



HAL
open science

Identification of spatial and temporal features of EEG

Nisrine Jrad, Marco Congedo

► **To cite this version:**

Nisrine Jrad, Marco Congedo. Identification of spatial and temporal features of EEG. *Neurocomputing*, 2012, 90 (1), pp.66-71. 10.1016/j.neucom.2012.02.032 . hal-00777267

HAL Id: hal-00777267

<https://hal.science/hal-00777267>

Submitted on 17 Jan 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Identification of spatial and temporal features of EEG

Nisrine Jrad, Marco Congedo

*GIPSA-lab CNRS, Grenoble Univ., 961 rue de la Houille Blanche, 38402 GRENOBLE
Cedex, France*

Abstract

Classifying brain activities is a challenging task since Electroencephalography (EEG) recordings exhibit distinct and individualized spatial and temporal characteristics correlated with noise and various physical and mental activities. To increase classification accuracy, it is thus crucial to identify discriminant spatio-temporal features. This paper presents a method for analyzing the spatio-temporal characteristics associated with Event related Potentials (ERP)s. First, a resampling procedure based on Global Field Power (GFP) extracts temporal features. Second, a spatially weighted SVM (sw-SVM) is used to learn a spatial filter optimizing the classification performance for each temporal feature. Third, the so obtained ensemble of sw-SVM classifiers are combined using a weighted combination of all sw-SVM outputs. Results indicate that inclusion of temporal features provides useful insight regarding classification performance and physiological understanding.

Keywords:

Brain computer interface (BCI), Electroencephalography (EEG), Global Field Power (GFP), Support Vector Machine (SVM), Event related Potentials (ERP), spatial and temporal feature extraction.

1. Introduction

Brain Computer Interface (BCI) is a tool for direct communication between human and external devices [1]. Many BCIs make use of Electroencephalography (EEG) signals to categorize two or more classes and associate them to simple computer commands. EEG is a popular technology for BCI

Email address: `firstname.lastname@gmail.com` (Nisrine Jrad, Marco Congedo)

purposes due to its low-cost, portability and non-invasiveness. Classification of EEG is not an easy task because this latter can only measure blurred cortical activities due to the diffusion of the skull and the skin [2]. Furthermore, recordings are highly contaminated by noise of various sources. Interestingly, EEG signals often reveal various spatial and temporal characteristics. Thus, it is important to characterize both spatial and temporal dynamics of EEG data to provide reliable BCI control.

Usually spatial decomposition is performed to extract spatial features [3, 4, 5, 6]. Signal-processing criteria like the Signal-to-Noise Ratio and ratio of class variances [3, 6, 7] have been often employed for Evoked Response Potential (ERP). These methods are based on the assumption of the instantaneous and approximately linear relation between the amplitudes of the generating cerebral electrophysiological current sources and the amplitude of the observed scalp potential field. The idea is to find a linear transformation of the data optimizing the extraction of the relevant EEG spatial feature and the noise suppression. The performance of such filters mainly depends on the accuracy of spatial covariance estimations and is jeopardized by the non-stationary nature of the noise. Although a relation might be found between the objective functions of these aforementioned criteria and class separability, this relation has, to the best of our knowledge, never been addressed explicitly. Hence, spatial filter can be suboptimal from the classification point of view. Another limitation of these methods is that spatial decomposition is performed in a specific time window. However, ERPs reflect several temporal components, thus, spatial decomposition should be performed for each interesting interval occurring in the pre-fixed window.

Recently, some algorithms have been proposed to study where the discriminative information lies into the spatio-temporal plane. They visualize a matrix of separability measures into the spatio-temporal plane of the experimental conditions. The matrix is obtained by computing a separability index for each pair of spatial electrode measurement and time sample. Several measures of separability have been used, for instance the signed- r^2 -values [3], Fisher score and Student's t-statistic [8], or the area under the ROC curve [9]. Separability matrix should be sought as to automatically determine intervals with fairly constant spatial patterns and high separability values. This proves difficult and heuristic techniques are often employed to approximate interval borders. In addition, the three first aforementioned measures rely on the assumption that class distributions are Gaussian, which is seldom verified.

