Motor Imagery Classification for BCI Using Common Spatial Patterns and Feature Relevance Analysis

Luisa F. Velásquez-Martínez, A.M. Álvarez-Meza, and C.G. Castellanos-Domínguez*

Universidad Nacional de Colombia sede Manizales,
Signal Processing and Recognition Group, Universidad Nacional de Colombia, Campus La Nubia, km 7 via al Magdalena, Manizales-Colombia.
{lfvelasquezma,amalvarezme,cgcastellanosd}@unal.edu.co

Abstract. Recently, there have been many efforts to develop Brain Computer Interface (BCI) systems, allowing to identify and discriminate brain activity. In this work, a Motor Imagery (MI) discrimination framework is proposed, which employs Common Spatial Patterns (CSP) as preprocessing stage, and a feature relevance analysis approach based on an eigendecomposition method to identify the main features that allow to discriminate the studied EEG signals. The CSP is employed to reveal the dynamics of interest from EEG signals, and then we select a set of features representing the best as possible the studied process. EEG signals modeling is done by feature estimation of three frequency-based and one time-based. Besides, a relevance analysis over the EEG channels is performed, which gives to the user an idea about the channels that mainly contribute for the MI discrimination. Our approach is tested over a well known MI dataset. Attained results (95.21 ± 4.21 [$\%$] mean accuracy) show that presented framework can be used as a tool to support the discrimination of MI brain activity.

Keywords: Motor Imagery, Common Spatial Patterns, Feature Relevance Analysis.

1 Introduction

The electroencephalography (EEG) is the most commonly employed method for monitoring brain activity and it has been used for several applications, such as: epilepsy detection, analysis of cognitive behaviors, game controlling, among others. Brain Computer Interfaces (BCI) take advantage of the extracted information from EEG signals to establish a direct communication channel between the human brain and the machine [1]. BCI is used to help people with disability by means of the analysis of the human sensorimotor functions, which are based on the paradigm in cognitive neuroscience named as Motor Imagery (MI). However, the analysis of the EEG signals requires to develop suitable preprocessing, feature representation, feature selection, and classification methodologies to improve the performance of real-world BCI applications. Regarding to the preprocessing stage, Common Spatial Pattern (CSP) is a popular algorithm for MI-based BCI systems [1]. CSP method constructs spatial filters that

* Under grants provided by a Msc. scholarship funded by Universidad Nacional de Colombia, and by a PhD. scholarship and the project 111045426008 funded by Colciencias.

J.M. Ferrández Vicente et al. (Eds.): IWINAC 2013, Part II, LNCS 7931, pp. 365–374, 2013.

maximize the variance of one kind of task and simultaneously minimize the variance of another. In order to achieve high classification accuracy, a pre-filtered broad band or subject-specific frequency bands are fixed to highlight the dynamics of interest. To find such optimal bands, several algorithms have been proposed, such as: Common Spatio-Spectral Pattern, Sub-band CSP and Filter Bank Common Spatial [1,2]. In [1] the CSP preprocessing stage is matched with an Empirical Mode Decomposition (EMD) based method to select informative frequency bands from EEG. However, a direct and an automatic framework that allows to find such bands of interest is still an open issue.

Now, with respect to feature representation methodologies for BCI systems, the attributes are estimated by different methods such as Adaptive Autoregressive (AAR) coefficients, Hjorth parameters, Power Spectral Density (PSD), and continuous and discrete wavelet transforms (CWT and DWT) [3]. Although, many features may be extracted from different methods, several features may not contain relevant information introducing redundancy. Therefore, it is necessary to find a subset of attributes that preserving, as well as possible, the input data variability, allows to identify the most important information that helps to recognize different classes from EEG data. Several approaches have been used to identify the relevance of the computed features in BCI systems [3,4]. Nevertheless, most of these feature selection methods are computationally expensive and they are not able to find directly a measure that relates each feature with its discriminative contribution.

In this work, an MI discrimination framework is proposed, which employs CSP as preprocessing stage, and a feature relevance analysis approach based on an eigendecomposition method to identify discriminative features. The CSP is matched with EMD to reveal the dynamics of interest. Then, we select a set of features representing the best as possible the studied process. For such purpose, a variability analysis is presented to identify relevant features. EEG signals modeling is done by feature estimation of three frequency-based and one time-based. Moreover, a relevance analysis over the EEG channels is performed, which gives to the user an idea about the channels that mainly contribute for the MI discrimination.

