A new feature and associated optimal spatial filter for EEG signal classification: Waveform Length

Fabien Lotte

A new feature and associated optimal spatial filter for EEG signal classification: Waveform Length

Fabien LOTTE
Inria Bordeaux Sud-Ouest, France
fabien.lotte@inria.fr

Abstract

In this paper, we introduce Waveform Length (WL), a new feature for ElectroEncephaloGraphy (EEG) signal classification which measures the signal complexity. We also propose the Waveformlength Optimal Spatial Filter (WOSF), an optimal spatial filter to classify EEG signals based on WL features. Evaluations on 15 subjects suggested that WOSF with WL features provide performances that are competitive with that of Common Spatial Patterns (CSP) with Band Power (BP) features, CSP being the optimal spatial filter for BP features. More interestingly, our results suggested that combining WOSF with CSP features leads to classification performances that are significantly better than that of CSP alone (80% versus 77% average accuracy respectively).

1. Introduction

Brain-Computer Interfaces (BCI) are communication systems that enable users to send commands to a computer by using only their brain activity [8], this activity being generally measured using ElectroEncephaloGraphy (EEG). Most EEG-based BCI are designed around a pattern recognition approach: In a first step, features describing the relevant information contained in EEG signals are extracted [2]. They are then used as input to a classifier in order to identify the class of the mental state [2]. Therefore, the efficiency of a BCI, in terms of recognition rate, depends mostly on the choice of appropriate features and classifiers.

For BCI based on mental tasks, e.g., motor imagery (imagination of limb movements) [8], one of the most popular and efficient feature is Band Power (BP) [2], i.e., the EEG signal power in a given frequency band. Their efficiency has been further increased by the design of the Common Spatial Patterns (CSP) algorithm, an optimal spatial filter for EEG classification based on BP features [3]. Using CSP for spatial filtering and BP as features has become a gold standard to design BCI-based on mental tasks [2]. For instance, this setup was the basis of most of the winning entries of the last BCI competitions for motor imagery data sets (http://www.bbci.de/competition/). Despite the availability of such algorithms, current BCI performances are still far from being satisfactory, and the BCI community keeps stressing the need to further explore and design alternative features for improved performance and robustness [6][9].

In this paper, we introduce such an alternative feature: Waveform Length (WL). WL measures the length of a given waveform, which is also a measure of the signal complexity [10]. WL was initially designed to classify ElectroMyoGraphy (EMG) signals, and has been proved to be one of the most robust and efficient feature for this task [10]. As both EEG and EMG measure an electrical signal resulting from the activity of neuron populations (cortical neurons for EEG, motor ones for EMG), it seems promising to explore whether a feature that can successfully classify EMG could also successfully classify EEG. Finally, since spatial filtering is a key element in EEG signal analysis [3], we also propose an algorithm to obtain optimal spatial filters to classify EEG signals based on WL features.

This paper is organized as follows: Section 2 gives more details about BP features and CSP. Then, Section 3 describes WL features and the algorithm we propose to obtain optimal spatial filters for such features. They are evaluated and compared with CSP and BP features in Section 4, on 2 different data sets (a motor imagery data set and a mental rotation one). Finally, a conclusion is provided in Section 5.

2 Band Power features and CSP

A BP feature is defined as the EEG signal power for a given channel and frequency band. A classical way to obtain such a feature is to 1) band-pass filter the signal in this given frequency band, 2) square it, 3) average it over a given time window and 4) take its logarithm, to make the feature distribution more normal-like [8]. Formally, bp, the BP feature of a band-pass filtered signal $x = (x_1, \dots, x_N)$ can be obtained as follows:

$$bp = \log(\frac{1}{N}\sum_{i=1}^{N} x_i^2) = \log(\frac{1}{N}\|x\|_2^2) = \log(var(x)) \tag{1}$$

with $\|\|_2$ being the $l2$-norm. The last part of this equation assumes a zero mean for EEG signals, which is usually the case after band-pass filtering.

