Biomedical Signal Processing and Control 57 (2020) 101812

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Improved motor imagery brain-computer interface performance via adaptive modulation filtering and two-stage classification

Eliana M. dos Santos [a], Raymundo Cassani [b], Tiago H. Falk [b], Francisco J. Fraga [a,*]

[a] UFABC, Federal University of ABC, Santo André, Brazil
[b] INRS-EMT, University of Quebec, Montreal, QC, Canada

ARTICLE INFO

Article history:
Received 19 June 2019
Received in revised form 24 October 2019
Accepted 23 November 2019
Available online 7 December 2019

Keywords:
Brain-computer interface
Motor imagery
Adaptive modulation filtering
Multi-stage classification

ABSTRACT

Electroencephalogram (EEG) based brain-computer interfaces (BCI) monitor neural activity and translate these signals into actions and/or decisions, with the final goal of enabling users to interact with a computer using only their thoughts. To this end, users must produce specific neural activity patterns that are used by the system as control signals. A common task used to elicit such signals is motor imagery (MI), where specific patterns are elicited in the sensorimotor cortex during imagination of movements (e.g., of the hands, arms, feet or tongue). The processing pipeline typically used in EEG-BCIs consists of three stages: pre-processing, feature extraction, and classification. Here, we propose innovations in pre-processing and classification and quantify the gains achieved on 4-class MI-based BCI performance. More specifically, for the pre-processing stage, we propose the concept of spectro-temporal filtering as we show that MI-elicited neural patterns have varying amplitude modulation variations relative to artifacts. For the classification stage, in turn, a two-step classification method is proposed. First, LDA classifiers are used to discriminate between different pair-wise MI tasks. Next, a naive Bayes classifier is used to predict the final task performed by the user based on the weighted outputs of the LDA classifiers. Experimental results showed that the proposed system outperformed the first-place winner of the BCI competition IV by 3.5 %.

1. Introduction

An EEG-based brain-computer interface (BCI) is a communication system that captures the electrical signals emitted by the cerebral cortex and translates these signals into actions and/or decisions, with the final goal of enabling the user to interact with a given equipment without using his/her muscles [1]. To operate a BCI, the user must produce different patterns of brain activity that will be identified and used by the system as control signals [2]. Motor imagery is one of the most widely used BCI paradigms and comprises of the use of imagined movements to elicit neural activity in the sensorimotor cortex, just as observed when one performs the movement [3]. The typical signal processing pipeline in current BCI systems consists of three stages: pre-processing, feature extraction, and classification [4].

Pre-processing aims at simplifying subsequent processing operations without losing relevant information. An important function of pre-processing is to improve signal quality by increasing the signal-to-noise ratio (SNR). A low SNR means that the brain patterns are buried in the signal (e.g. background EEG), which makes relevant patterns hard to detect [1]. A high SNR, on the other hand, tends to improve BCI classification accuracy. It is known that EEG signals are affected by various types of artifacts, such as power line interference, involuntary movements, muscular contractions, amongst others. For artifact removal, several manual, semi-automated or fully-automated methods have been proposed, such as time-frequency filtering, independent component analysis (ICA) and others [3]. Considering that EEG signals present non-stationary behavior, i.e., their spectral content changes over time, we explore the use of a recent technique called modulation filtering (also called spectro-temporal filtering) as a new and fully-automated EEG artifact removal method [5]. Modulation filtering is based on amplitude modulation analysis and requires the application of two transformation mappings: in the first, the EEG signal is transformed into a time-frequency representation (e.g., via short-time Fourier transforms or wavelet transforms). A second transform is then taken across the time dimension to result in

* Corresponding author at: UFABC – Federal University of ABC, CECS – Engineering, Modelling and Applied Social Sciences Center, Avenida dos Estados, 5001, Bairro Santa Terezinha, Bloco B, 9.o andar, sala 940, CEP: 09210-580, Santo André, SP, Brazil.
E-mail address: francisco.fraga@ufabc.edu.br (F.J. Fraga).

https://doi.org/10.1016/j.bspc.2019.101812
1746-8094/

a frequency-frequency representation termed "modulation spectrogram" [5]. Modulation filtering has been successfully employed in the past to blindly separate heart and lung sounds from breathing sound signals [6], to enhance noisy speech for hearing devices [7] and to enhance noisy electrocardiograms (ECG) [8]. To the best of our knowledge, this is the first time it is being explored within an EEG-based MI BCI paradigm [9].

In order to extract features for classification of motor-imagery tasks, we used the common spatial pattern (CSP), one of the most used and more effective methods for MI BCI. This technique provides spatial filters that allow to separate two conditions by maximizing differences in variance between them. Since variance of band-pass filtered signals is equivalent to band-power, CSP filters are well suited to discriminate mental states that are characterized by motor sensory rhythm effects [10]. Although the CSP algorithm is very efficient, it is noise sensitive for limited datasets [11]. To address such drawback, several CSP variants have been proposed to make it more robust. The interested reader is referred to [10] for more details. Here, we employed CSP combined with Tikhonov regularization [11].