2 Materials and Methods

2.1 Preprocessing

Let $\boldsymbol{Y}_r \in \mathbb{R}^{C \times T_Y}$ represents the raw EEG signal of the r-th single trial; being $r = 1, \ldots, n$; C the number of channels, and T_Y the length of the samples. The CSP method is employed to analyze multi-channel EEG data based on recordings from two classes [5], producing spatial filters $\boldsymbol{W} \in \mathbb{R}^{C \times C}$, which project the original signal to a space where the differences in variances of two kinds of tasks can be maximized [1]. The projected signal $\boldsymbol{Z}_r \in \mathbb{R}^{C \times T_Y}$ is given by $\boldsymbol{Z}_r = \boldsymbol{W}\boldsymbol{Y}_r$. Given the projected signal $\boldsymbol{Z}_r$ by CSP, an EMD is performed to find out the main components of each $\boldsymbol{Z}_r$. Thus, EMD decomposes $\boldsymbol{Z}_r$ into a residual and intrinsic modes as $\boldsymbol{Y}_r = \sum_{i=1}^{N} \boldsymbol{c}_i + \epsilon_n$, where $\boldsymbol{c}_i \in \mathbb{R}^{C \times T_Y} : i = 1, \ldots, N$ stands for Intrinsic Mode Functions (IMFs) and $\epsilon_n \in \mathbb{R}^{C \times T_Y}$ indicates a residual. The zero-mean amplitude IMFs are obtained by a sifting process according to the characterizing conditions of the IMFs. The process can be finished when residual becomes a monotonic component or a constant [1]. In this regard, the main bands $\hat{\boldsymbol{Y}}_r \in \mathbb{R}^{C \times T_Y}$ of $\boldsymbol{Z}_r$ can be highlighted by considering N_c IMFs as $\hat{\boldsymbol{Y}}_r = \sum_{i=1}^{N_c} \boldsymbol{c}_i$.

2.2 Feature Representation

From the preprocessed EEG signal matrix $\hat{\boldsymbol{Y}}_r$, the Power Spectral Density (PSD), the Continuous and Discrete Wavelet Transforms (CWT)-(DWT), and the Hjorth parameters are computed for each row vector $\hat{\boldsymbol{y}}_r \in \mathbb{R}^{1 \times T_Y}$ as follows. Let $\boldsymbol{p} = \{p_f : f = 0, \ldots, F_s/2\}$ the PSD of input signal $\hat{\boldsymbol{y}}_r$ that, in the concrete case, is computed by means of the nonparametric Welch's method, being F_s the sample frequency [4]. Particularly, the fast Fourier transform algorithm is employed to estimate the PSD, by dividing the time-series into M overlapped segments of length L, and applying a smooth time weighting window $\boldsymbol{w} = \{w_i : i = 1, \ldots, L\}$, obtaining the windowed segments $\boldsymbol{v}^{(m)} = \{v_i^{(m)} : i = 1, \ldots, L\}$, with $m = 1, \ldots, M$. The main goal is to deal with the non-stationary nature of the EEG, assuming a piece-wise stationarity into each overlapped segment. So, inspired by singular spectrum analysis-based approaches for analyzing one-dimensional time-series, the length of the segments is fixed as $L > F_s/F_r$, with F_r the minimum frequency to be considered within the analysis [6]. Thus, the modified periodogram vector $\boldsymbol{u} = \{u_f : f = 0, \ldots, F_s/2\}$ is calculated by the Discrete Fourier Transform as $u_f = \sum_{m=1}^{M} |\sum_{i=1}^{L} v_i^{(m)} \exp(-j2\pi i f)|^2$. Afterwards, each element of PSD vector $\boldsymbol{p}$ can be computed as $p_f = u_f/(M\,LU)$, with $U = \mathbf{E}\{|w_i|^2 : \forall i \in L\}$, where notation $\mathbf{E}\{\cdot\}$ stands for expectation operator. The motor imagery discrimination analysis is mostly provided for μ ($8-13$ *Hz*) and β ($13-30$ *Hz*) bands. Therefore, their PSD bands (noted as S_μ and S_β, respectively) are calculated from $\boldsymbol{p}$, for which the PSD magnitude is parameterized based on the first and second statistical moments.

