Feature Extraction for Multi-class BCI using Canonical Variates Analysis

Ferran Galán†, Pierre. W. Ferrez†, Francesc Oliva‡, Joan Guàrdia§, and José del R. Millán†
†IDIAP Research Institute, 1920 Martigny, Switzerland. E-mail: Ferran.Galan@idiap.ch
‡Dept. of Statistics, University of Barcelona, Barcelona, Spain
§Dept. of Methodology of the Behavioural Sciences, University of Barcelona, Barcelona, Spain

Abstract **– Objective: To propose a new feature extraction method with canonical solution for multi-class Brain-Computer Interfaces (BCI). The proposed method should provide a reduced number of canonical discriminant spatial patterns (CDSP) and rank the channels sorted by power discriminability (DP) between classes. Methods: The feature extractor relays in Canonical Variates Analysis (CVA) which provides the CDSP between the classes. The number of CDSP is equal to the number of classes minus one. We analyze EEG data recorded with 64 electrodes from 4 subjects recorded in 20 sessions. They were asked to execute twice in each session three different mental tasks (left hand imagination movement, rest, and words association) during 7 seconds. A ranking of electrodes sorted by power discriminability between classes and the CDSP were computed. After splitting data in training and test sets, we compared the classification accuracy achieved by Linear Discriminant Analysis (LDA) in frequency and temporal domains. Results: The average LDA classification accuracies over the four subjects using CVA on both domains are equivalent (57.89% in frequency domain and 59.43% in temporal domain). These results, in terms of classification accuracies, are also reflected in the similarity between the ranking of relevant channels in both domains. Conclusions: CVA is a simple feature extractor with canonical solution useful for multi-class BCI applications that can work on temporal or frequency domain.**

Keywords **– Electroencephalogram, Brain-computer interfaces, Canonical Variates Analysis, Linear Discriminant Analysis.**

I. Introduction

Brain-computer interfacing (BCI) research enables a new interaction modality with the environment. Many applications have been explored in recent years [1], [2], [3], [4], [5], [6]. Our work is focused on asynchronous and non-invasive electroencephalogram (EEG) based BCI to control robots and wheelchairs [7], [8]. It means that the users drive such devices by learning to voluntary control specific EEG features. To facilitate this learning process it is necessary to select those subject-specific features that allow to generate the maximum number of discriminant patterns. This process becomes crucial to facilitate the generation of those patterns that will permit an easier execution of those commands needed to drive the different devices. To this end, Common Spatial Patterns (CSP) [9] and his extension Commom Spatio Spectral Patterns (CSSP) [10] have been proven very useful. However, *there is no canonical way to choose the relevant CSP patterns for multi-class CSP and only approximative solutions can be obtained* [11]. In the present paper we propose a new feature extraction method with canonical solution for multi-class BCI. The feature extractor utilized relays on Canonical Variates Analysis (CVA) [12], also known as Multiple Discriminant Analysis [13], that provides the canonical discriminant spatial patterns (CDSP) between the classes. The number of CDSP is equal to the number of classes minus one.

The paper is structured as follows: Section II describes CVA and the experimental setup, preprocessing and analysis carried out to assess its usability for multi-class BCI feature extraction; Section III reports the results; and finally in Section IV gives some conclusions and discusses future work.

II. Methods

A. Canonical Variates Analysis

In our BCI research the user employs the voluntary modulation of different oscillatory rhythms [7] by executing of different mental tasks (motor and cognitive) to drive robots and wheelchairs in virtual [8] and real environments. In these

This work was supported in part by Agència de Gestió d'Ajuts Universitaris i de Recerca, Departament d'Universitats Recerca i Societat de la Informació, Generalitat de Catalunya, under Grants 2001SGR00139 and 2001SGR00067, by the Spanish Ministerio de Educación y Ciencia, under Grant MTM2004-00440, by the Swiss National Science Foundation through the National Centre of Research on "Interactive Multimodal Information Management (IM2)", and by the European IST Programme FET Project FP6-003758. This paper only reflects the author's views and founding agencies are not liable for any use that may be made of the information contained herein.

