the normalized covariances trials

$$\overline{\mathbf{R_t}} = \frac{1}{N_t} \sum_{i=1}^{N_t} \mathbf{R_t}(i) \quad \overline{\mathbf{R_{nt}}} = \frac{1}{N_{nt}} \sum_{i=1}^{N_{nt}} \mathbf{R_{nt}}(i) \quad (4)$$

where N_t and N_{nt} are the number of target and non-target trials in the training set. The averaged covariance matrix $\overline{\mathbf{R}}$

is factorized through the application of the PCA

$$\overline{\mathbf{R}} = \overline{\mathbf{R_t}} + \overline{\mathbf{R_{nt}}} = \mathbf{A}\lambda\mathbf{A}' \quad (5)$$

where A is the orthogonal matrix of eigenvectors of $\overline{\mathbf{R}}$ and λ is the diagonal matrix of eigenvalues of $\overline{\mathbf{R}}$. A whitening transformation matrix $\mathbf{W} = \lambda^{-\frac{1}{2}}\mathbf{A}'$ transforms the covariance matrix $\overline{\mathbf{R}}$ to I (identity matrix)

$$\mathbf{S} = \mathbf{W}\overline{\mathbf{R}}\mathbf{W}' = \mathbf{I}. \quad (6)$$

Applying the whitening transform to each individual class, we obtain

$$\mathbf{S_t} = \mathbf{W}\overline{\mathbf{R_t}}\mathbf{W}' \quad \mathbf{S_{nt}} = \mathbf{W}\overline{\mathbf{R_{nt}}}\mathbf{W}'. \quad (7)$$

From the above two equations it is straightforward that

$$\mathbf{S_t} + \mathbf{S_{nt}} = I \quad (8)$$

The PCA factorization of S_t and S_{nt} is performed by

$$\mathbf{S_t} = \mathbf{A_t}\lambda_\mathbf{t}\mathbf{A_t'} \quad \mathbf{S_{nt}} = \mathbf{A_{nt}}\lambda_\mathbf{nt}\mathbf{A_{nt}'} \quad (9)$$

From (8) and (9) then $\mathbf{A_t} = \mathbf{A_{nt}}$ and $\lambda_\mathbf{t} = \mathbf{I} - \lambda_\mathbf{nt}$.

Both classes share the same eigenvectors and the respective eigenvalues are reversely ordered. The eigenvector with largest eigenvalues for one class has the smallest eigenvalue for the other class and vice versa. The first and last eigenvector are optimal eigenvectors to discriminate the two classes. Defining $A^{(1)}$ and $A^{(N)}$ as the first and last eigenvectors, each with dimension $N \times 1$, the following spatial filters are designed

$$\mathbf{H_t} = \mathbf{A^{(1)}}'\mathbf{W} \quad \mathbf{H_{nt}} = \mathbf{A^{(N)}}'\mathbf{W}. \quad (10)$$

The spatial projections are then given by $\mathbf{Y} = \mathbf{HX}$ where $\mathbf{H}$ is the matrix or vector with the selected filters.

E. Bayesian Classification

The projections from the two filters are concatenated respectively for each class, creating a new vector for classification (see Fig. 2):

$$\mathbf{Z}(j) = \quad [\mathbf{Y}^{(1)}(j) \quad \mathbf{Y}^{(N)}(j)] \quad (11)$$

where j is the index trial and the superscripts represent the EEG projections using filters $\mathbf{H_t}$ and $\mathbf{H_{nt}}$ (10). The conditional density function of the class i (target, non-target) is modeled as a multivariate distribution under gaussian assumption

$$p(\mathbf{Z}|w_i) = \frac{1}{(2\pi)^{n/2}|\Sigma_i|^{1/2}} \exp(-(\mathbf{Z} - \mu_i)^T \Sigma_i^{-1} (\mathbf{Z} - \mu_i)/2) \quad (12)$$

where μ_i and Σ_i are the mean and covariance matrices computed for each class w_i. Each model takes into account the temporal structure of P300 pattern and therefore this is a different approach of typical feature representation used for μ and β rhythms in motor imagery.

The posterior probability $p(w_i|\mathbf{Z})$ is computed from the conditional probabilities using the Bayes rule [22]

$$P(w_i|\mathbf{Z}) = \frac{P(w_i)p(\mathbf{Z}|w_i)}{p(\mathbf{Z})} \quad (13)$$

The prior probabilities $P(w_i)$ are respectively 1/15 and 14/15 for target and non-target (probability of the events).