Now, regarding to CWT, noted that this inner-product-based transformation quantifies the similarity between a given signal ($\hat{\boldsymbol{y}}_r$) and the considered base function (termed mothers wavelets). Therefore, the wavelet transform of a EEG signal, at time t and frequency f, is provided by their convolution with the scaled and shifted wavelet [4]. The short-time instantaneous amplitude of the CWT of EEG data is accomplished, where two Morlet wavelets centered at the bands of interest (10 *Hz* and 22 *Hz*) to highlight the μ and β bands, respectively. After that, the first and second statistical moments, as well as the maximum value of the coefficients magnitude are estimated; those values are considered as the CWT based features. With respect to DWT, this transformation is assumed to provide a multi-resolution decomposition and non-redundant representation of the input signal $\hat{\boldsymbol{y}}_r$. DWT has a wide application in biomedical signal processing, especially, for non-stationary signals such as EEG [7]. A seventh order Symlet mother wavelet is used, for which the detail coefficients of the third and fourth level are obtained (DWT4 and DWT3) to compute the required frequency bands α and β. Namely, the estimated frequency bands for each wavelet level are $62.5-125$ *Hz*; $31.3-62.5$ *Hz*; $15.7-31.3$ *Hz* (including the β rhythm); and $7.9-15.7$ *Hz*; $0.5-7.9$ *Hz* (including the α rhythm) [4]. From the detail coefficient sets, DWT4 and DWT3, the first and second statistical moments, and the maximum value are estimated.

Lastly, a time-domain based characterization is also employed to describe the EEG data. Particularly, from the input signal $\hat{\boldsymbol{y}}_r$, the following short-time Hjorth parameters are estimated: activity, mobility, and complexity [3]. The activity is directly described by the variance that is related to the signal power, $\sigma^2(\hat{\boldsymbol{y}}_r)$. The mobility is a measure of the signal mean frequency, defined as $\phi(\hat{\boldsymbol{y}}_r) = \sqrt{\sigma^2(\hat{\boldsymbol{y}}_r')/\sigma^2(\hat{\boldsymbol{y}}_r)}$, being $\hat{\boldsymbol{y}}_r'$ the derivative of $\hat{\boldsymbol{y}}_r$. Finally, the complexity measures the deviation of the signal from the sine shape, that is, the change in frequency and it can be computed as $\vartheta(\hat{\boldsymbol{y}}_r) = \phi(\hat{\boldsymbol{y}}_r')/\phi(\hat{\boldsymbol{y}}_r)$. From the estimated short-time Hjorth parameter sets, the first and second statistical moments, and the maximum value are obtained as features.

2.3 Feature Relevance Analysis

From the above mentioned EEG representations, a feature space matrix $\boldsymbol{X} \in \mathbb{R}^{n \times D}$ is obtained, assuming that a set of preprocessed EEG signals $\{\hat{\boldsymbol{Y}}_r : r = 1, \ldots, n\}$ is provided, being n the number of training trails of a given subject in a BCI system, and D the number of estimated features. Particularly, each row, $\hat{\boldsymbol{y}}_r$ of $\hat{\boldsymbol{Y}}_r$ holds the c-th studied EEG channel, with $c = 1, \ldots, n_c$ and being n_c the number of analyzed channels. To carry out a low-dimensional representation of the original feature representation space, this work uses Principal Component Analysis (PCA) as a statistical eigendecomposition, searching for directions with greater variance to project the data. Although, PCA is commonly used as a feature extraction method, it can be useful to properly select a relevant subset of original features that better represent the studied process [8]. In this sense, given a set of features $\varXi = \{\boldsymbol{\xi}_d : d = 1, \ldots, D\}$, where $\boldsymbol{\xi}_d$ corresponds to each column of the input data matrix $\boldsymbol{X}$, the relevance of each feature can be analyzed by the PCA mapping. More precisely, the relevance of $\boldsymbol{\xi}_d$ can be identified by computing the corresponding weighting term $\boldsymbol{\rho} = \{\rho_d : d = 1, \ldots, D\}$, where $\boldsymbol{\rho}$ is defined as $\boldsymbol{\rho} = \mathbf{E}\{|\lambda_d \boldsymbol{\alpha}_d| : \forall d \in D'\}$, being λ_d and $\boldsymbol{\alpha}_d$ the eigenvalues and eigenvectors of the covariance matrix $\boldsymbol{\varSigma} \in \mathbb{R}^{p \times p}$, which is estimated as $\boldsymbol{\varSigma} = \boldsymbol{X}^\top \boldsymbol{X}$. The main assumption is that the largest values of ρ_d point out to the best input attributes, since they exhibit higher overall correlations with principal components. The D' value is fixed as the number of dimensions needed to conserve a percentage of the input data variability.