To overcome all these drawbacks, we develop a spatio-temporal data

driven decomposition technique. A two-stage feature extraction technique is proposed. First, a time feature extraction is performed based on Global Field Power (GFP) [10], defined for each time sample as the sum of the square potential across electrodes. Second, a spatially weighted SVM (sw-SVM) is used to learn for each time interval a sparse spatial filter optimizing directly the classification performance [11]. Finally, the ensemble of sw-SVMs obtained on selected temporal features are combined using a weighted average, to get a robust decision function.

The remainder of this paper is organized as follows. The proposed method is introduced in Section 2. Section 3 accounts for experiment and data set description. Section 4 shows and discusses experimental results. Finally, Section 5 holds our conclusions.

2. Method

Background brain activities, irrelevant to BCI tasks, continuously generate EEG signals that can be recorded almost anywhere over the scalp. These signals interfere with the EEG signals triggered by stimuli. Interestingly, post-stimulus signals are generated in particular regions of the brain at a given interval of time. Hence, they present specific spatial and temporal characteristics. This section describes a method for analyzing ERPs spatial and temporal characteristics considering the sequence :

- a temporal feature extraction,
- a spatial feature extraction embedded in a classification scheme,
- an ensemble of classifiers learning technique.

2.1. Problem description

In the following, we consider BCI applications with two classes of action providing a training set of labeled trials from which a decision function is learned. The decision function should correctly classify unlabeled trials. Let us consider an EEG post-stimulus trial recorded over S electrodes in a short time period of T samples as a matrix $\tilde{\mathbf{X}}_p \in \mathbb{R}^{S \times T}$. Hence, the entire available set of data can be denoted $\{(\tilde{\mathbf{X}}_1, y_1), \dots, (\tilde{\mathbf{X}}_p, y_p), \dots, (\tilde{\mathbf{X}}_P, y_P)\}$ with $y_p \in \{-1, 1\}$ the class labels. Our task consists on finding the spatio-temporal features that maximize discrimination between two classes.

2.2. Temporal features

To select temporal intervals in the ERP where discriminative peaks appear, Global Field Power (GFP) [10] is computed on the difference of the grand averages of the two class post-stimulus trials as following:

$$(GFP^2)_t = \frac{1}{S} \sum_{s=1}^S \left[\left(\sum_{p|y_p=1} (\tilde{\mathbf{X}}_p)_{s,t} - \sum_{p|y_p=-1} (\tilde{\mathbf{X}}_p)_{s,t} \right) \right]^2, \quad (1)$$

$(GFP^2)_t$ is the square of Global Field Power at a given time sample t and $(\tilde{\mathbf{X}}_p)_{s,t}$ is the EEG recording of trial p at electrode s and p^{th} post-stimulus time sample t .

Pronounced deflections with large peaks, denoting big dissimilarities between the two activities, are associated with large GFP values. Windows involving significant temporal features are chosen as intervals where GFP is high relative to the background EEG activity.

To select significant windows we require a statistical threshold for the observed GFP of the difference grand average trials in the two classes. Such threshold is estimated with a resampling method as the 95th percentile (5% type I error rate) of the appropriate empirical null distribution. For Q and R observed single trials in classes labeled 1 and -1 , respectively, we resample Q and R trials with random onset from the entire EEG recording. We compute the difference of the grand average of the Q and R random trials and retain the maximum value of GFP. The procedure is repeated 1000 times and the sought threshold is the 95th percentile of such max-GPF null distribution after 10% trimming. The trimming makes the estimated GFP more robust with respect to outliers given by eye blinks and other large-amplitude artefacts. Taking the max-GPF at each resampling ensures that the nominal type I error rate is preserved regardless the number of windows that will be declared significant.

Noteworthy, contiguous samples with high GFP coincide with stable deflection configurations where spatial characteristics of the field remains unchanged [10]. Since within each selected time window the spatial pattern is fairly constant, average across time is calculated. Averaging over time rules out aberrant values, reduces signal variability and attenuates noise. Besides, it reduces dramatically time dimensionality to I where I is the number of significant time features.

2.3. Spatial features and classifier : sw-SVM

Temporal filtering provides us with $\mathbf{X}_p \in \mathbb{R}^{S \times I}$ trials. Each column vector $(\mathbf{x}_p)_i \in \mathbb{R}^S$ reflects a spatial characteristic at a temporal feature $i \in \{1, \dots, I\}$. Hereafter, \mathbf{x}_p will refer to $(\mathbf{x}_p)_i$ for convenience. Hence, I spatial filters are learned over the different time components.

