A Brain Computer Interface methodology based on a visual P300 Paradigm

Gabriel Pires, Urbano Nunes and Miguel Castelo-Branco

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. Features are selected through a Fisher criteria and EEG channels are selected through a coherence measure. 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. The classifier follows a Bayesian methodology for feature combination that overcomes the limitations of the feature method used in motor imagery. Offline classification results 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 on 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 was 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. The stimulus has to 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 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 to a 2-class discrimination problem, i.e. the discrimination between target events (desired command) and non-target events. 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 with a minor frequency.

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

Gabriel Pires is with the Institute for Systems and Robotics, University of Coimbra, 3000 Coimbra, Portugal, and with the Department of Electrical Engineering, Institute Polytechnic of Tomar, 2300 Tomar, Portugal gpires@isr.uc.pt

Urbano Nunes is with the Institute for Systems and Robotics, University of Coimbra, 3000 Coimbra, Portugal urbano@isr.uc.pt

Miguel Castelo-Branco is with the Biomedical Institute for Research in Light and Image (IBILI), University of Coimbra, Portugal mcbranco@ibili.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

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 respectively due to continuous user focus and flickering. 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. The user should be able to provide sparse commands indicating low-level steering directions or high-level commands or tasks, at the same time as an intelligent navigation module provides a safe navigation, using the same concept described in [8]. 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 several 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. The development of a BCI system should provide a systematic procedure to obtain robust classifiers for realtime operation and subject short time training phase. Some of the main properties of this procedure are:

1) Pre-processing and normalization steps for robustness

to inter-trial variability;

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.

II. BCI SYSTEM BASED ON P300 PARADIGM

The overall training BCI system is described in Fig. 1. 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. 2 and was already presented in [10]. It is composed by 8 direction arrows, a stop square, a ON/OFF switch and 5 small squares with no special meaning. The paradigm was specifically designed to steer a robotic device though the detection of the desired steering direction. However, each arrow can be associated to a high level command such as room, toilet, desk, door, etc. Each symbol is randomly flashed with uniform distribution, therefore the event target occurs once on each 15 flashes providing an oddball paradigm.

B. EEG acquisition and pre-processing

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,
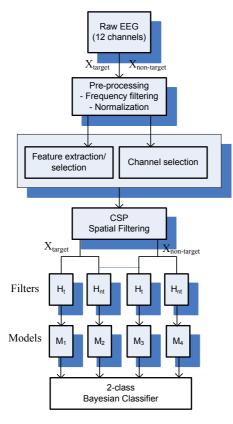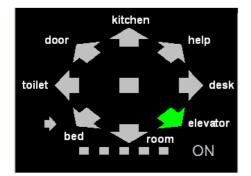
Fig. 1. Overall training system.

Fig. 2. 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.

POz and Oz according to the internacional extended 10-20 standard system using a g.tec cap. 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 80 target epochs and 1120 non-target epochs, which takes about 4 minutes. Each epoch has a duration of 1 second and is synchronized with the start 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
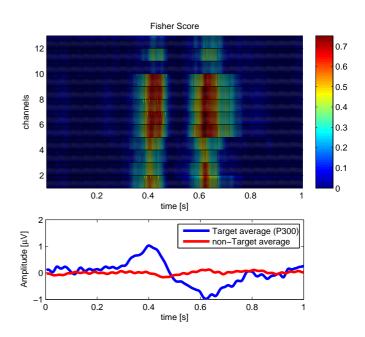

Fig. 3. Top: color map indicating level of discrimination between target and non-target classes (channels are ordered according to Fz, Cz, C3, C4, CPz, Pz, P3, P4, PO7, PO8, POz and Oz); bottom: Average of target events (P300) and non-target events for Fz channel.

deviation.

C. Feature selection - Fisher Criterion

The P300 features correspond to the amplitudes of the P300 temporal pattern. The Fisher criterion (FC) [11] 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 time-window features (time segment within the 1 second epoch) and the best channels that present best discrimination. The Fisher score FS_j of the feature j (instant time) is given by

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

where $\mu_j(\mathbf{X_t})$ and $\mu_j(\mathbf{X_{nt}})$ are respectively the target and non-target means 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. 3 shows a color map representing the Fisher score. High level colors denote high levels of discrimination. The bottom Figure presents the average 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. Channel selection

The method presented in the preceding section can be used to select the channels with best target vs. non-target discrimination. However, for application of CSP, experiments showed that selecting channels with strong linear correlation

provided improved results. Therefore, a coherence measure is used to identify channels with strong inter-correlation.