applications the users utilize more than two commands. To facilitate this voluntary modulation it is necessary to find those subject-specific spatial patterns that maximize the separability between the patterns generated by executing the different mental tasks. In this way, from band-pass filtered EEG signals, the CSP algorithm extracts canonical discriminant spatial patterns which directions maximizes the differences in variance between two classes. *Since the variance of a band-pass filtered signal is a measure for the energy in the corresponding frequency band, the patterns reflect the spatial distributions of event-related (de)synchronization effects* [14]. However, *there is no canonical way to choose the relevant CSP patterns for multi-class CSP and only approximative solutions can be obtained* [11]. This limitation can be avoided in two ways, namely working in frequency domain or working with the squared band-pass filtered EEG signal. In the former case, the energy in the corresponding frequency band is measured by its spectral power. In this domain the spatial distributions of event-related (de)synchronization effects are identified by changes on the spectral power. In the later case, the spatial distributions of event-event-related (de)synchronization effects are identified by changes on the mean, given that the variance of a band-pass filtered EEG signal becomes the mean when the signal is squared (see proof in the appendix). Thus, using CVA it is easy to extract CDSP which directions maximizes the differences in mean, either spectral power in the first case or energy of the original band-pass filtered EEG signal in the second case, between a given number of classes.

Given the $n_i \times c$ matrix, either with the estimated spectral power of a frequency band or the squared band-pass filtered EEG signal, $\mathbf{S}_i = (\mathbf{s}_{i1}, ..., \mathbf{s}_{in_i})^{'}$ of class $i = 1, ..., k$, where n_i is the number of samples and c is the number of channels, and $\mathbf{S} = (\mathbf{S}_1^{'}, ..., \mathbf{S}_k^{'})^{'}$, the $k-1$ CDSP of $\mathbf{S}$ are the eigenvectors $\mathbf{A}$ of $\mathbf{W}^{-1}\mathbf{B}$ which eigenvalues $\lambda_u, (u = 1, ..., k-1)$ are larger than 0. Note that the direction of eigenvectors $\mathbf{A}$ maximize the quotient between the between-classes dispersion matrix

$$\mathbf{B} = \sum_{i=1}^{k} n_i(\mathbf{m}_i - \mathbf{m})(\mathbf{m}_i - \mathbf{m})^{'} \tag{1}$$

and the pooled within-classes dispersion matrix

$$\mathbf{W} = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(\mathbf{s}_{ij} - \mathbf{m}_i)(\mathbf{s}_{ij} - \mathbf{m}_i)^{'} \tag{2}$$

where

$$\mathbf{m}_i = \frac{1}{n_i}\sum_{j=1}^{n_i}\mathbf{s}_{ij} \tag{3}$$

and

$$\mathbf{m} = \frac{1}{n}\sum_{i=1}^{k} n_i\mathbf{m}_i \tag{4}$$

are the class and total centroids respectively. Thus, the new features are obtained by the projection

$$\mathbf{Y} = \mathbf{SA} \tag{5}$$

Once the CDSP are computed, it is useful to know how the original features (electrodes) are contributing in the separability between the classes. It also permits to interpret the space generated by the CDSP, specially when the number of classes is high. In this way, it is possible to rank the channels given their contribution on the new space. We define a new *Discriminant Power* (DP) [15] measure for each channel from the *structure matrix*, pooled correlation matrix between original channels in $\mathbf{S}$ and the new features in $\mathbf{Y}$. Given the $c \times k-1$ structure matrix $\mathbf{T}$, where $\mathbf{T} = \sum_{i=1}^{k}\mathbf{T}_i$, $e = 1, ..., c$, and the normalized eigenvalues $\gamma_u = \lambda_u / \sum_{u=1}^{k-1}\lambda_u$, the proposed DP can be computed as follows

$$DP_e = (\sum_{u=1}^{k-1}\gamma_u t_{eu}^2 / \sum_{e=1}^{c}\sum_{u=1}^{k-1}\gamma_u t_{eu}^2) \times 100 \tag{6}$$