For classification, we proposed a novel two-stage design: the well-known and widely used linear discriminant analysis (LDA) algorithm [12] was used first as a binary classifier to generate weighed outputs of the six possible pairwise combinations among four motor-imagery tasks (right hand, left hand, foot and tongue) performed by the subjects. LDA classifier outputs were then directed as inputs to a naive Bayes classifier, which made the final four-class decision.

The remainder of this article is organized as follows. Section 2 describes the methods and materials, including the modulation representation, standard database used, feature extraction, classifiers, and performance figure of merit. Sections 3 and 4, in turn, present the experimental results and discusses them, respectively. Lastly, conclusions are drawn in Section 5.

2. Methods and materials

2.1. Amplitude modulation analysis

A non-stationary signal, such as the EEG, can be modeled as the result of the interaction between two independent signals: a low-frequency signal (modulator) that changes or modulates the amplitude of a higher frequency (carrier) signal [5,7]. This corresponds to the well-known concept of amplitude modulation. In amplitude modulation analysis, a signal $x(t)$ can be expressed as the product of a low-frequency modulating signal $m(t)$, and a high-frequency carrier signal, $c(t)$:

$$x(t) = m(t)\, c(t), \tag{1}$$

assuming: i) there is no spectral overlap between $m(t)$ and $c(t)$ and ii) $m(t)$ is a real and non-negative signal. The modulator $m(t)$ can be easily extracted by applying an envelope detector to $x(t)$ [7]. As a simple graphical example, Fig. 1 depicts the modulating signal $m(t)$ (1-A), the carrier signal $c(t)$ (1-C) and the resulting signal $x(t)$ (1-E) in the time domain. Fig. 1D, B and F show, respectively, their frequency-domain representations.

2.2. Modulation spectrogram

The modulation spectrogram is a complex-valued function of a signal that portrays the over-time amplitude changes of its spectral components in the frequency domain [5]. The modulation spectrogram allows us to explore amplitude modulation changes, revealing undetectable hidden periodicities in the temporal representation of a biomedical signal. This technique has been successfully applied to ECG signals [8], breathing sound signals [6] and speech [7,13]. In the experiments described herein, we used the open-source modulation spectrogram toolbox for MATLAB described in [5]. In Fig. 2, we show some useful graphical representations performed by this toolbox. Fig. 2(a), for example, shows a time-domain single-channel EEG signal. A shorter segment of the same signal (highlighted by the green box) can then be zoomed in, as seen in Fig. 2(c). The traditional (time-frequency) spectrogram of the entire signal is depicted in Fig. 2(b), whereas for the shorter segment in subplot (d). Lastly, the modulation (frequency-frequency) spectrogram of the signal segment is shown in Fig. 2(e).

The amplitude modulation analysis pipeline is illustrated in Fig. 3. The input signal is windowed into short-time segments and a first transformation (e.g. Continuous Wavelet Transform – CWT) is applied to generate the time-frequency representation of the signal. A second transformation, called modulation transform (e.g. Fourier transform – FT), is then applied to the spectral representation of the short-time signal segments, i.e., FT is applied at each frequency bin of the magnitude spectrogram, discarding the phase information.

2.3. Modulation filtering

The assumption here is that the spectral content of the EEG signals can be enhanced by removing noise in the modulation domain. If both transformations used are invertible, then filtering can be performed in the modulation domain and the enhanced signal can be reconstructed back into the time domain via two invertible transformations. Fig. 4 illustrates the modulation filtering process used in this study. We applied a Continuous Wavelet Transform (CWT) to short-time segments of EEG signal channels. Then, frequency-domain magnitude signals underwent a second transformation, performed through Fast Fourier Transform (FFT), converting the signal to the modulation frequency domain. After that, we enhanced the signal by filtering out some regions in the modulation spectrogram that we have found to be dominated by artifacts (more details in Section 4.1).

For the CWT implementation, we used the complex Morlet wavelet, also known as Gabor wavelet. It presents good resolution both in frequency and in time, and has been widely used in biological signal analysis [14,15]. The complex Morlet wavelet consists of a complex oscillation with a fixed frequency tapered by a Gaussian window [5]. The mother wavelet for the complex Morlet wavelet family is defined by

$$\psi_f(t) = A(\theta_t) e^{-t^2/2\theta_t^2} e^{-j2\pi f_0 t}, \tag{2}$$

where $A\left(\theta_t\right)$ is used to ensure that the wave energy is equal to one, $e^{-t^2/2\theta_t^2}$ represents the Gaussian window with a time standard deviation of θ_t and $e^{-j2\pi f_0 t}$ corresponds to the complex oscillatory component with frequency of $f_0\,Hz$. The number of cycles in the frequency at the Gaussian bell is defined by N_c, i.e., $\theta_t(f_0) = \frac{N_c}{2\pi f_0}$. Smaller values assigned to the number of cycles increase the temporal resolution, whereas higher values assigned to the number of cycles increase the spectral resolution. For our experiment, we used six cycles to obtain an adequate resolution both in the time and frequency domains [5]. In summary, the spectrotemporal representation, X(t; f), was computed using the discrete-time CWT with the complex Morlet wavelet as mother wavelet, for frequencies from 0.5–125 Hz with 0.5-Hz step, and number of cycles equal to 6.