Due to volume conduction, EEG signals inherently have a low spatial resolution, and the relevant information they contain is generally spread over several channels. To alleviate this issue and improve the signal-to-noise ratio, spatial filtering algorithms such as CSP have been proposed. CSP aims at learning spatial filters which can maximize the variance of band-pass filtered EEG signals from one class while minimizing the variance of those from the other class [3]. As the variance of EEG signals filtered in a given frequency band corresponds to the signal power in this band (see Eq. 1), CSP aims at achieving optimal discrimination for BCI based on BP features. Formally, CSP uses the spatial filters w which extremize the following function:

$$J_{CSP}(w) = \frac{\|w^T X_1\|_2^2}{\|w^T X_2\|_2^2} = \frac{w^T X_1^T X_1 w}{w^T X_2^T X_2 w} = \frac{w^T C_1 w}{w^T C_2 w}$$

where T denotes transpose, X_i is the training band-pass filtered signal matrix for class i (with the samples as rows and the channels as columns) and C_i the spatial covariance matrix from class i. In practice, the covariance matrix C_i is defined as the average covariance matrix of each trial from class i [3]. $J_{CSP}(w)$ appears to be a generalized Rayleigh quotient, and as such, the spatial filters w that maximize or minimize it are the eigenvectors corresponding to the largest and lowest eigenvalues, respectively, of the Generalized EigenValue Decomposition (GEVD) of matrices C_1 and C_2. Once the spatial filters w_i obtained, extracting feature bp_i is simply achieved by computing the BP feature of the EEG signals X after spatial projection, in other words:

$$bp_i = \log(w_i^T X^T X w_i) = \log(var(w^T X))$$

3 WL and its optimal spatial filter

The WL feature that we introduce in this paper can be extracted from an EEG signal x as follows [10]:

$$wl = \log(\textstyle\sum_{i=1}^{N-1} |x_{i+1} - x_i|) = \log(\displaystyle\sum_{i=1}^{N-1} |\Delta x_i|) = \log(\|\Delta x\|_1)$$

with $|x|$ being the absolute value of x, and $\|\|_1$ the $l1$-norm. This feature measures the cumulative length of the EEG signal analyzed. To maximize the efficiency of this feature for EEG classification, it should be extracted after appropriate spatial filtering, in the same way as BP efficiency is maximized by CSP spatial filtering. Therefore, we also propose a spatial filter that is optimal for classification based on WL features. We denote this new spatial filter Waveformlength Optimal Spatial Filter (WOSF). In order to derive such an algorithm, we have to find spatial filters w which maximize the waveform length of spatially projected EEG signals from one class, while minimizing it for the other class. Formally, this means extremizing the following objective function:

$$\begin{aligned} J_{WOSF1}(w) &= \frac{\|w^T X_1^{2:N} - w^T X_1^{1:(N-1)}\|_1}{\|w^T X_2^{2:N} - w^T X_2^{1:(N-1)}\|_1} \\ &= \frac{\|w^T \Delta X_1\|_1}{\|w^T \Delta X_2\|_1} \end{aligned}$$

with $\Delta X = X^{2:N} - X^{1:(N-1)}$ and $X^{i:j}$ being the signal matrix X with only rows i to j, i.e., with only EEG samples from indexes i to j. Unfortunately, the $l1$-norm is not differentiable. This makes the optimization of J_{WOSF1} unconvenient, iterative, complex and computationally expensive. Therefore, we decided to optimize the spatial filters using the $l2$-norm rather than the $l1$-norm, which, as we will see later on, leads to a closed-form and computationally efficient solution, similar to that of CSP. Thus, our objective function becomes:

$$\begin{aligned} J_{WOSF2}(w) = \frac{\|w^T \Delta X_1\|_2}{\|w^T \Delta X_2\|_2} &= \frac{w^T \Delta X_1^T \Delta X_1 w}{w^T \Delta X_2^T \Delta X_2 w} \\ &= \frac{w^T D_1 w}{w^T D_2 w} \end{aligned}$$

with $D_i = \Delta X_i^T \Delta X_i$. This is again a generalized Rayleigh quotient, and as such, the spatial filters which maximizes or minimizes J_{WOSF2} are the eigenvectors corresponding to the largest and lowest eigenvalues obtained by GEVD of matrices D_1 and D_2. As for CSP, the ΔX_i matrices used in practice are the average ΔX matrices computed for each trial of class i. It is worth noting here the similarities between CSP and WOSF. Indeed, both spatial filters are obtained using GEVD of two matrices, their difference lying in the definition of these two matrices. Once the WOSF spatial filters are obtained, extracting feature wl_i for the i^{th} spatial filter w_i is simply achieved as follows:

$$wl_i = \log(\|w_i^T \Delta X\|_1)$$

4 Evaluation

We compared WL features and WOSF to BP features and CSP on 15 subjects, from two different EEG data sets. This section describes these data sets, the evaluation methodology and the obtained results.