B. Data Acquisition and Task

Data were recorded from 4 subjects with a portable Biosemi acquisition system using 64 channels sampled at 512Hz and high-pass filtered at 1Hz. The subjects were sitting in a chair looking at a fixation cross placed at the center of a monitor. The subjects were instructed to execute three different mental tasks (left hand imagination movement, rest, and words association) in a self-paced way. The mental task to be executed was previously specified by the operator in order to counterbalance the order, the subjects specify when they started to execute the mental task. Each subject participated in 20 sessions integrated by 6 trials each, 2 trials of each class. The duration of each trial was 7 seconds but only the last 6 seconds were utilized in the analysis to avoid preparation periods. Subjects 1 and 2 had previous experience with the selected mental tasks, while it was the first time for subjects 3 and 4.

C. Preprocessing

To work in frequency domain the signal was spatially filtered using common average reference (CAR) previous to the estimation every 62.5 ms. (16 times per second) of the power spectral density (PSD) in the band 10-14Hz with 2Hz resolution over the last 1-second windows. PSD was estimated by Welch method with 5 overlapped (25%) Hanning windows of 500 ms. length. To work in temporal domain the signal was also spatially filtered by CAR, band-pass filtered in the frequency range 8-16Hz (to get a band-pass filtered signal in the same frequency ranges analyzed in the frequency domain, taking in account the FIR filter transition band) and finally squared. Single trials were obtained by averaging samples within last 1-second window. In both cases only 45 electrodes were utilized, namely: F1, F3, F5, FC1, FC3, FC5, C1, C3, C5, CP1, CP3, CP5, P1, P3, P5, P7, PO3, PO7, O1, Fz, FCz, Cz, CPz, Pz, POz, Oz, F2, F4, F6, FC2,

TABLE I.

LDA CLASSIFICATION ACCURACY OVER THE FOUR SUBJECTS ACCORDING TO THE DIFFERENT TEST SESSIONS USING CVA IN FREQUENCY AND TEMPORAL DOMAINS

Subject	Domain	Test Session					Average
		1	2	3	4	5	
1	F[a]	66.25%	76.04%	71.04%	70.41%	62.92%	69.33%
	T[b]	60.34%	87.05%	74.13%	73.54%	72.42%	73.50%
2	F	72.71%	59.79%	73.54%	69.37%	64.38%	67.95%
	T	62.36%	56.70%	69.81%	61.76%	71.14%	64.35%
3	F	43.54%	49.38%	55.00%	60.21%	50.63%	51.75%
	T	60.32%	60.04%	61.41%	50.28%	55.83%	57.57%
4	F	35.83%	61.45%	48.33%	33.54%	34.16%	42.66%
	T	31.24%	62.17%	35.71%	46.57%	35.95%	42.33%
Average	F						57.89%
	T						59.43%