2.4. Database

The BCI Competition IV (BCIC IV) dataset 2a was used in this study [16]. This is a well-known publicly-available database and was already used in several MI-BCI studies [17–22]. Data is separated in training and testing EEG with data recorded from nine

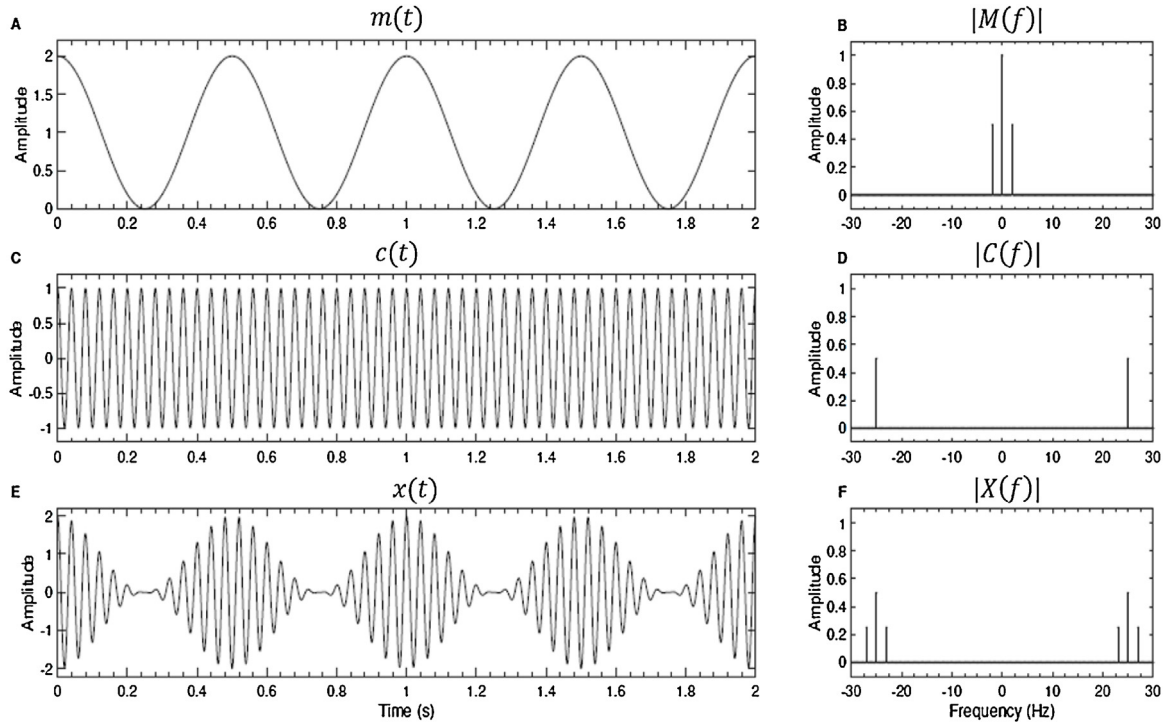

Fig. 1. Time- and frequency-domain representation of a simple amplitude modulation process [5].