4.1 Data set 1: Motor Imagery

We used data set IIa [7] from BCI competition IV (http://www.bbci.de/competition/iv/). It comprises EEG signals from 9 subjects who performed left hand, right hand, foot and tongue Motor Imagery (MI). EEG signals were recorded using 22 electrodes located above motor areas. For the purpose of this study, only EEG signals corresponding to left and right hand MI were used. A training and a testing set were available for each subject, both sets containing 72 trials for each class. For this data set, performances were measured by optimizing the spatial filters and classifier on the training set, and using them to predict the labels of the test set.

4.2 Data set 2: Mental Rotation

This data set was collected in-house from 6 subjects during Mental Rotation (MR) tasks. The protocol used was similar to that of data set 1, except that on cue presentation, instead of performing MI tasks, subjects were instructed to either imagine continuous rotations of a 3D geometric figure displayed on screen or to relax while fixating a dot displayed in the screen center. EEG were collected using 15 electrodes (C3, C1, Cz, C2, C4, F3, F1, Fz, F2, F4, P3, P1, Pz, P2, P4). Each subject participated to 4 runs, a run comprising 20 trials from each class (relax and MR), except subject B3, who participated to 3 runs only, due to fatigue. Due to the smaller number of trials per subject in this data set, performances were evaluated using leave-one-run-out cross-validation, i.e., the spatial filters and classifier were optimized on the EEG signals from all runs except one, and tested on this remaining run. The process was repeated by using each run as the testing run. The performance measure is the classification accuracy averaged over each testing run.

4.3 Methods

All EEG signals were band-pass filtered in 8-30Hz, using a 5^{th} order Butterworth filter. Indeed, this frequency band has been shown to be suitable to classify EEG signals corresponding to both MI and MR tasks [4]. For both data sets, for each trial, we extracted the features from the time segment located from 0.5s to 2.5s after the cue instructing the subject to perform a mental task (as done by the winner of BCI competition IV, data set IIa). With both CSP and WOSF, we used 3 pairs of spatial filters for feature extraction, as recommended in [3] for CSP. The 6 features hence extracted were classified using Linear Discriminant Analysis, one of the most efficient classifier for BCI design [2].

We compared the performance of BP and WL features extracted from all available channels, and their performance after CSP and WOSF spatial filtering, respectively. Since we expected WL features to extract a different information from EEG signals than BP features, we also explored whether combining them could further improve performances. Thus, we concatenated into a single feature vector BP and WL features extracted after CSP and WOSF spatial filtering respectively. To keep a similarly low dimensionality despite the feature concatenation, we used only 2 pairs of spatial filters for both CSP and WOSF, i.e., 8 features in total, rather than 3 pairs when they were used alone.

4.4 Results

The results described in Table 1 highlight several interesting points. First, they confirmed the need for spatial filtering in EEG-based BCI, with CSP and WOSF clearly outperforming BP and WL without spatial filtering. Then, they showed that WL outperformed BP on the MR signals, whereas it obtained lower performance than BP on the MI signals. Overall, a paired t-test revealed no significant difference between them ($p > 0.05$). More interesintgly, WOSF reached similar performance than CSP on the MR data, and outperformed it on the MI data. This difference was not significant though, maybe due to the limited number of subjects. Even more interestingly, our results suggested that combining WOSF together with CSP leads to the best classification accuracy, 80% on average, versus 77% for CSP alone, this difference being statistically significant ($p < 0.05$). 11 out of 15 subjects reached their best performance using CSP and WOSF combined. This suggests that WL and WOSF is a valuable feature for EEG classification as well as a robust alernative and complement to CSP-based features.

To try to understand why WOSF might be more efficient than or complementary to CSP, it is worth looking at the spatial filters obtained (see Figure 1). This figure reveals that CSP and WOSF filters are generally rather similar (see, e.g., subject A1), although spatial filters leading to higher performances tend to be spatially smoother and more focused. It should be mentioned that an EEG power increase also means an EEG waveform length increase, which could explain the spa-

Table 1. Classification accuracy (%) obtained by each method on the two data sets

	Motor Imagery Data Set										Mental Rotation Data Set						
subject	**A1**	**A2**	**A3**	**A4**	**A5**	**A6**	**A7**	**A8**	**A9**	**Mean**	**B1**	**B2**	**B3**	**B4**	**B5**	**B6**	**Mean**
BP	69,4	54,2	84	63,8	59,7	59	59	84,7	85,4	68,8	71,3	50	60	65	80	73,8	66,7
WL	70,8	50,7	63,9	59,7	50	59	58,3	84,7	87,5	65	81,9	53,8	67,5	62,5	80	70,4	69,3
CSP	93,1	50	**96,5**	70,8	56,9	68,8	80,6	93,8	92,4	78,1	93,8	59,4	68,3	66,9	85	79,4	75,5
WOSF	92,4	**55,6**	95,1	71,5	**74,3**	70,1	**83,3**	**95,8**	91,7	**81,1**	98,8	**60,6**	70,8	58,1	85,6	76,9	75,1
CSP+ WOSF	**94,4**	51,4	94,4	**73,6**	72,2	**70,8**	**83,3**	**95,8**	**93,1**	81	**100**	57,5	**75**	**69,4**	**88,1**	**82,5**	**78,8**