[a] frequency domain, [b] temporal domain

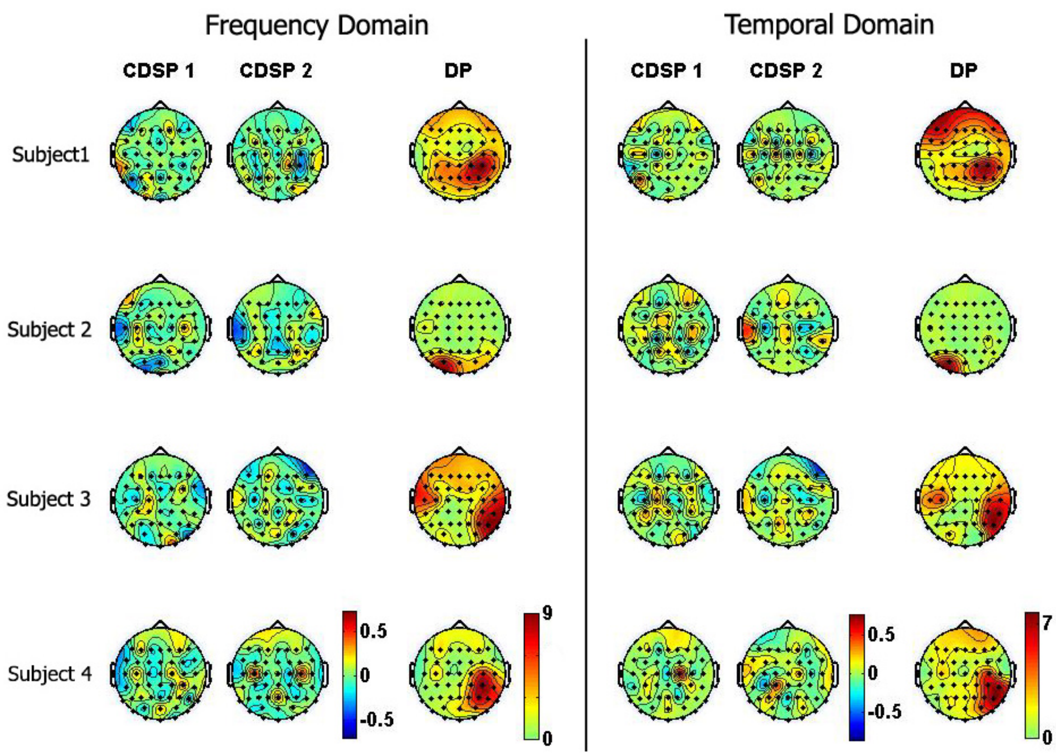

Figure 1. CDSP AND DP FOR EACH SUBJECT IN FREQUENCY AND TEMPORAL DOMAINS COMPUTED FROM TRAINING SET. NOTE THAT DP SCALE IS IN %.

FC4, FC6, C2, C4, C6, CP2, CP4, CP6, P2, P4, P6, P8, PO4, PO8, O2.

D. Analysis

To assess the canonical discriminant spatial patterns stability over time, data were split in two sets, the training set integrated by the trials from the first 15 sessions, and the test set integrated by the trials from the last 5 sessions. In frequency domain a trial was defined by each PSD estimation whereas in temporal domain each trial was defined as the averaged squared band-pass signal over the last second. After obtain the CDSP from the training set of each domain, training and test trials where projected in the new space using eq. 5. Then, we built one Linear Discriminant Analysis (LDA) classifier per subject and per domain whose parameters are estimated on the corresponding training sets. Finally, we used these LDA classifiers to assess the generalization performances of each subject. Given that the main problem in BCI research is to deal with EEG unstability over time, the use of k-fold crossvalidation was avoided. This non-parametric classification error estimator uses as training and test sets data from all sessions, what never occurs in on-line