In this work, spatial filtering is learned jointly with a classifier in the theoretical framework of SVM. This method was proposed in [11] and called sw-SVM for spatially weighted SVM. It has the advantage of learning a spatial filter so as to improve separability of classes whilst reducing classification errors. It involves spatial feature weights in the primal SVM optimization problem and tunes these weights as hyper-parameters of SVM. We denote by $\mathbf{d} = [d_1, d_2, \dots, d_S]^T \in \mathbb{R}^S$ the spatial filter and \mathbf{D} a matrix with \mathbf{d} on the diagonal. Matrix \mathbf{D} is learned by solving the sw-SVM optimization problem:

$$\begin{aligned} & \min_{\mathbf{w}, b, \xi, \mathbf{D}} \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{p=1}^P \xi_p \\ \text{subject to } & y_p (\langle \mathbf{w}, \mathbf{D} \mathbf{x}_p \rangle + b) \geq 1 - \xi_p \quad \text{and} \quad \xi_p \geq 0 \quad \forall p \in \{1, \dots, P\} \\ & \text{and} \quad \sum_{s=1}^S d_s^2 = 1 \quad \forall s \in \{1, \dots, S\}, \end{aligned} \quad (2)$$

where $\mathbf{w} \in \mathbb{R}^{d \times 1}$ is the normal vector, $b \in \mathbb{R}$ is an offset, ξ_p are slack variables that ensure a solution in case data are not linearly-separable, and C is the regularization parameter that controls the trade-off between a low training error and a large margin.

By setting to zero the derivatives of the partial associated Lagrangian according to the primal variables \mathbf{w} , b and ξ_p the optimization problem of the dual formulation can be written as :

$$\min_{\mathbf{D}} J(\mathbf{D}) \quad \text{subject to} \quad \sum_{s=1}^S d_s^2 = 1. \quad (3)$$

with

$$J(\mathbf{D}) = \begin{cases} \max_{\alpha} \mathbf{1}^T \alpha - \frac{1}{2} \alpha^T \mathbf{Y}^T \mathbf{X}^T \mathbf{D}^T \mathbf{D} \mathbf{X} \mathbf{Y} \alpha \\ \text{subject to} \quad \mathbf{y}^T \alpha = 0 \\ \text{and} \quad 0 \leq \alpha_p \leq C \quad \forall p \in \{1, \dots, P\}, \end{cases}$$

with $\boldsymbol{\alpha}$ is the vector of Lagrangian multipliers, $\mathbf{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_P\}$, $\mathbf{y}^T = \{y_1, \dots, y_P\}$ and $\mathbf{Y} = \text{Diag}(\mathbf{y})$. The optimization problem, as stated in (3), is inherently non-convex since the boundary of the unit ball given by $\{\sum_{s=1}^S d_s^2 = 1\}$ is not a convex set. To remove this non-convexity, we perform a variable substitution $\tilde{\mathbf{D}} = \mathbf{D}^T \mathbf{D}$. We arrive at the following optimization problem which is convex in $\tilde{\mathbf{D}}$:

$$\min_{\tilde{\mathbf{D}}} J(\tilde{\mathbf{D}}) \quad \text{subject to} \quad \sum_{s=1}^S \tilde{d}_s = 1. \quad (4)$$

This is clearly an instance of the Multiple Kernel Learning (MKL) problem proposed in [12] where one homogeneous degree one polynomial kernel is used over electrode s samples and \tilde{d}_s is its corresponding positive mixing coefficient. Authors of [12] prove that the search for the optimal $\tilde{\mathbf{D}}$ is convex, yielding fast convergence toward the optimal conditional solution. Hence, the optimization problem can be solved efficiently using a gradient descent as in SimpleMKL [12]. For effect of the ℓ_1 constraints over \tilde{d}_s , the sought spatial filters will be sparse. Linear sw-SVM can be extended to a non-linear sw-SVM by replacing inner products with a suitable kernel.