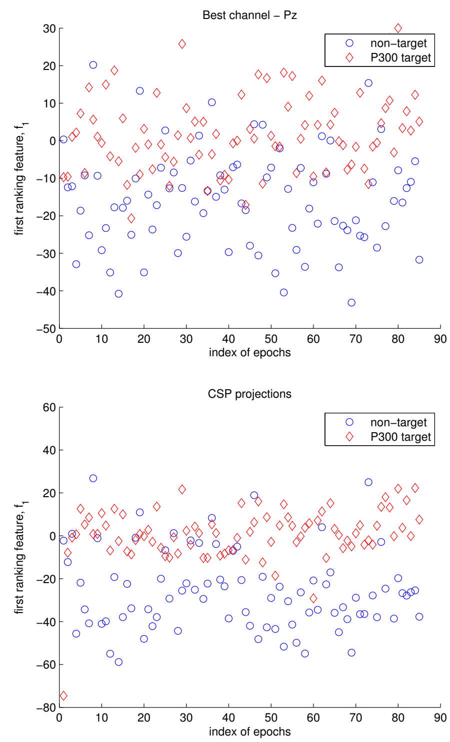

Fig. 6. First ranked feature for best channel (top) and CSP projection (bottom).

III. Results

The results were obtained through experiments performed with 3 healthy subjects as described in section II-B. The data were divided into a training set and a testing set. From the training set it was applied the Fisher criterion (1), from which the temporal features were ranked and the EEG channels were selected (see Fig. 5).

The CSP filter is applied to the selected channels and the projections from the optimum filters (10) are concatenated (11). Then, using the FS, the features with higher relevance are selected. The conditional probabilities in (12) are computed for each class from these data. The estimated class is reached using the Bayes decision function through the posterior probabilities (13).

The plots in Fig. 6 compares the CSP projection and the best channel for several epochs (it was used the same number of target and non-target epochs for a better illustration). Using the first ranked feature for both cases, it can be seen that the CSP projection improves the discrimination between the 2 classes. The first ranked feature was obtained from single epochs. Table II shows the achieved classification results for the 3 subjects. Classification tests were performed using the best channel, the filtered CSP projections using all 12 channels and the filtered CSP projections using the selected channels. The algorithms were tested for a single epoch (1 trial) and for the average of 5 epochs (5 trials). The results confirm that the application of CSP reduces the classification error rate when compared with best channel selection. Furthermore, the application of CSP to the se-

TABLE II

ERROR RATE CLASSIFICATION (%)

Nr. trials	Method	Subjects S1	S2	S3
1	best ch	17.0	19.6	15.0
	CSP all ch	10.0	15.8	9.2
	CSP sel. ch	10.0	15.6	7.0
5	best ch	5.3	6.5	5.0
	CSP all ch	4.5	6.1	4.4
	CSP sel. ch	3.0	6.5	2.1

lected channels also improve the classification accuracy when compared with CSP applied to all channels. Subject S3 has the better performance and subject S2 the worst. For this subject there is no significant improvement with CSP. This occurs because this subject only presents one channel with significant discrimination. Therefore the application of CSP to several channels does not improve the results.

The classification performance in Table II shows state of the art results [5] [23]. As the algorithms were applied only offline, we don't have yet a value for the online effective transfer rate. The offline transfer rate (bits/min) can however be obtained through the widely used measure [24]

$$B = M[\log_2(N_s) + P\log_2(P) + (1-P)\log_2\frac{(1-P)}{(N_s-1)}] \tag{14}$$

where N_s is the number of possible selections, P is the accuracy probability and M is the average number of decisions per minute (related with the time needed to make a decision). For example, for our case, considering $N = 15$, $P \approx 0.95$ (classification rate for 5 trials average), and $M = 3.75$ decisions/min (obtained for 5 trials and 200 ms stimuli interflash), the transfer rate B is 7.44 bit/min. This measure takes into consideration not only the classification accuracy but also other parameters related to the paradigm, such as the number of symbols, the way they are flashed, and the time between flashes. These parameters can be further improved to increase the transfer rate.

IV. CONCLUSION

This paper summarizes several BCI approaches showing their strong and weak points for real world application. A full methodology to implement a P300 BCI system is described. The system uses a temporal feature selection and feature extraction prior to classification which improves the classification accuracy. The CSP method combined to a Bayesian classifier showed to be an effective approach for P300 detection. The results were reached offline and therefore further developments are needed to the online implementation. Several parameters of the proposed P300 paradigm can be improved to increase the transfer rate.

REFERENCES

[1] G. Fabiani, D. J. McFarland, J. R. Wolpaw, and P. Pfurtscheller, "Conversion of eeg activity into cursor movement by brain-computer interface (bci)," *IEEE Trans. on Neural Systems and Rehabilitation Eng.*, vol. 12, no. 3, pp. 331–338, September 2004.

[2] G. Pfurtscheller, N. Christa, A. Scholgl, and K. Lugger, "Separability of eeg signals recorded during right and left motor imagery using adaptive autoregressive parameters," *IEEE Trans. on Rehabilitation Eng.*, vol. 6, no. 3, pp. 316–324, September 1998.

[3] L. Farwell and E. Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event related brain potentials," *Electr. and Clinical Neuroph.*, vol. 70, no. 6, pp. 510–523, 1988.

[4] E. Donchin, K. Spencer, and W. R., "The mental prosthesis: Assesing the speed of a p300-based brain-computer interface," *IEEE Trans. on Rehabilitation Eng.*, vol. 8, no. 2, pp. 174–179, June 2000.