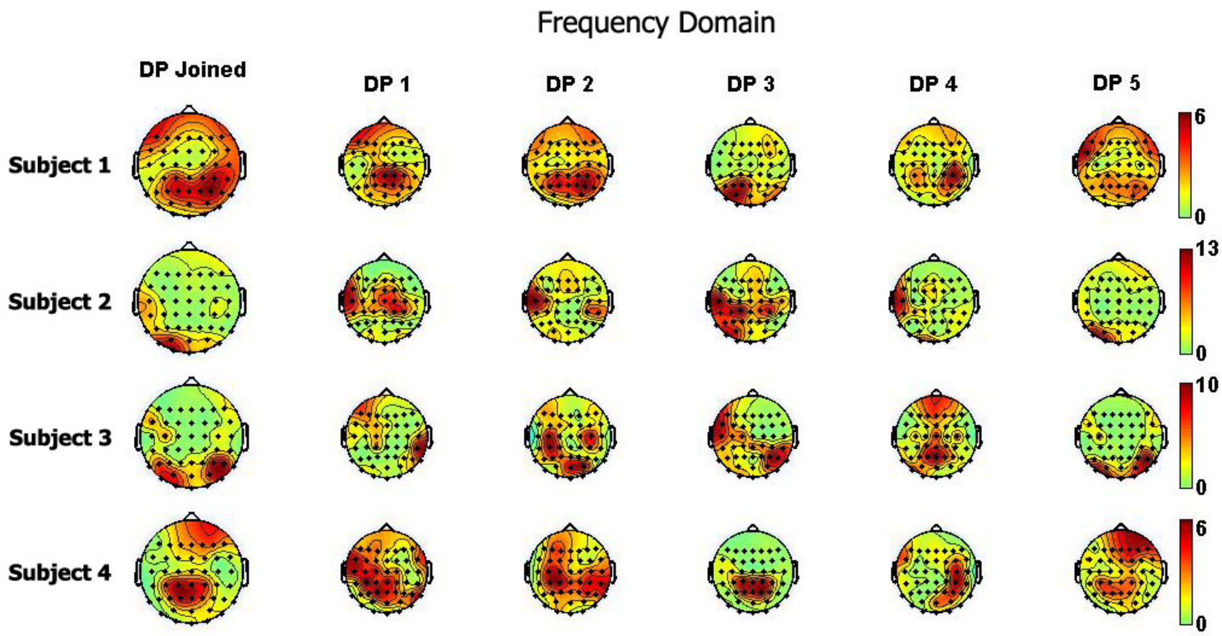

Figure 2. DP for each subject in frequency domain computed joining all test sessions and from every single test session. Note that DP scale is in %.