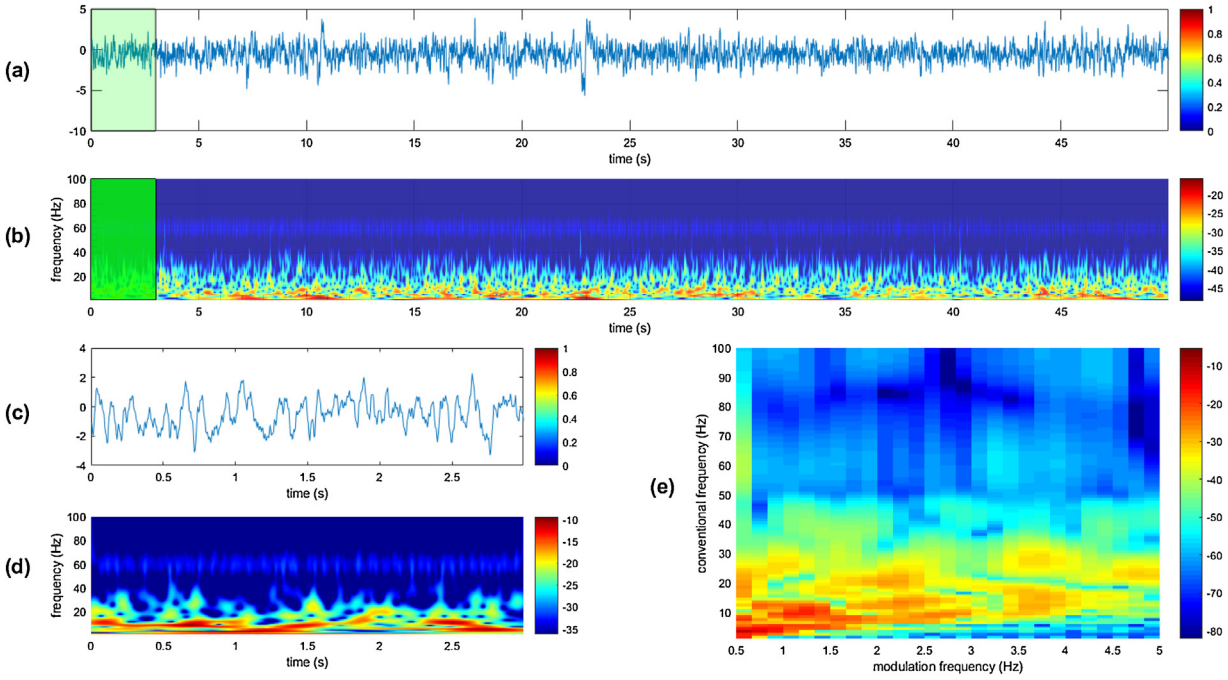

Fig. 2. Modulation spectrogram toolbox applied to a single-channel EEG signal.

subjects. Participants were positioned comfortably in front of a computer and performed four motor-imagery tasks: left hand, right hand, feet and tongue [16]. At the beginning of each trial, a fixation cross appeared on screen. After two seconds, a cue in the form of an arrow pointing left or right (hand task), down (feet task) or up (tongue task) appeared and remained on screen for about 1.25 s, signaling to the subjects which motor-imagery task of they should perform. Participants were instructed to perform the task until the fixation cross disappeared from the screen, which occurred four seconds after the appearing of the cue (arrow). Then the screen turned black, marking a brief pause. After that, the fixation cross appeared again, indicating the beginning of a new trial. The protocol and timing are shown in Fig. 5. EEG recording was divided into short runs, each run containing 48 trials from each of the four motor-imagery tasks. Training and test sessions occurred in different days and were comprised of six runs each. Thus, each session contains a total of 288 trials from each of the four motor-imagery tasks.

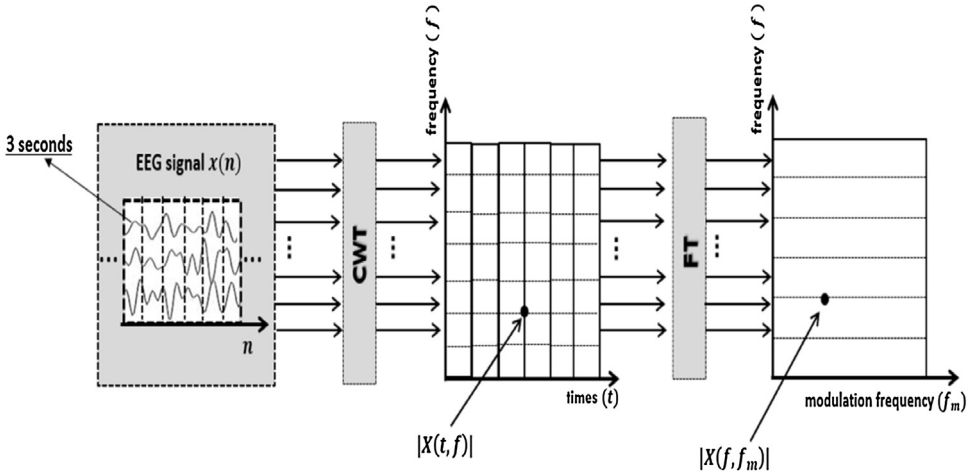

Fig. 3. Amplitude Modulation analysis pipeline.

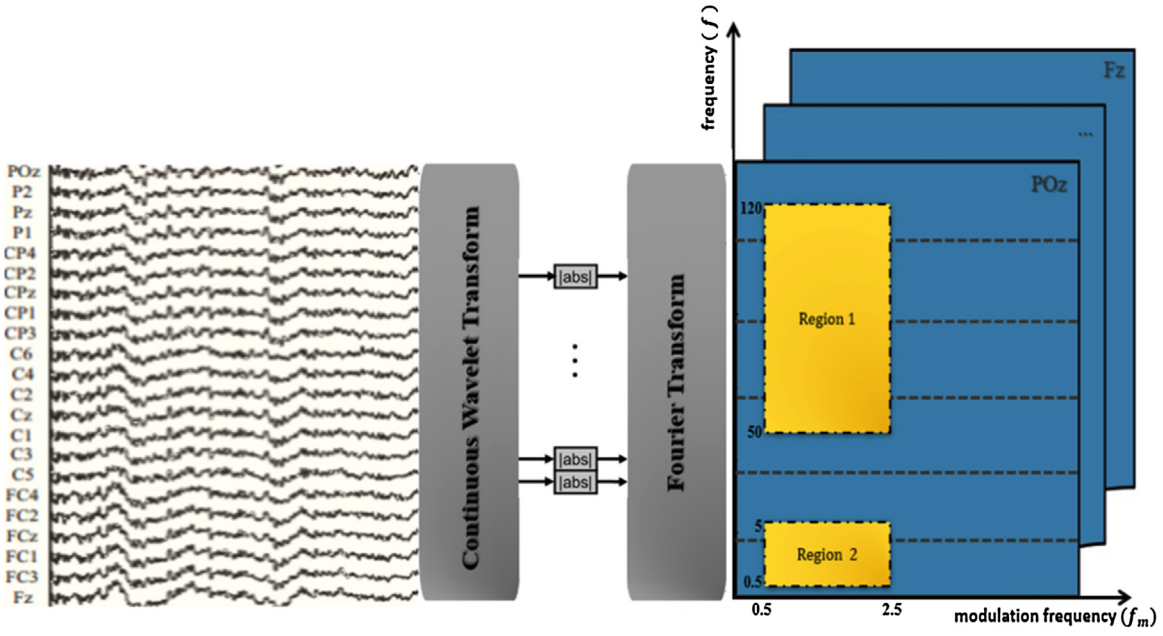

Fig. 4. Modulation filtering process.