[5] H. Serby, E. Yom-Tov, and G. Inbar, "An improved p300-based brain-computer interface," *IEEE Trans. on Neural Systems and Rehabilitation Eng.*, vol. 13, no. 1, pp. 89–98, March 2005.

[6] X. Gao, D. Xu, M. Cheng, and S. Gao, "A bci-based environmental controller for the motion disabled," *IEEE Trans. on Neural Systems and Rehabilitation Eng.*, vol. 11, no. 2, pp. 137–140, June 2003.

[7] M. Naeem, C. Brunner, R. Leeb, and G. Graimann, B. Pfurtscheller, "Separability of four-class motor imagery data using independent component analysis," *IOP - Journal of Neural Engineering*, vol. 3, pp. 208–216, 2006.

[8] G. Pires and U. Nunes, "A wheelchair steered through voice commands and assisted by a reactive fuzzy logic controller," *Journal of Intelligent and Robotic Systems*, vol. 34, no. 3, pp. 301–314, 2002.

[9] J. Philips, J. Millan, G. Vanacker, E. Lew, F. Galan, P. Ferrez, H. Van Brussel, and M. , Nuttin, "Adaptive shared control of a brain-actuated simulated wheelchair," in *IEEE 10th International Conference on Rehabilitation Robotics ICORR2007*, 2007, pp. 408–414.

[10] G. Pires, M. Castelo-Branco, and U. Nunes, "Visual p300-based bci to steer a wheelchair: A bayesian approach," *30th Annual Int. IEEE EMBS, EMBC08*, pp. 658–661, Aug. 2008.

[11] G. Pires, U. Nunes, and M. Castelo-Branco, "P300 spatial filtering and coherence-based channel selection," *4th Int. IEEE EMBS Conf. on Neural Engineering, NER09*, pp. 311–314, Apr 2009.

[12] K. Fukunaga, *Introduction to Statistical Pattern Recognition, Second Edition*. Morgan Kaufmann - Academic Press, 1990.

[13] D. J. McFarland, L. McCane, S. V. David, and J. R. Wolpaw, "Spatial filter selection for eeg-based communication," *Electroencephalography and clinical Neurophys.*, vol. 103, pp. 386–394, 1997.

[14] A. Soong and Z. Koles, "Principal-component localization of the sources of the background eeg," *IEEE Trans. on Biomedical Eng.*, vol. 42, no. 1, pp. 59–67, Jan. 1995.

[15] H. Ramoser, J. Muller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial eeg during imagined hand movement," *IEEE Trans. on Rehabilitation Eng.*, vol. 8, no. 4, pp. 441–446, Dec 2000.

[16] J. Muller-Gerking, G. Pfurtscheller, and H. Flyvbjerg, "Designing optimal spatial filters for single-trial eeg classification in a movement task," *Clinical Neurophysiology*, vol. 110, no. 5, pp. 787–798, 1999.

[17] Y. Li, X. Gao, H. Liu, and S. Gao, "Classification of single-trial electroencephalogram during finger movement," *IEEE Trans. on Biomedical Eng.*, vol. 51, no. 6, pp. 1019–1025, June 2004.

[18] G. Dornhege, B. Blankertz, G. Curio, and K.-R. Muller, "Boosting bit rates in noninvasive eeg single-trial classifications by feature combination and multiclass paradigms," *IEEE Trans. on Biomedical Eng.*, vol. 51, no. 6, pp. 993–1002, June 2004.

[19] Y. Wang, P. Berg, and M. Scherg, "Common spatial subspace decomposition applied to analysis of brain responses under multiple task conditions: a simulation study," *Clinical Neurophysiology*, vol. 110, pp. 604–614, 1999.

[20] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K.-R. Muller, "Optimizing spatial filters for robust eeg single-trial analysis," *IEEE Signal Processing Magazine*, vol. 25, no. 1, pp. 41–56, 2008.

[21] D. J. Krusienski, W. Sellers, and T. M. Vaughan, "Common spatio-temporal patterns for the p300 speller," *3rd IEEE EMBS Intern. Conf. on Neural Engineering*, pp. 421–424, May 2007.

[22] F. Heijden, R. Duin, D. Ridder, and D. Tax, *Classification, Parameter Estimation and State Estimation*. John Wiley & Sons, 2004.

[23] A. Rakotomamonjy and V. Guigue, "Bci competition iii: Dataset ii-ensemble of svms for bci p300 speller," *IEEE Trans. on Biomedical Eng.*, vol. 55, no. 3, pp. 1147–1154, March 2008.

[24] J. Wolpaw, N. Birbaumer, W. Heetderks, D. McFarland, P. Peckham, G. Schalk, E. Donchin, L. Quatrano, C. Robinson, and T. Vaughan, "Brain-computer interface technology: a review of the first international meeting," *IEEE Trans. on Rehabilitation Eng.*, vol. 8, no. 2, pp. 164–173, Jun 2000.