3 Experiments and Results

In order to test the proposed framework, experimental tests were done over the a well known Motor Imagery (MI) dataset. The EEG data collection is provided by the Berlin Brain-Computer Interface (BCI competition IV 2008 - Data sets 1)[1]. This database is based on the paradigm in cognitive neuroscience of MI, e.g. imagination of hand movements, whole body activities, relaxation, etc. For each subject the first two classes of motor imagery were selected from the three classes left hand, right hand, and foot (side chosen by the subject). The EEG signals were obtained from seven subjects. For each subject, the signals from 59 EEG positions were measured, being the sensorimotor

[1] `http://bbci.de/competition/iv/desc_1.html`

area the most densely covered area by the electrodes. Signals were band-pass filtered between 0.05 and 200 Hz and then digitized at 1000 Hz. Moreover, the database was downsampled at $F_s = 100$ Hz, but previously a low.-pass Chevyshev II filter (order 10) was employed with stopband ripple 50dB down and stopband edge frequency 49 Hz. The whole motor imagery session was performed without feedback. The data base contains 100 repetitions of each MI class per person. Particulary, the EEG segments were extracted while a cue (indicating a side) is presented, i.e. an arrow pointing left or right were presented on a screen, the duration of each extracted segment is $4\ s$ during which the subject was instructed to perform the cued motor imagery task. These periods were interleaved with $2\ s$ of blank screen and $2\ s$ with a fixation cross shown in the center of the screen. All EEG channels per subject of the above mentioned MI dataset were used. We test four frameworks mainly changing the preprocessing stage, in order to validate the performance of the proposed approach (see Fig. 1). The first one does not uses any kind of preprocessing method of the data before the characterization stage, the next two frameworks are conceived to use either CSP or EMD techniques (Framework 2 (FW2) and Framework 3 (FW3) respectively) as part of the preprocessing stage of the EEG recordings. The last framework (FW4 - proposed approach) uses together EMD and CSP techniques as a preprocess of the data.

For a given subject, a set of signals $\{\boldsymbol{Y}_r : r = 1, \ldots, 200\}$ was obtained, with $\boldsymbol{Y}_r \in \mathbb{R}^{C \times T_Y}$, $C = 59$ and $T_Y = 400$. For FW3 and FW4 the number of IMFs in EMD is fixed as $N_c = 3$ [1]. Thereby, for each framework the following analysis is performed. According to the described features in section 2.2, three frequency-based (PSD, CWT, and DWT) and one time-based (Hjorth parameters) kind of features are estimated for each channel of a given trial $\hat{\boldsymbol{y}}_r$. Hence, a feature space representation matrix $\mathbf{X} \in \mathbb{R}^{400 \times 1593}$ is calculated. It is important to note that for the segment length value L in PSD and Hjorth parameters based features, the minimum frequency to be analyzed is fixed as $F_r = 8$ Hz, taking into account that the band of interest for the analyzed BCI application is $8 - 30$Hz (containing the α and β rhythms).

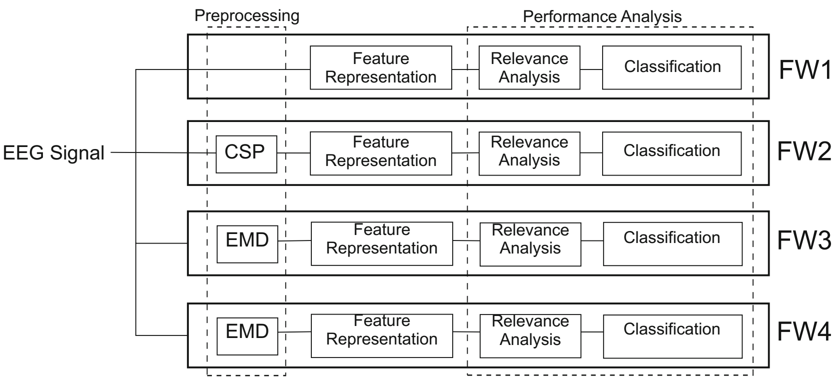


Fig. 1. Tested Frameworks. FW4-proposal.