To solve this sw-SVM optimization problem, we proceed by alternating the search for $\boldsymbol{\alpha}$ and $\tilde{\mathbf{D}}$. For $\tilde{\mathbf{D}}$ fixed, the problem is reduced to a ℓ_1 soft margin SVM with the only difference being that \mathbf{x}_p is replaced by $\tilde{\mathbf{D}}\mathbf{x}_p$ in the inequality constraint. Primal and dual objective functions of SVM are convex, and their solution is obtained by any of the available SVM solver [13]. For the so obtained $\boldsymbol{\alpha}$, the minimization problem over $\tilde{\mathbf{D}}$ is smooth and convex [14]. Hence, we can use a reduced gradient method which converges for such functions [15]. Once the gradient of $J(\tilde{\mathbf{D}})$ is computed, $\tilde{\mathbf{D}}$ is updated by using a descent direction ensuring that the equality constraint and the non-negativity constraints on $\{\tilde{d}_s\}$ are satisfied. These two steps are iterated until a stopping criterion is reached. The stopping criterion we chose is based on a norm variation of the sensor weights.

2.4. Ensemble of sw-SVM classifiers

As seen above, a way to reduce EEG variability is to perform signal averaging across time. Another way to reduce this influence, from a classification point of view, is to use an ensemble of classifiers. This technique has proven efficient in [16, 17] where a set of classifiers were learned on different subsets of a given training set. Classifier outputs were then averaged. In this work,

we propose to learn an ensemble of sw-SVMs, each one on a given temporal feature and to process a weighted average on sw-SVM outputs. Weights are used to keep significant classifiers and discard insignificant ones.

Weights are set as the product of two functions growing proportionally with the accuracies of the two (positive and negative) classes, evaluated on a validation set. If TP , FP , TN and FN hold for True Positive, False Positive, True Negative and False Negative of a given sw-SVM classifier, respectively, then Positive Predictive Value (PPV) and Negative Predictive Value (NPV) can be defined as :

$$PPV = \frac{TP}{TP + FP}, \quad NPV = \frac{TN}{FN + TN},$$

and weights associated to the classifier are such as:

$$\begin{cases} 0 & \text{if } PPV < 0.5 \text{ or } NPV < 0.5 \\ \tan(2(PPV - 0.5)) \times \tan(2(NPV - 0.5)) & \text{otherwise} \end{cases} \quad (5)$$

This weighting strategy is ideal for unbalanced data sets since it privileges classifiers that jointly present good accuracies for both classes.

3. Experiment

The proposed method was evaluated on a visual feedback Error Related Potentials (ErrP) [18] experiment. Eight healthy volunteers (including three women) participated in this experiment. All subjects were BCI-naive at the time of the experiment. Subjects had to retain the position of an ensemble of two to nine digits. The digits were displayed as a sequence in square boxes and evenly distributed along a circle. When the sequence disappeared, a target digit was shown and subjects were asked to click on the box where it previously appeared. A visual feedback indicates whether the answer was wrong or correct.

The experiment involved two sessions that lasted together approximately half an hour. Each session consisted of six blocks of six trials, for a total of $6 \times 6 \times 2 = 72$ trials.

The temporal order of each trial, illustrated in Figure 1, is detailed next. The score, initially zero, was displayed for 3000 ms followed by a fixation cross, which was in turn displayed for 3000 ms. Then the memorization sequence started with variable duration depending on the number of digits the



Figure 1: Temporal diagram of one ErrP trial.

subject had to memorize. When it ended the subject was asked to click on the box where the target digit had appeared. Once the subject had answered the interface was paused for 1500 ms and then turned the clicked box into green upon a correct answer or into red upon an erroneous answer. This feedback lasted for 2500 ms. The 1500 ms preceding the feedback was introduced to avoid any contamination of ErrP by beta rebound motor phenomena linked to mouse clicking [19]. The subject was then asked to report if the feedback (correct/error) matched his expectation by a mouse click ("yes"/"no"). Following his answer a random break of 1000 ms to 1500 ms preceded the beginning of the new trial. The number of digits was adapted with an algorithm tuned to allow about 20% errors for all subjects. The mean error rate (standard deviation) was equal to $17.87(\pm 4.64)\%$ of the trials.

Recordings of the EEG were made using 31 sensors from the extended 10/20 system. The Mitsar 202 DC EEG amplifier was used for signal acquisition. We used one ground positioned on the forehead and two earlobe references. The acquisition software links the two references digitally and computes one common reference for all cephalic sensors. During acquisition, EEG was band-pass filtered in the range 0.1 – 70 Hz and digitized at 500 Hz.