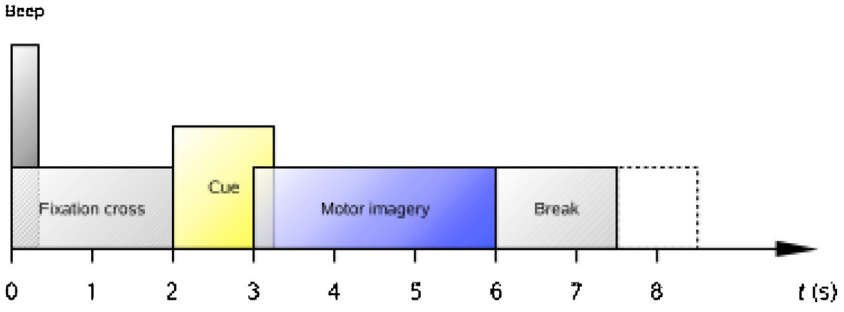

Fig. 5. Timing of the applied paradigm [16].

For EEG signal recording, 22 electrodes were used according to the international 10–20 system (Fz, FC3, FC1, FCz, FC2, FC4, C5, C3, C1, Cz, C2, C4, C6, CP3, CP1, CPz, CP2, CP4, P1, Pz, P2, POz), with the left and right mastoids as reference and ground, respectively. In addition, three monopolar electrooculography (EOG) channels were recorded. EEG and EOG signals were sampled at 250 Hz and bandpass filtered (0.5–100 Hz). Also, a notch filter centered in 50 Hz was used to remove power-grid interference. The interested reader can find more details about the subjects profile, experiment protocol and signal acquisition procedures in [16].

2.5. Signal enhancement

We defined two regions in the modulation spectrograms of EEG signals, defined by rectangles in the frequency-frequency analysis graph. The first region was delimited by 50–120 Hz (conventional frequency, y-axis) and 0.5–2.5 Hz (modulation frequency, x-axis). Region 2, in turn, was defined through the rectangle 0.5–5 Hz, 0.5–2.5 Hz. These regions were selected empirically, after a thorough visual inspection of motor imagery trials extracted from the training dataset, looking for the best filtering regions based on our previous experience on modulation filtering of EEG signals. Such regions were filtered out from the modulation spectrograms, as a pre-processing step aiming to improve the SNR. Fig. 6 shows an example of time-domain signals from subject A09 (training session, feet MI, trial #19) before (blue) and after (red) modulation filtering.

Following modulation filtering, the three-second segment of motor-imagery EEG (see Fig. 5) underwent an additional (three-band) bandpass filtering using 5th-order Butterworth filters, where we used overlapping frequency bands specifically designed for this study: theta + alpha (4–14 Hz), alpha + beta (8–30 Hz) and beta + gamma (15–40 Hz). As occurred with the design of modulation spectrogram filters, we made several tests with the training database alone in order to find out which was the best combination of overlapping bands.

2.6. Feature extraction

For feature extraction, we used CSP with Tikhonov regularization (TRCSP). CSP is a spatial filter widely used to classify non-stationary EEG signal activities [23]. It is a method used to learn spatial filters that maximize differences in EEG signals belonging to different classes [10] Variance of bandpass EEG signals corresponds to signal power, so CSP attempts to achieve an optimal discrimination for BCIs based on spatial power distribution [10,24]. Formally, CSP uses the spatial filter w which maximize the following function:

$$J(w) = \frac{w^T X_i^T X_i\, w}{w^T X_j^T X_j\, w} = \frac{w^T C_i\, w}{w^T C_j w}, \tag{3}$$

where T represents transpose, $X_i$and X_j are the data matrices for class i and class j respectively (number of channels × number of samples), and $C_{i,j}$ are the spatial covariance matrices of classes i and j, assuming zero mean for EEG signals.

Despite being widely used in BCIs, CSP is sensitive to noise and tends to overfit for small training sets. In order to overcome such problems, several types of regularization were proposed [10]. These methods consist in adding a regularization term to the objective function of the CSP to penalize solutions (spatial filter outputs) that do not satisfy a given prior. Thus, adding the regularization parameter to Eq. (3) leads to

$$J_{P_1}(w) = \frac{w^T C_i\, w}{w^T C_j\, w + \alpha P(w)} \tag{4}$$

where $P(w)$ is a penalty function measuring how much the spatial filter w satisfies a given prior, which means that the closer the filter w satisfies the prior, the lower the penalty $P(w)$ can be. In order to ensure maximization of spatial filters satisfying the prior, we have to maximize $J_{P_1}(w)$ and minimize $P(w)$ ($\alpha \geq 0$, the higher the α, more satisfied the prior is). In this study, we implemented the Tikhonov solution for CSP regularization, since it was the one which offered the best classification results [10]. The Tikhonov regularization algorithm uses a quadratic penalty function. This is a classic form of regularization, initially introduced for regression problems, and consists in penalizing solutions with heavy weights. The penalty function is then given by $P(w) = \|w\|^2 = w^T w$. This type of regularization gives preference to solutions with smaller norms, thus mitigating the influence of artifacts and outliers.