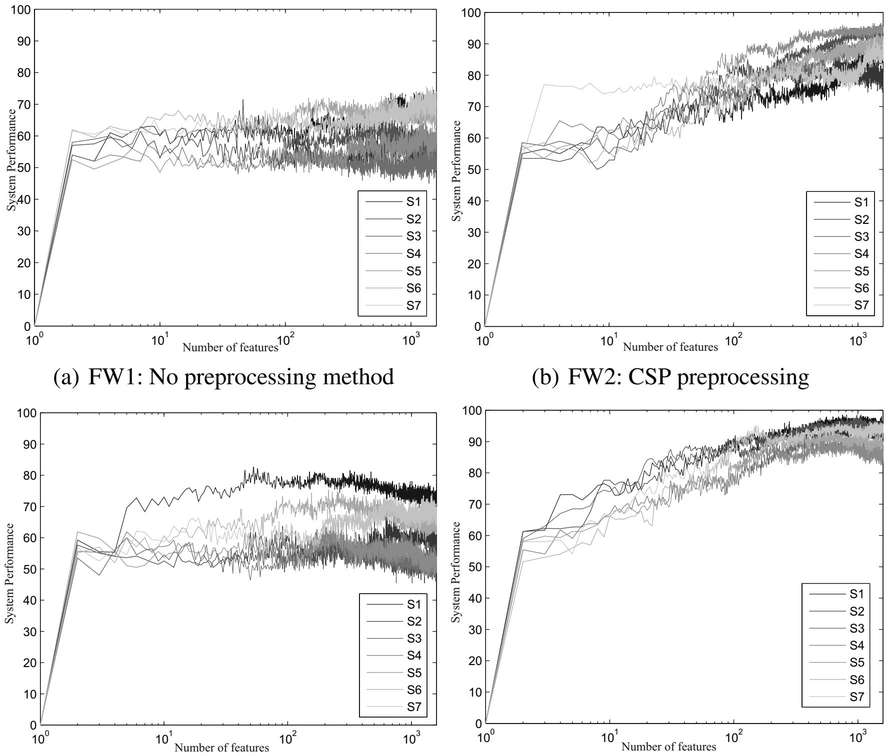
Regarding to the eigendecomposition-based feature relevance analysis presented in section 2.3, the number of dimensions D' in PCA is calculated looking for a 95% of the input data variability. Therefore, the inferred relevance vector $\rho \in \mathbb{R}^{1593\times 1}$ is employed to sort the original features. In addition, a soft-margin Support Vector Machine (SVM) classifier is trained using a regularization parameter $C \in \mathbb{R}^{+}$, and a Gaussian kernel $k(\mathbf{x}_a, \mathbf{x}_b) = \exp(-||\mathbf{x}_a - \mathbf{x}_b||/2\delta^2)$, with band-width $\delta \in \mathbb{R}^{+}$; and being $\mathbf{x}_a, \mathbf{x}_b \in \mathbb{R}^{1\times D}$ two given samples of the feature representation space. We generate a curve of performance adding one by one the characteristics obtained in each subspace representation based on the order given by the relevance vector $\boldsymbol{\rho}$. For a given subset, the optimum working point has been searched using a 10-fold cross validation scheme to fix the C and δ values. The C value is selected from the set $\{1, 10, 100, 1000\}$; and the δ value from $\{\delta_s, 10\delta_s, 100\delta_s, 1000\delta_s\}$; being $\delta_s = 0.9\min(\mathbf{E}\{\sigma(\Xi)\}, (1/1.34)\mathbf{E}\{\mathrm{iqr}(\Xi)\})$ the Sylverman rule based Gaussian kernel band-width estimation. Note that $\sigma(\cdot)$ computes the standard deviation and $\mathrm{iqr}(\cdot)$ the interquartile range of a provided set of features, respectively. Table 1 shows the best BCI system performance for each subject according to each training framework. Fig. 2 presents the system performance for the four tested frameworks, thus is, these figures show the accuracy as a function of the number of chosen features according to proposed relevance analysis. Finally Fig. 4 presents the distribution relevance information per channel extracted to each method.