4. Experimental results

4.1. Preprocessing

Raw EEG potentials were first re-referenced to the common average by subtracting from each sensor the average potential (over the 31 sensors) for each time sample. Many studies report two peaks, Ne and Pe, as the main components of Error Related Potential components [18]. Ne shows up about 250 ms after the response as a sharp negative peak and Pe shows up about 300 to 500 ms after the response as a broader positive peak. According to these knowledge, only a window of 1000 ms posterior to the stimulus has been considered for each trial, which results in 500 samples per sensor. A 1 – 10 Hz 4th order Butterworth filter was applied as error related potentials

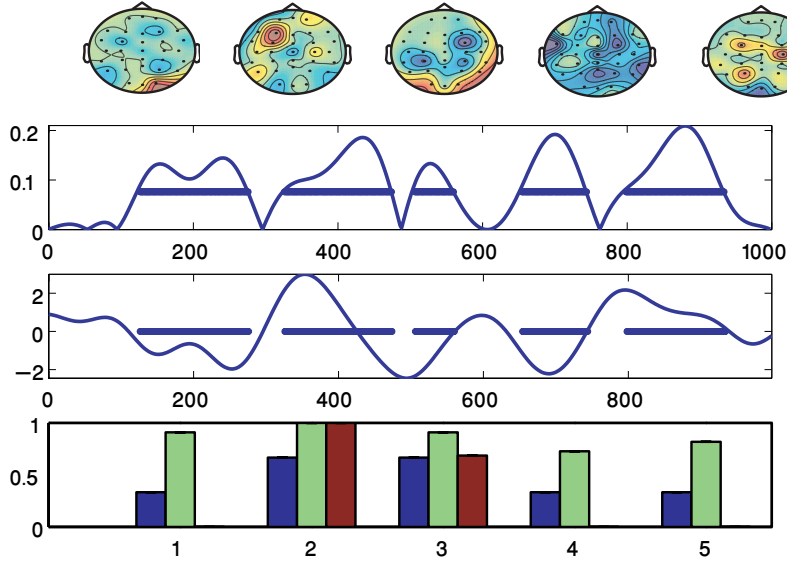


Figure 2: Top : GFP computed on the difference of the grand average error-minus-correct for 1s trials, selected intervals and topographies associated. Middle : the difference computed on electrode FCz. Bottom : accuracies for error (blue bar) and correct (green bar) classes and sw-SVM associated weights (red bar, normalized between 0 and 1).

are known to be a relatively slow cortical potential. No artifact rejection algorithm was applied and all trials were kept for analysis.

4.2. Cross Validation

Single trial classification of ErrPs was assessed using a 5-Cross Validation technique. Because of the limited number of trials, temporal and spatial extraction were performed on four subsets. sw-SVM was trained with a given regularization parameter C (selected from a set of pre-defined values of C) and weighted using (5). Performances were computed on the remaining subset. For each value of C , this process was repeated five times for a given subject and averaged. The highest average accuracy was reported.

4.3. Results

Figure 2 shows the average of the difference error-minus-correct for channel FCz of subject 7. It also reports GFP computed on the difference average. Five components are to be noted. A negative deflection can be seen around 240ms after the feedback and a second positive component occurs about

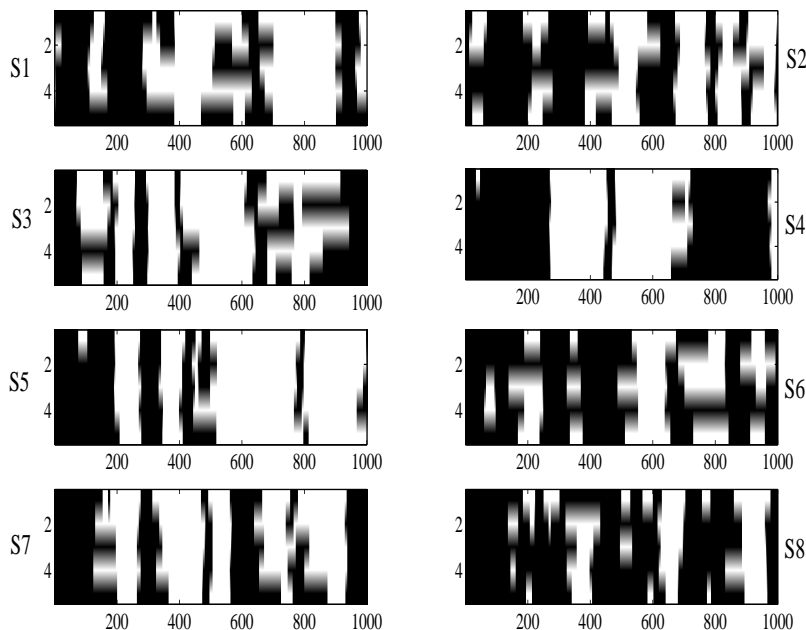


Figure 3: Selected time intervals are shown in white pixels, for each of the 8 subjects and 5 partitions. Each matrix refers to a subject where columns hold time-course and rows hold partitions.