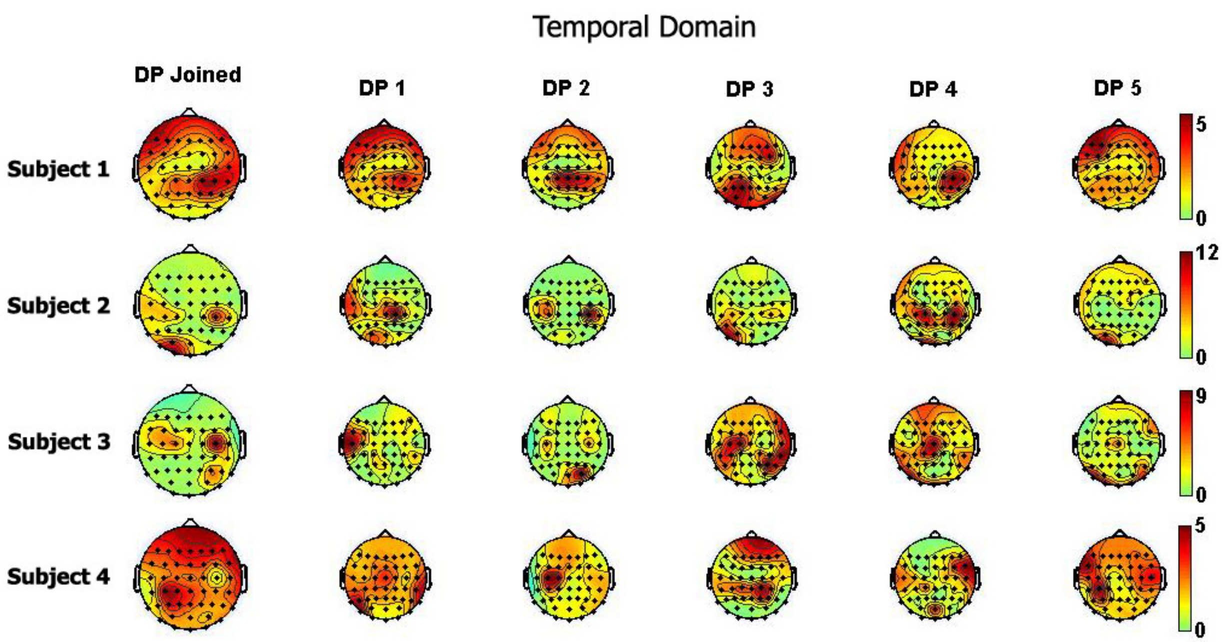

Figure 3. DP for each subject in temporal domain computed joining all test sessions and from every single test session. Note that DP scale is in %.

applications and yields optimistic error estimations.

III. Results

Table I reports the LDA classification accuracy over the 5 test sessions using CVA in frequency and temporal domain. In average, the classification accuracies for both domains are equivalent (57.89% in frequency domain vs. 59.43% in temporal domain, random level is 33.3% for a 3-class problem). In the temporal domain, we obtained higher classification accuracies for two subjects, namely subjects 1 and 3 (73.50% and 57.57% vs. 69.33% and 51.75%). In the frequency domain, we obtained higher classification accuracies only for one subject, namely subject 2 (67.95% vs. 64.35%). The performance is equivalent on subject 4 (42.66% vs. 42.33%). Fig. 1 depicts the two CDSP and the DP obtained for each subject in frequency and temporal domains computed on the training set. The CDSP interpretation as a whole it is facilitated by DP maps. DP maps show the electrodes contribution, in percentage, on the space defined by the CDSP. As expected according to the results obtained in terms of classification accuracy, DP maps obtained from both domains show a similar distribution of electrodes contribution in all subjects. Fig. 2 and Fig. 3 depict the DP for each subject in the frequency and temporal domains, respectively, computed joining all test sessions (first column) and also from every single test session (next five columns). These figures show the origin of the intersession variability and allow also to understand the results in terms of classification accuracy (see Table I). In both domains, the classification accuracy is related to the level of similarity between DP maps obtained from the training set (see DP maps in Fig. 1) and DP maps obtained from test sessions (see Fig. 2, frequency domain, and Fig. 3, temporal domain), either joining all test sessions or for each single test session. Higher classification accuracies correspond to higher similarity between the maps, what means that the canonical spaces defined by the CDSP estimated on the training sets are more stable over time. It is also worth noting that the similarity between DP maps obtained from both domains (DP joined in Fig. 2 and Fig. 3, first column) decreases on those subjects with lower classification accuracies.

IV. Conclusion and Further Research

The objective of this paper is to propose a new feature extraction method with a canonical solution for multi-class BCI. The estimated CDSP yield the space of maximum separability between event-related (de)synchronization effects involved in the execution of different mental tasks. The proposed DP measure rank the electrodes sorted by their contribution in the new space. The average LDA classification accuracies obtained working on frequency and temporal domains are equivalent. Performances are not very high for a 3-class problem because, for comparative purposes, we have classified every single trial obtained from the last second window. The equivalent results, in terms of classification accuracies, are also reflected in the similarity between the DP maps obtained from the training sets of both domains. On the other hand, the level of similarity between DP maps obtained from the testing sets of both domains decreases for those subjects with lower classification accuracies (subjects 3 and 4). A possible explanation that needs to be explored is that energy (temporal domain) and PSD estimation (frequency domain) do not reflect the same phenomena when the signal is less stationary, what occurs when the subject have difficulties to generate stable EEG patterns during the execution of the mental tasks. Future work will focus on testing different extensions of CVA, assessing the sources of performance variability between both domains on different subjects, and exploring the relation between energy and spectral estimation.