Table 1. Classification results (average accuracy ± standard deviation, 10-fold cross validation)

Subject	Framework 1	Framework 2	Framework 3	Framework 4
	Acc. (%)	Acc. (%)	Acc. (%)	Acc. (%)
S1	74.50±09.26	89.50±07.62	82.66±11.36	98.50±03.37
S2	67.00±15.49	86.50±04.74	67.74±09.41	95.97±03.95
S3	71.50±12.03	96.50±05.29	60.50±08.32	98.50±03.37
S4	59.00±13.70	93.00±06.75	63.37±08.48	91.84±04.84
S5	65.00±08.16	96.50±04.74	65.55±07.18	91.76±05.98
S6	73.50±12.03	93.50±04.74	75.29±05.38	93.42±06.28
S7	75.50±08.96	92.00±05.37	74.53±09.19	96.50±4.11
Mean	**69.43±11.38**	**92.50±5.61**	**69.95±08.47**	**95.21±04.42**

4 Discussion

According to Table 1, it is possible to notice that carry out a preprocessing stage improves both the performance and the BCI system stability. The best mean discrimination performances are obtained by FW4 and FW2, respectively. The above statement can be explained by the fact that FW4 and FW2 use signal decomposition methods (CSP, CSP-EMD) working as filters, which remove information that can decrease the classification performance. Although some subjects (S4 and S5) present lower classification performance in FW4 than FW2, it is explained by the fact that each subject presents different cognitive characteristics, not mentioning the non-stationary nature of the signals. Additionally, the quality of the EEG trials is perturbed by the artifacts, and by the brain



(a) FW1: No preprocessing method

(b) FW2: CSP preprocessing

(c) FW3: EMD preprocessing

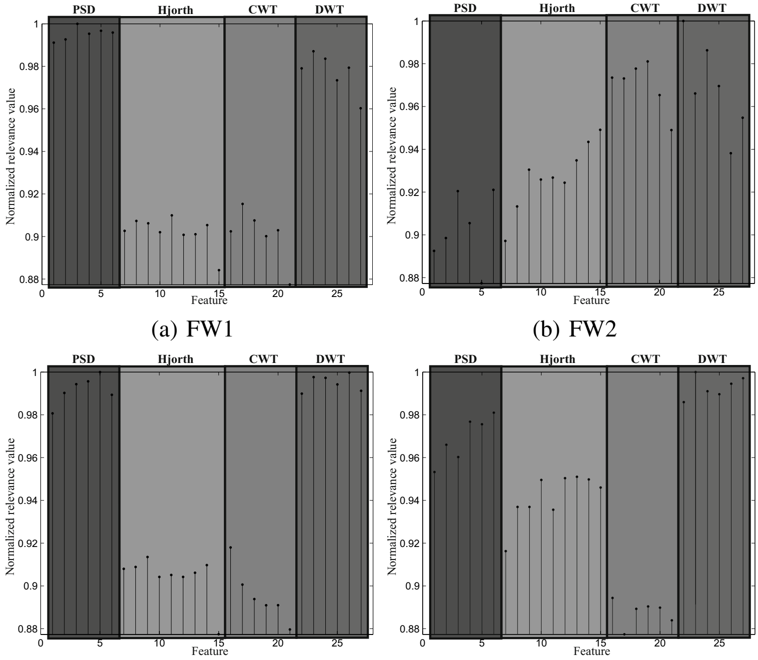
(d) FW4: EMD - CSP preprocessing

Fig. 2. Performance curves

response capability of each subject. Even though FW1 and FW3 present a similar accuracy, the performance computed by FW3 is more stable than FW1, because FW3 is calculated only using the first three IMFs which are mainly related to the optimal informative frequency bands of interest (α and β rhythms) for MI classification [1].

Moreover, from the attained performance of each framework (Fig. 2), note that for FW1 (Fig.2(a)) and FW3 (Fig. 2(c)), overall, the first 10 relevant features achieved the maximum system performance without a notable gain by increasing the number of features. For FW2 (Fig. 2(b)) and FW4 (Fig. 2(d)) the performance notedly increases by adding features. Also, in some cases, the BCI performance curves present local minimums when adding new relevant features, and then the classification accuracy grows up again. This behavior is explained by the fact that some features may represent highly relevant attributes, but they involve redundant information, i.e. needling phenomenon.