350ms. Three more peaks are also detected; a negative deflection around 500ms, a less pronounced negative deflection around 700ms and a small positive deflection around 800ms.

Scalp potentials topographies associated with the five extracted temporal features are also shown on Figure 2. The first negative peak seems to be occipital whereas the second positive peak covers a rather fronto-central area. The third peak covers a parieto-central area, the fourth peak covers the whole right hemisphere and the last one is more central.

Figure 2 shows accuracies for error and correct classes for each sw-SVM and their corresponding weights (normalized between 0 and 1). Only sw-SVMs learned on the most pronounced peaks (second and third) show good accuracies in both classes and are thus retained.

An important question is whether time interval selection, found by GFP, are consistent across different partitions of the data and across subjects.

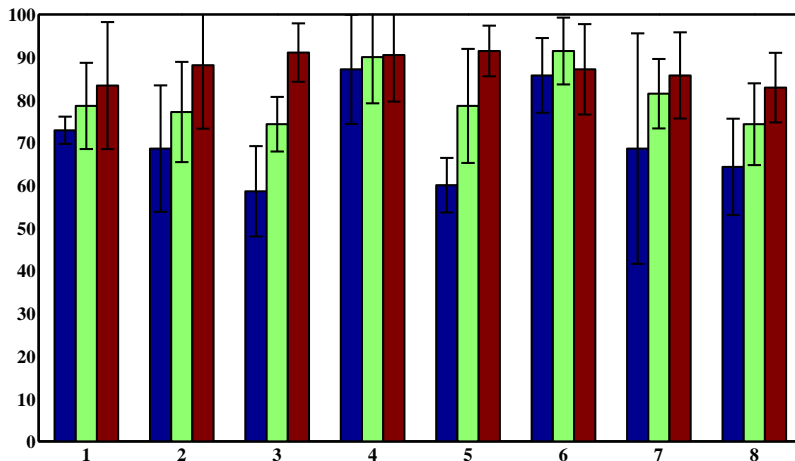


Figure 4: From left to right: performances of classical SVM, sw-SVM and the proposed method for the 8 subjects. Mean (std) accuracies across the 8 subjects are of 70.71(10.77), 80.71(6.61) and 87.80(3.63) respectively.

Figure 3 shows, for each subject, and each of the five partitions, temporal intervals (in white) selected on the ErrP data set. Because of the very small number of trials used in each partition, some inter-partition differences can be noted in these data, but overall, the procedure appears robust and meaningful. Latencies, thus selected time intervals, are different from subject to subject, which is not surprising. However, for almost all subjects, an important activity is noted between 400 and 600ms. These findings confirm those of [18] where a negative deflection, following an incorrect visual feedback of a time-production task, peaked at 330ms with a duration of 260ms. This witnesses in favor of the effectiveness and the consistency of the proposed temporal feature extraction.

Concerning classification results, figure 4 shows the 5-Cross Validation performance provided by a classical SVM approach where all electrodes were used, the sw-SVM where only one spatial filter was used on the whole trial duration and the proposed method where spatio-temporal features were extracted. The proposed method proved constantly superior to SVM and sw-SVM. A paired Wilcoxon signed-rank test was evaluated to compare the proposed method to SVM and sw-SVM and p-values of 0.0078 and 0.0391 were obtained. We conclude that, inclusion of temporal features along with

learning an ensemble of classifiers, provide with superior performance.

5. Conclusion

Spatio-temporal feature identification was addressed. An analysis of Global Field Power highlighted time periods of interest where effects are likely to be the most robust yielding to a data-driven temporal feature extraction. For each temporal feature, a spatial filter was learned jointly with a classifier in the SVM theoretical framework. Spatial filters were learned to optimize classification performance. A weighted averaging on the so obtained ensemble of classifiers yielded to a robust final decision function. Experimental results on Error-related Potentials illustrate the efficiency of the method from a physiological and a machine learning points of view. Further research may extract all relevant aspects of brain post-stimulus dynamics recorded in EEG (spatio-temporal-frequential).