1) Coherence: Coherence gives a linear correlation between two signals as a function of the frequency. In the context of neurophysiology, it is used to measure the linear dependence and functional interaction between different brain regions. Mathematically, the estimated coherence between signals x and y is computed by the magnitude squared coherence [12]

$$k_{xy}^2(f) = \frac{|\langle S_{xy}(f)\rangle|^2}{|\langle S_{xx}(f)\rangle||\langle S_{yy}(f)\rangle|} \tag{2}$$

where $S_{xx}(f)$ and $S_{yy}(f)$ are respectively the power spectra of x and y, and S_{xy} is the cross-power spectrum. The spectra is estimated from the average ($\langle . \rangle$) of the periodogram over the set of trials.

E. Common Spatial Patterns

EEG signals are characterized by a very poor spatial resolution and the signal of interest has a very low signal-to-noise ratio. Working toward an effective BCI motivates the researchers to investigate methods that can increase communication bandwidth (ideally with single-trial) and reduce the channel dimension. To reach these goals, spatial filtering plays an important role and is an indispensable processing step for the feature extraction and pattern classification. Spatial filters can accentuate the signal of interest and at the same time attenuate the ongoing EEG and the non-EEG artifacts. In [13], McFarland et al. compare reference filter methods such as Common Average Reference, Small Laplacian and Large Laplacian with conventional ear reference, and show the improvement obtained with these spatial filters. They act as high-pass spatial filters that enhance local activity and decrease the distributed activity.

A different approach named Common Spatial Patterns was proposed by Koles [14]. It was applied in clinical electroencephalography for localization of sources of specific neurophysiologic components and to extract high frequency spike and sharp wave components from the EEG of neurologic patients. The CSP method is based on the simultaneous diagonalization of two real symmetric matrices proposed by Fukunaga [11]. 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. At the same time, a dimension reduction is achieved which is an important step for posterior classification.

This method has been successfully applied in BCI research for extraction and enhancement of ERD/ERS and Readiness Potential features associated with motor imagery paradigms [15], [16], [17]. Some variants of CSP were already proposed for the multiclass problem [18] [19]. A survey on CSP methods can be found in [20].

There are however very few applications of the CSP method on the detection of Event Related Potentials (ERP) such as the P300. As a relevant work, we point out the work in [21] where an extension of the CSP method is suggested, namely the Common Spatio-Temporal Patterns (CSTP). This approach incorporates spatio-temporal covariance matrices into CSP to extract more prominent spatio-temporal patterns.

We propose here the application of standard CSP combined with a new approach of feature combination based on probabilistic models of spatial filtered data embedded in a Bayesian classifier. The CSP is an attractive method because it is computationally efficient as it will be seen in the following.

Within the P300 oddball principle context, we 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

$$\mathbf{R} = \mathbf{R_{nt}} + \mathbf{R_t} \tag{3}$$

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

$$\mathbf{R_t} = \frac{\mathbf{X_t}\mathbf{X_t'}}{tr(\mathbf{X_t}\mathbf{X_t'})} \tag{4}$$

$$\mathbf{R_{nt}} = \frac{\mathbf{X_{nt}}\mathbf{X_{nt}'}}{tr(\mathbf{X_{nt}}\mathbf{X_{nt}'})} \tag{5}$$

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

$$\overline{\mathbf{R_t}} = \frac{1}{N_t}\sum_{i=1}^{N_t}\mathbf{R_t}(i) \quad \overline{\mathbf{R_{nt}}} = \frac{1}{N_{nt}}\sum_{i=1}^{N_{nt}}\mathbf{R_{nt}}(i) \tag{6}$$

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 factored through the application of the PCA as follows [22]

$$\overline{\mathbf{R}} = \overline{\mathbf{R_t}} + \overline{\mathbf{R_{nt}}} = \mathbf{A}\lambda\mathbf{A'} \tag{7}$$