Figure 3 shows the relevance of features to each one of the frameworks. In FW1 (Fig. 3(a)), FW2 (Fig. 3(b)), and FW4 (Fig. 3(d), both PSD and DWT methods provides a better relevance value than the other analyzed features. This is because the PSD features are estimated into a restricted frequency band-width (α and β bands), in order to take advantage of the prior knowledge about the studied phenomenon [4], besides, the signal



(a) FW1 (b) FW2

(c) F3 (d) F4

Fig. 3. Feature relevance values

windowing procedure allows to deal with the non-stationary nature of the EEG. Regarding to DWT, they also bring relevant information, since, this method allows to extract features of interest from μ and β rhythms. As expected, the DWT transformation provides a multi-resolution decomposition, which is able to deal with non-stationary signals. In this sense, the Hjorth features, generally, can not captured the non-stationarity behavior of the signals, because the just analyze second-order statistical moments.

From figures 4(a), 4(b), and 4(c) (FW1, FW2, and FW3, respectively) it is possible to see how most of the channels are considered as high relevance channels. In these frameworks the frontal cortex (i.e. the AF,F,FC electrodes) seems to present discriminant information. However, the Primary Motor Cortex – PMC (FC electrodes) is related to movement mode but not to imagery mode. On the other hand, the FW4 (Fig. 4(d)) exhibits low relevance on the frontal cortex (AF, F, and FC electrodes). A human lesion study suggests that PMC does not play a fundamental role in motor imagery process, although individual subjects may show PMC activity during motor imagery, depending on their thinking strategy. Besides the activity associated to the task performance for the imagery mode was localized in the precentral sulcus, indicating the significance of this region in motor imagery [9].

FW1, FW2, and FW3 show high relevance for the parietal cortex, however, activity in the anterior parts of the parietal cortex (i.e. CP electrodes), most likely reflects somatosensory-motor association and sensory feedback from movement mode, but not