Acknowledgment

This work has been supported through the project OpenViBE2 of the ANR (National Research Agency), France. Authors would like to acknowledge Sandra Rousseau from GIPSA-Lab Grenoble INP for providing the data.

References

- [1] G. Dornhege, M. Kraudelat, K.-R. Müller, B. Blankertz, Towards Brain-Computer interfacing, MIT Press.
- [2] P. Nunez, R. Srinivasan, Electric Field of the Brain, New York: Oxford Univ Press, 2nd edition, 2006.
- [3] B. Blankertz, S.Lemm, M. Treder, S. Haufe, K. Müller, Single-trial analysis and classification of ERP components – a tutorial, *NeuroImage* (2010). In press.
- [4] J. Müller-Gerking, G. Pfurtscheller, H. Flyvbjerg, Designing optimal spatial filters for single-trial EEG classification in a movement task, *Clin. Neurophys.* 110 (1999) 787–798.

- [5] C. Gouy-Pailler, M. Congedo, C. Brunner, C. Jutten, G. Pfurtscheller, Nonstationary brain source separation for multiclass motor imagery., *IEEE Transactions on Biomedical Engineering* 57 (2010) 469–478.
- [6] B. Rivet, A. Souloumiac, V. Attina, G. Gibert, xDAWN Algorithm to Enhance Evoked Potentials: Application to Brain Computer Interface., *IEEE Trans Biomed Eng* (2009) 2035–2043.
- [7] U. Hoffmann, J. Vesin, T. Ebrahimi, Spatial filters for the classification of event-related potentials, in: *Proceedings of ESANN 2006*, pp. 47–52.
- [8] K. Müller, M. Krauledat, G. Dornhege, G. Curio, B. Blankertz, Machine learning techniques for brain-computer interfaces, *Biomed Tech* 49 (2004) 11–22.
- [9] M. D. Green, J. Swets, *Signal detection theory and psychophysics*, Krieger, Huntington, NY, 1966.
- [10] D. Lehmann, W. Skrandies, Reference-free identification of components of checkerboard-evoked multichannel potential fields, *Electroencephalogr Clin Neurophysiol* 48 (1980) 609–21.
- [11] N. Jrad, M. Congedo, R. Phlypo, S. Rousseau, R. Flamary, F. Yger, A. Rakotomamonjy, sw-SVM : sensor weighting Support Vector Machines for EEG-based Brain-Computer Interfaces, *Journal of Neural Engineering* in press (2011).
- [12] A. Rakotomamonjy, F. Bach, S. Canu, Y. Grandvalet, SimpleMKL, *Journal of Machine Learning Research* 9 (2008) 2491–2521.
- [13] S. Canu, Y. Grandvalet, V. Guigue, A. Rakotomamonjy, SVM and Kernel Methods Matlab Toolbox, *Perception Systèmes et Information*, INSA de Rouen, Rouen, France, 2005.
- [14] C. Lemaréchal, C. Sagastizábal, Practical aspects of the moreau–yosida regularization: Theoretical preliminaries, *SIAM J. on Optimization* 7 (1997) 367–385.
- [15] D. Luenberger, *Linear and Nonlinear Programming*, Addison-Wesley, 1984.

- [16] Y. Grandvalet, Bagging equalizes influence, *Machine Learning* 55 (2004) 251–270.
- [17] A. Rakotomamonjy, V. Guigue, BCI Competition III : Dataset II - Ensemble of SVMs for BCI P300 speller, *IEEE Trans. Biomedical Engineering* 55 (2008) 1147–1154.
- [18] W. Miltner, C. Braun, M. Coles, Event-related brain potentials following incorrect feedback in a time-estimation task: Evidence for a generic neural system for error detection., *Journal of Cognitive Neuroscience* 9 (1997) 788–798.
- [19] M. Jurkiewicz, W. Gaetz, A. Bostan, D. Cheyne, Post-movement beta rebound is generated in motor cortex: Evidence from neuromagnetic recordings, *NeuroImage* 32 (2006) 1281–1289.