Table 1
Pairwise combinations of MI tasks.

Task 1		Task 2
Left Hand	X	Right Hand
Right Hand	X	Foot
Foot	X	Left Hand
Foot	X	Tongue
Tongue	X	Left Hand
Tongue	X	Right Hand

2.7. Two-stage classification

We used LDA, a binary classifier widely used by motor-imagery BCI researchers [4,12], to train six two-class classifiers for all possible pairwise combinations of MI tasks, as show in Table 1. LDA assumes normal distribution with equal covariance matrix for both classes. The separating hyperplane is obtained by seeking the projection that maximizes the distance between the means of the two classes. For binary classification tasks, a discriminant function has a decision hyperplane defined by

$$\hat{w}.f(v) + b, \tag{5}$$

where v is the feature vector, $\boldsymbol{f}(\boldsymbol{v}) = \boldsymbol{v}$ in the linear case, $\hat{\boldsymbol{w}}$ is a vector of classification weights, and b is the bias term [12]. The weights vector $\hat{\boldsymbol{w}}$ is calculated through

$$\hat{w} = (V^T V)^{-1} V^T y, \tag{6}$$

where V is the matrix of observed feature vectors and y is the class labels vector. This technique has a very low computational requirement, which makes it suitable for online BCI systems. Moreover, this classifier is easy to use and generally provides good results [4,12]. The main drawback of LDA is its linearity, which provides poor results on complex nonlinear EEG features [12].

In most papers related to BCI, classification is achieved using a single classifier. However, recently there is a trend to use multiple classifiers aggregated in different ways [25]. Thus, our proposed strategy was to perform two-stage classification combining LDA with Naive Bayes classification. More specifically, we propose to use the six two-class LDA outputs (six pairwise combinations, according to Table 1) for each of the three features, thus resulting in an 18-dimensional feature vector, that are then into a Naive Bayes classifier for final four-class classification.

Bayesian classification aims at assigning to a feature vector the class it belongs to with the highest probability. The Bayes' rule is used to compute the so-called *a posteriori* probability, i.e., the probability that a feature belongs to a given class [25]. Even though this classifier is not as widely used for BCI as the LDA, it has been previously applied with success to motor imagery and mental task classification [26]. The well-known Bayes' rule is

$$\rho(A\boldsymbol{v}) = \frac{\rho(\boldsymbol{v}A)\, P(A)}{\rho(\boldsymbol{v})}, \tag{7}$$

where $\rho(A\boldsymbol{v})$ is the conditional probability of class A given feature vector $\boldsymbol{v}$; $\rho(\boldsymbol{v}A)$ is the conditional probability of $\boldsymbol{v}$ given A, and $P(A)$ is the *a priori* probability of class A.

The calculation of $\rho(v|A)$ becomes feasible by a naive assumption that features $v_1, v_2, \ldots v_d$ forming the feature vector $\boldsymbol{v}$ are independent, thus following