tial filter similarities. Naturally, the performance difference between CSP and WOSF are also due to the different information extracted from the signals as well as to the use of the l1-norm for WL rather than the l2-norm for BP. Indeed, the l2-norm squaring the EEG signal values, it tends to amplify the influence of high amplitude artifacts more than the l1-norm.

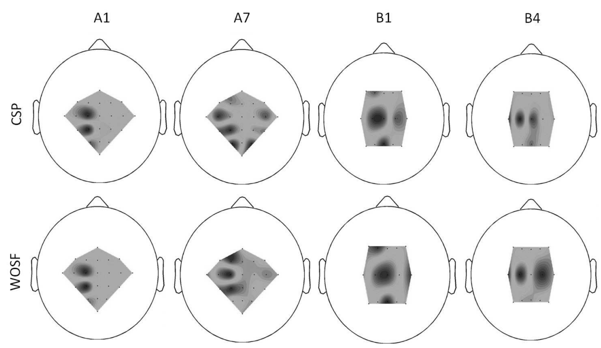

Figure 1. CSP and WOSF spatial filters

5 Conclusion

In this paper, we introduced the waveform length as a new feature for EEG classification, as well as an optimal spatial filter for classification based on such features. Our evaluations suggested that these WL and WOSF algorithms are simple to use and implement, computationally efficient and provide competitive performance with CSP. Combining WOSF with CSP even significantly outperformed CSP alone. Thus, this approach can potentially become a new and useful tool in the feature repertoire of EEG-based BCI designers.

The work presented in this paper opens the door to many potential future works. For instance, WOSF being based on the same optimization framework as CSP, namely GEVD, it could benefit from several works extending CSP, such as regularization [5], filter-bank approaches [1] or optimization based on the l1-norm [12]. It would also be interesting to study any potential theoretical link between WL features and time-domain parameters (TDP) [11], TDP being defined as the variance of EEG signal derivatives. This highlights the quick evolution potential of WL and WOSF approaches.

References

[1] K. Ang, Z. Chin, H. Zhang, and C. Guan. Filter bank common spatial pattern (FBCSP) in brain-computer interface. In *IEEE IJCNN*, pages 2391–2398, 2008.

[2] A. Bashashati, M. Fatourechi, R. K. Ward, and G. E. Birch. A survey of signal processing algorithms in brain-computer interfaces based on electrical brain signals. *J Neur Eng*, 4(2):R35–57, 2007.

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

[4] E. Friedrich, R. Scherer, and C. Neuper. The effect of distinct mental strategies on classification performance for braincomputer interfaces. *Int J Psychophys*, 2012.

[5] F. Lotte and C. Guan. Spatially regularized common spatial patterns for EEG classification. In *ICPR*, 2010.

[6] D. J. McFarland, C. W. Anderson, K.-R. Müller, A. Schlögl, and D. J. Krusienski. BCI meeting 2005-workshop on BCI signal processing: feature extraction and translation. *IEEE Trans Neur Sys Rehab*, 2006.

[7] M. Naeem, C. Brunner, R. Leeb, B. Graimann, and G. Pfurtscheller. Seperability of four-class motor imagery data using independent components analysis. *J Neur Eng*, 3:208–216, 2006.

[8] G. Pfurtscheller and C. Neuper. Motor imagery and direct brain-computer communication. *proc of the IEEE*, 89(7):1123–1134, 2001.

[9] The Future BNCI consortium. Future BNCI: A roadmap for future directions in brain / neuronal computer interaction. Technical report, 2012.

[10] D. Tkach, H. Huang, and T. Kuiken. Study of stability of time-domain features for electromyographic pattern recognition. *J NeuroEng Rehab*, 7(1):21, 2010.

[11] C. Vidaurre, N. Krämer, B. Blankertz, and A. Schlögl. Time domain parameters as a feature for EEG-based brain computer interfaces. *Neural Networks*, 22, 2009.

[12] H. Wang, Q. Tang, and W. Zheng. L1-norm-based common spatial patterns. *IEEE Trans Biomed Eng*, 59(3):653 –662, march 2012.