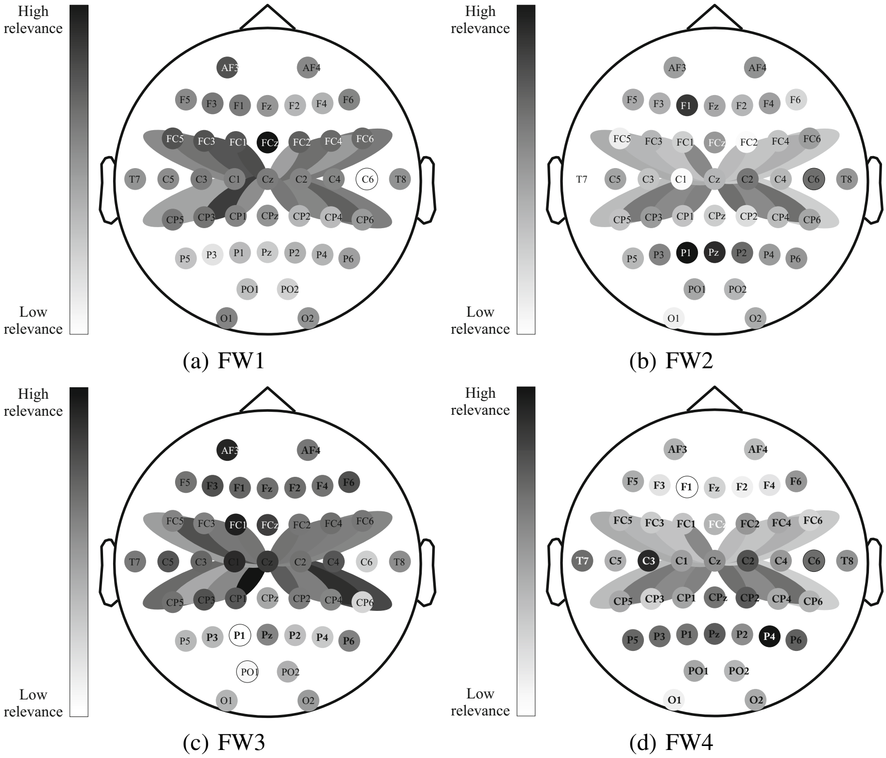


Fig. 4. Channels relevance

the imagery mode, which implies a decrease of the performance of the classification stage in these frameworks. Moreover, detailed neuropsychological examination supports the role of the parietal cortex in generating mental movement representations. The posterior part of the parietal cortex, including the precuneus (P and PO electrodes), has been reported to be active during tasks involving motor imagery [9], which corresponds with the relevance configuration found by FW4 (Fig. 4(d)). The imagery-predominant activity showing the overactivity of the posterior parietal cortex during motor imagery of finger movement. Thus, the channels relevance analysis, computed by FW4, resemble to the clinic findings of the state of the art.

5 Conclusions and Future Work

In this paper we develop a BCI based motor imagery classification framework using CSP and feature relevance analysis. The CSP is matched with EMD to reveal the dynamics of interest. We select a set of features representing the best as possible the studied process. For such purpose, a variability analysis is presented to identify relevant features. Four frameworks of training for MI classification were compared. Experimental results showed that the precision of the MI system significantly increases when a preprocessing stage is done, gaining accuracy and reducing the variance among experiments, achieving a major system reliability and stability. In order to model the studied phenomenon, three frequency-based (PSD, DWT, and CWT) and one time-based (Hjorth parameters) features were used. Moreover, a soft-margin SVM based classifier was employed, and the BCI-system was validated by a 10-fold cross validation methodology. Achieved results showed that in general the PSD based features provided a better relevance value than the other analyzed features. Furthermore, DWT attributes also brought relevant information to the BCI-system. Also, the relevance per EEG channel was computed, which found that the proposed framework (FW4) presents a great similarity with the clinical findings about the brain function pointed in the state of the art, that is, the frontal cortex shows low relevance, and high relevance in the central cortex, and an average relevance in the parietal hemisphere, highlighting the motor imagery functions of the brain. As future work, it would be interesting to identify other movements and to analyze other decomposition methodologies. Besides, it would be interesting to test our methodology in other kind of EEG applications as Epilepsy detection and monitoring.

References

1. Wei-He, Wei, P., Wang, L., Zou, Y.: A novel emd-based common spatial pattern for motor imagery brain-computer interface. In: IEEE EMBC (2012)
2. Ang, K.K., Chin, Z.Y., Zhang, H., Guan, C.: Filter bank common spatial pattern (fbcsp) in brain-computer interface. In: IEEE International Joint Conference on Neural Networks, IJCNN 2008 (IEEE World Congress on Computational Intelligence) (2008)
3. Rodríguez, G., García, P.J.: Automatic and adaptive classification of electroencephalographic signals for brain computer interfaces. Medical Systems 36(1), 51–63 (2012)
4. Corralejo, R., Hornero, R., Álvarez, D.: Feature selection using a genetic algorithm in a motor imagerybased brain computer interface. In: IEEE EMBC (2011)
5. Bankertz, B., Tomioka, R., Lemm, S., Kawanabe, M., Müller, K.-R.: Optimizing spatial filters for robust eeg single-trial analysis. IEEE Signal Processing Magazine 08, 41–56 (2008)
6. Teixeira, A.R., Tomé, A.M., Boehm, M., Puntonet, C., Lang, E.: How to apply nonlinear subspace techniques to univariate biomedical time series. IEEE Trans. on Instrument. and Measur. 58(8), 2433–2443 (2009)
7. Li, M.-A., Wang, R., Hao, D.-M., Yang, J.-F.: Feature extraction and classification of mental eeg for motor imagery. In: 5th Int. Conf. on Nat. Comp., ICNC 2009, vol. 2 (2009)
8. Daza-Santacoloma, G., Arias-Londoño, J.D., Godino-Llorente, J.I., Sáenz-Lechón, N., Osma-Ruíz, V., Castellanos-Domínguez, G.: Dynamic feature extraction: An application to voice pathology detection. Intel. Aut. and Soft Comp. (2009)
9. Hanakama, T., Immisch, I., Toma, K., Dimyan, M.A., Van Gelderen, P., Hallett, M.: Functional properties of brain areas associated with motor execution and imagery. J. Neurophysiol. 89, 989–1002 (2003)