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 $\mathbf{W}$

$$\mathbf{W} = \lambda^{-\frac{1}{2}}\mathbf{A'} \tag{8}$$

transforms the covariance matrix $\overline{\mathbf{R}}$ to I (identity matrix)

$$\mathbf{S} = \mathbf{W}\overline{\mathbf{R}}\mathbf{W'} = \mathbf{I}. \tag{9}$$

Applying the whitening transform to each individual class, we obtain

$$\mathbf{S_t} = \mathbf{W}\overline{\mathbf{R_t}}\mathbf{W'} \tag{10}$$

$$\mathbf{S_{nt}} = \mathbf{W}\overline{\mathbf{R_{nt}}}\mathbf{W'}. \tag{11}$$

From the above three equations it is straightforward that

$$\mathbf{S_t} + \mathbf{S_{nt}} = I \tag{12}$$

Performing a PCA factorization to (10) and (11) then

$$\mathbf{S_t} = \mathbf{A_t}\lambda_t\mathbf{A_t^{'}} \quad \mathbf{S_{nt}} = \mathbf{A_{nt}}\lambda_t\mathbf{A_{nt}^{'}} \quad (13)$$

From (12) and (13)

$$\mathbf{A_t} = \mathbf{A_{nt}} \quad (14)$$

and

$$\lambda_t = \mathbf{I} - \lambda_{nt}. \quad (15)$$

It means that both class patterns share the same eigenvectors and the respective eigenvalues are reversely ordered. The eigenvector with largest eigenvalue 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_t and A_{nt} as the first and last eigenvectors with dimension $N \times 1$ the following spatial filters are designed

$$\mathbf{H_t} = \mathbf{A_t^{'}W} \quad (16)$$

$$\mathbf{H_{nt}} = \mathbf{A_{nt}^{'}W}. \quad (17)$$

The spatial filtered data is given by

$$\mathbf{Y} = \mathbf{HX}. \quad (18)$$

F. Bayesian Classification

The Bayesian classifier uses the features from spatially filtered data projections. Four probabilistic models are created as shown in Fig. 1. 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.

For each projected sequence a conditional probability is computed, namely

$$p(\mathbf{x_{ft}}|M_1) \quad p(\mathbf{x_{fnt}}|M_2) \quad p(\mathbf{x_{ft}}|M_3) \quad p(\mathbf{x_{fnt}}|M_4) \quad (19)$$

where the vector $\mathbf{x_f}$ represents the filtered projection. Let define w_t and w_{nt} respectively the class of target and nontarget events. The posterior probability $p(w_i|\mathbf{X})$ ($i = t, nt$), i.e., the probability of a non spatially filtered data pattern $X_{N \times T}$ belong to class w_i is obtained through the Bayes rule [23]

$$p(w_t|\mathbf{X}) = \frac{P(w_t)p(\mathbf{x_{ft}}|M_1)p(\mathbf{x_{fnt}}|M_2)}{p(\mathbf{X})} \quad (20)$$

$$p(w_{nt}|\mathbf{X}) = \frac{P(w_{nt})p(\mathbf{x_{ft}}|M_3)p(\mathbf{x_{fnt}}|M_4)}{p(\mathbf{X})} \quad (21)$$

where $p(\mathbf{X})$ is the unconditional density of $\mathbf{x}$ and $P(w_t), P(w_{nt})$ are the prior probabilities of each of the classes. The conditional probability $p(\mathbf{x_f}|M)$ is computed from the likelihood function under a gaussian distribution assumption

$$p(\mathbf{x_f}|(\mu, \Sigma)) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp(-\frac{(\mathbf{x_f} - \mu)^T(\mathbf{x_f} - \mu)}{2\Sigma}) \quad (22)$$

where μ and Σ are the mean and covariance matrices computed for each class.
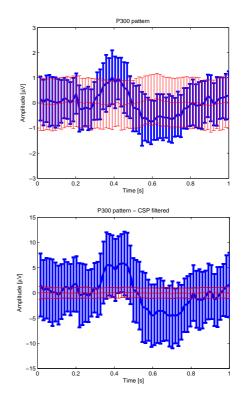

Fig. 4. Top: best channel mean and standard deviation over the set of target trial (blue) and nontarget trials (red); down: average and standard deviation of CSP filtered data from CPz, Pz, P3 and P4 channels.