Appendix

Theorem 1: Given a band-pass filtered signal $x(t)$, $(t = 1,...,T)$, its variance is equal to the squared signal's mean:

$$\sigma_{x(t)} = \mu_{x^2(t)} \tag{7}$$

Proof: Given that

$$\mu_{x(t)} = 0 \tag{8}$$

$$\sigma_{x(t)} = \frac{\sum_{t=1}^{T}(x(t) - \mu_{x_{(t)}})^2}{T} \tag{9}$$

substituting (8) in (9) yields

$$\sigma_{x(t)} = \frac{\sum_{t=1}^{T} x^2(t)}{T} \tag{10}$$

that, by definition, it is $\mu_{x^2(t)}$ ■

References

[1] J. D. Bayliss, 'Use of the evoked potential P3 component for control in a virtual apartment,' *IEEE Trans. Neural Sys. Rehab. Eng.*, vol. 11, pp. 113–116, 2003.

[2] M. A. Nicolelis and J. K. Chapin, 'Controling robots with the mind,' *Sci. Am.*, vol. 287, pp. 46–53, 2002.

[3] J. M. Carmena, M. A. Lebedev, R. E. Crist, J. E. O'Doherty, D. M. Santucci, D. F. Dimitrov, P. G. Patil, C. S. Henriquez, and M. A. L. Nicolelis, 'Learning to control a brain-machine interface for reaching and grasping by primates,' *PLoS Biol.*, vol. 1, pp. 193–208, 2003.

[4] 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*, vol. 398, pp. 297–298, 1999.

[5] B. Obermaier, G. R. Müller, and G. Pfurtscheller, 'Virtual keyboard controlled by spontaneous EEG activity', *IEEE Trans. Neural Sys. Rehab. Eng.*, vol. 11, pp. 422–426, 2003.

[6] J. del R. Millán, 'Adaptive brain interfaces,' *Comm. of the ACM*, vol. 46, pp. 75–80, 2003.

[7] J. del R. Millán, F. Renkens, J. Mouriño, and W. Gerstner, 'Noninvasive brain-actuated control of a mobile robot by human EEG,' *IEEE Trans. Biomed. Eng.*, vol. 51, pp. 1026–1033, 2004.

[8] E. Lew, M. Nuttin, P. W. Ferrez, A. Degeest, A. Buttfield, and G. Vanacker, 'Noninvasive brain-computer interface for mental control of a simulated wheelchair,' *Proceedings of the 3rd International Brain-Computer Interface Workshop & Training Course*, 2006.

[9] H. Ramoser, J. Müller-Gerking, and G. Pfurtscheller, 'Optimal spatial filtering of single trial EEG during imagined hand movement,' *IEEE Trans. Neural Sys. Rehab. Eng.*, vol. 8, pp. 441–446, 2000.

[10] S. Lemm, B. Blankertz, G. Curio, and K-R. Müller, 'Spatio-spectral filters for improving the classification of single trial EEG,' *IEEE Trans. Biomed. Eng.*, vol. 52, pp. 1541–1548, 2005.

[11] G. Dornhege, B. Blankertz, G. Curio, and K-R. Müller, 'Increase information transfer rates in BCI by CSP extension to multi-class,' in: S. Thrun, L. Saul, and B. Schölkopf, eds., *Advances in Neural Inf. Proc. Systems*, vol. 16, pp. 733–740, 2004.

[12] W. J. Krzanowski, *Principles of multivariate analysis*, Oxford University Press, Oxford, 1998.

[13] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, John Wiley & Sons, 2nd ed., 2001.

[14] G. Dornhege, B. Blankertz, and G. Curio, 'Speeding up classification of multi-channel brain-computer interfaces: common spatial patterns for slow cortical potentials,' *In Proceedings of the 1st International IEEE EMBS Conference on Neural Engineering*, pp. 591–594, 2003.

[15] S. L. Gonzalez Andino, R. Grave de Peralta Menendez, G. Thut, J. del R. Millán, P. Morier, and T. Landis, 'Very high frequency oscillations (VHFO) as a predictor of movement intentions,' *Neuroimage*, vol. 32, pp. 170–179, 2006.