$$\rho(\boldsymbol{v}A) = \prod_{j=1}^{d} \rho\left(v_j A\right). \tag{8}$$

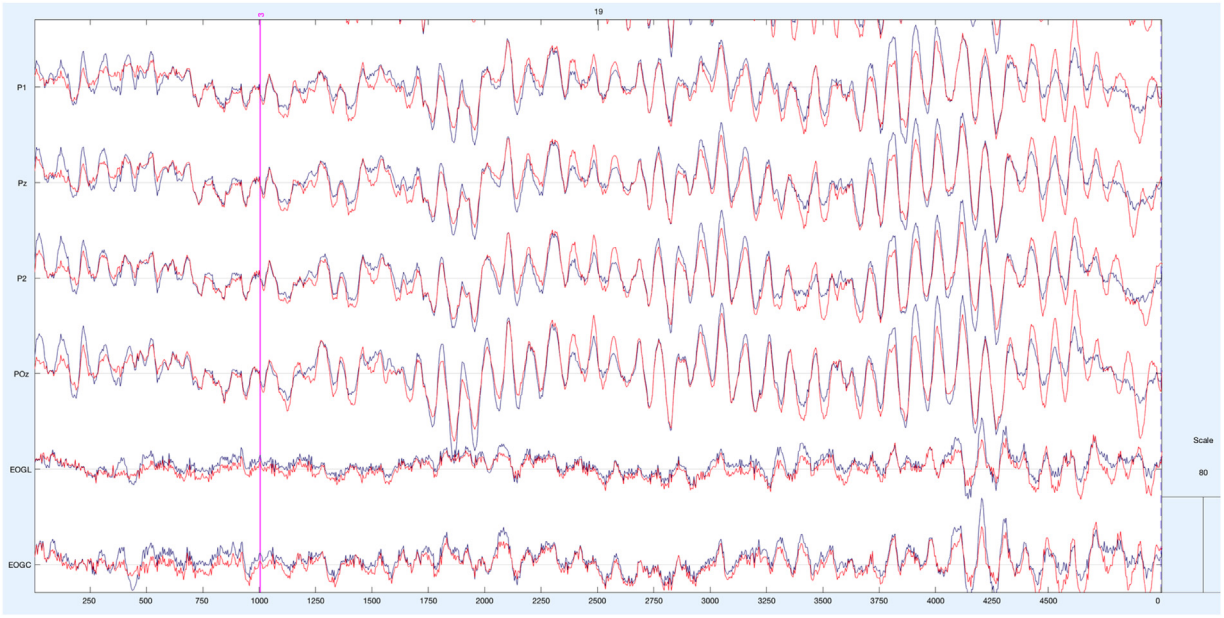

Fig. 6. EEG signals from subject A09 (training session, feet MI) before (blue) and after (red) modulation filtering.

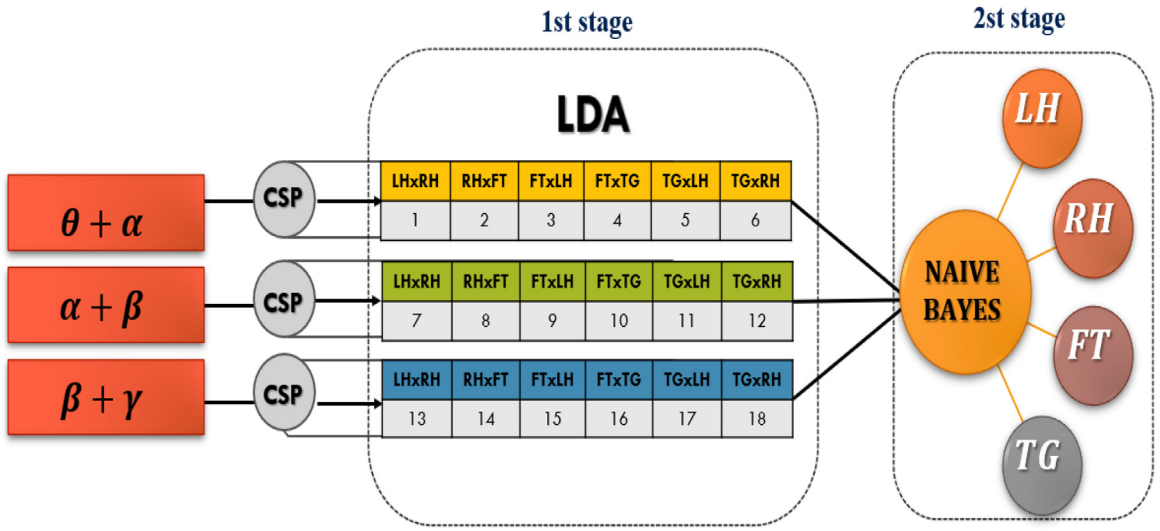

Fig. 7. Two-stage classification scheme after feature extraction.

MATLAB was used to implement the Naive Bayes classifier. We used the function *fitcnb.m*, which returns a prediction model for the features (outputs of binary LDA classifiers) extracted from the training dataset. Then, the model is informed as a parameter for the MATLAB *predict.m* function, which returns a vector of predicted class labels for the unseen features extracted from the test dataset. In Fig. 7 we show the two-stage classification scheme used in this study.

2.8. Figure of merit

Cohen's kappa, also known as kappa score, is considered a robust statistical measure for qualitative categorical items [26] and is used here as the BCI figure of merit. The maximum value of kappa is 1, which corresponds to perfectly correct classification, a value close to 0 indicates random performance and negative values mean performance below chance level. Cohen's kappa is computed as:

$$k = \frac{Pr\,(a)\ -\ Pr(e)}{1\ -\ Pr(e)}, \tag{9}$$

where $Pr(a)$ is the classification accuracy and $Pr(e)$ is the chance level. Basically, it informs how much better the classifier is in relation to the performance of a classifier that simply advents randomly according to the frequency of each class [27].

3. Results

Each of the nine subjects has his/her own two-stage classifier trained with 288 trials from each of the four motor-imagery tasks. Each subject's pair of classifiers (LDA followed by Naive Bayes) was then tested with other 288 trials (per MI task) from the same subject, recorded in a different day, as already pointed out in the

Table 2
Results obtained with the proposal of this study, compared to the BCIC IV results.