III. RESULTS

A cluster of contiguous channels that evidence a strong inter-correlation (a value above 0.9 for frequency in the range 0.5-7 Hz) was selected. Spatial filter CSP is then applied to this group of channels. Selecting the best time-window features, the conditional probabilities in (19) are computed through (22). The estimated class is reached using the Bayes decision function through the posterior probabilities (20) and (21) associated to each class.

The top of Fig. 4 shows the average and standard deviation of the P300 pattern of the best channel (in this case P07) over 80 target epochs (blue) and 720 nontarget epochs (red). The overlapping of the two classes is a measure of how low is the discrimination between them. Also it is important to see that the standard deviation is almost constant over the time sequence. The bottom of Fig. 4 represents the average and standard deviation of CSP filtered data. The blue color represents the projected target epochs using filter H_t and red color the projected nontarget epochs through the filter H_{nt}. The plot shows clearly that the CSP filters separate the two classes. Note however, that the current plot is the ideal situation where each epoch class is projected with filters that maximizes and minimizes the respective classes. In practice we will have to use the four models shown in Fig. 1 and not only two models.

Table II shows the achieved classification results with 3 subjects. Classification tests were performed using the best channel, the filtered CSP projections using all 12 channels

TABLE II

ERROR RATE CLASSIFICATION (%)

trials	method	Subjects		
		S1	S2	S3
1	best ch	18.3	18.4	11.22
	CSP all ch	15.0	14.2	10.1
	CSP sel. ch	15.0	18.5	8.3
2	best ch	12.5	10.6	6.5
	CSP all ch	10.2	8.0	3.0
	CSP sel. ch	10.0	12.8	2.8
5	best ch	3.0	6.0	2.7
	CSP all ch	1.9	5.4	1.0
	CSP sel. ch	1.2	5.4	0.5

and the filtered CSP projections using the channels selected through coherency. The use of CSP filter reduces the error rate when compared with single channel classification. In case of subject S2 the improvement of classification is not so significant. Actually, for this subject the individual channels showed a low target vs. nontarget discrimination. For subjects S1 and S3 the use of the selected channels demonstrates a better performance than the use of all channels which can confirm that CSP has a better performance when used with channels linearly correlated.

IV. CONCLUSION

The presented work describes a fully methodology for a P300 BCI system. The proposed methods for feature selection, channel selection, and spatial filtering presented effective results. It was introduced a novel Bayesian classification method based on probabilistic models created from spatial filtered data. Also, it was shown that channel selection through coherence measure is a suitable method to reduce data dimension and improve classification performance. Notwithstanding, the results were reached offline. It is expected a performance decrease on onroad online operation mainly due to the presence of muscular and other artifacts. The achieved offline results are significative when compared with state of the art.

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 transactions on Neural Systems and Rehabilitation Engineering*, 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 Transactions on Rehabilitation Engineering*, 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: Asensing the speed of a p300-based brain-computer interface," *IEEE transactions on Rehabilitation Engineering*, 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 Transactions on Neural Systems and Rehabilitation Engineering*, 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 transactions on Neural Systems and Rehabilitation Engineering*, 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. Conf. of the IEEE Eng. in Medicine and Biology Society, EMBS2008*, pp. 658–661, Aug. 2008.

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

[12] J. S. Bendat and A. G. Piersol, *Random data - Analysis and Measurement Procedure*. John Wiley & Sons, Inc, 2000.

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

[14] A. Soong and Z. Koles, "Principal-component localization of the sources of the background eeg," *IEEE Transactions on Biomedical Engineering*, 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 Engineering*, vol. 8, no. 4, pp. 441–446, Dec 2000.

[16] J. Mller-Gerking, G. Pfurtscheller, and H. Flyvbjerg, "Designing optimal spatial filters for single-trial eeg classification in a movement task," *Journal of 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 Transactions on Biomedical Engineering*, 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," *Biomedical Engineering, IEEE Transactions on*, 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," *Signal Processing Magazine, IEEE*, 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] I. T. Jolliffe, *Principal Component Analysis, Second Edition*. Springer Series in Statistics, 2002.

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