Kappa										
	A01	*A02*	*A03*	*A04*	*A05*	*A06*	*A07*	*A08*	*A09*	*Mean*
BCI Competition Results										
1 st.	0.68	0.42	0.75	0.48	0.40	0.27	0.77	0.75	0.61	0.57
2nd.	0.69	0.34	0.71	0.44	0.16	0.21	0.66	0.73	0.69	0.52
3rd	0.38	0.18	0.48	0.33	0.07	0.14	0.29	0.49	0.44	0.31
Our Results										
Test 1	0.69	0.40	0.70	0.56	0.24	0.32	0.70	0.76	0.68	0.56
Test 2	0.69	**0.42**	**0.71**	0.56	**0.34**	0.32	0.70	0.76	**0.71**	**0.58**
Test 3	**0.71**	0.33	**0.72**	0.54	0.26	0.30	0.69	0.74	**0.77**	0.56
Best Results	**0.71**	**0.42**	**0.72**	0.56	**0.34**	0.32	0.70	0.76	**0.77**	**0.59**

previous section. In Table 2 we present the results of our three experiments, compared to the results obtained in BCIC IV. All experiments focus on EEG signal enhancement in the pre-processing phase. In test 1 we present the results obtained without modulation filtering. In the second test, we added the modulation spectrogram filtering for enhancement, delimiting the (first) region with boundaries 50–120 Hz (conventional frequency) and 0.5–2.5 Hz (modulation frequency). For Test 3, we changed the modulation filtering settings by adding a second region (0.5–5 Hz, 0.5–2.5 Hz). The row labelled 'Best results' was obtained when we considered the best region per subject, based on the results seen from tests 2 and 3.

As one can see from Table 2, in Test 1 (without modulation filtering) we obtained a slightly lower performance (0.56) than the BCIC IV winner (0.57), but higher than the kappa value obtained by the second place (0.52). With Test 2, in turn, the proposed system outperformed the Test 1 setting, thus showing the advantages of modulation filtering. Moreover, the proposed solution achieved a kappa value slightly above the one obtained by the 1 st place winner in BCIC IV, which used the so-called Filter Bank Common Spatial Pattern (FBCSP) algorithm to optimize frequency bands for CSP calculation on a subject-specific basis [17]. With Test 3, filtering out two regions of the modulation spectrogram helped increase performance for three of the nine subjects (A01, A03 and A09). Lastly, by combining results from Tests 2 and 3 and assuming a system that is optimized for each subject, the best results achieved can be seen both in the Table, as well as depicted by Fig. 8. As can be seen, for five of the nine subjects, improved performance was achieved with the proposed system. Finally, in Table 3 we present the 'Best results' confusion matrix (on a trial-by-trial basis) of the subject who got the best performance (A09), the confusion matrix of the one who got the worst (A06) and the confusion matrix obtained by summing up together the trials of all subjects.

4. Discussion

To the best of our knowledge, our reported classification results on the four tasks of BCIC IV dataset 2a were only slightly surpassed by three recent studies. In 2017, Davoudi, Ghidary and Sadatnejad [20] used dimensionality reduction based on distance preservation to local mean (DPLM) and got a global kappa of 0.6 (averaged over the nine subjects). Last year (2018), Gaur et al. achieved the same kappa (0.6) through multivariate empirical mode decomposition based filtering and Riemannian geometry [21]. The most recent and also the best result obtained so far is from a study which used a decision tree framework and Riemannian geometry [22]. They got a total kappa value of 0.607, which is just a little bit higher than ours (0.59), but not enough to be statistically significantly better (5 % confidence interval = 0.0238). There are also other recent works which used the same BCIC IV dataset 2a [18,19], but none of them were able to beat the global 0.57 kappa obtained by the BCI competition winner [17].

In Test 2 (with modulation-filtering enhancement - one region), compared to the situation without SNR improvement (Test 1, see Table 2), we observed increased performance for subjects A02, A03, A05 and A09, which caused an increase of the overall kappa score to 0.58, overcoming the BCIC IV winner (0.57) [17]. In Test 3 (two regions), there was an increase in performance (kappa value) for subjects A01, A04, A06 and A09, respectively, the latter being the best performance among all subjects. It is important to note that the increase in kappa value due to modulation-filtering enhancement occurred for the majority of the subjects (five out of nine), and more significantly for subject A05, where there was an increase of 41.66 % in the individual kappa value (from 0.24 to 0.34). Notwithstanding, this does not mean much, as such poor performances are close to chance level. As for the remaining four participants, compared to the situation without signal enhancement (Test 1), the modulation spectrogram filtering using only the first region (Test 2) did not affect the performance (subjects A04, A06, A07 and A08).

Comparing the results of Test 2 (one region) and Test 3 (two regions), for subject A09 we had an increase of 8.5 % in the individual kappa value (from 0.71 to 0.77). However, due to the lower performance achieved by the other subjects, we had a reduction of the overall kappa to 0.56, that is, the same overall result obtained without enhancement (Test 1). Finally, when we considered individual modulation-filtering schemes, we were able to achieve an overall kappa of 0.59 (Best results, last line of Table 2), surpassing the best BCIC IV score by 0.02 points. In Fig. 8 we show a bar graph comparing our individual performance rates to the ones obtained by the BCIC IV winner. As shown in Fig. 8, we achieved an increase in the performance rate for subjects A8, A1, A4, A6 and A9. When it comes to the other four, the performance rate remained the same for A2 and decreased for A3, A5 and A7.

As mentioned in the first section, as far as we know, this is the first time the modulation filtering technique for signal enhancement was used in the context of MI-BCI. It is important to remark that, unlike many other EEG artifact removal techniques, once the best modulation spectrogram filters (rectangles) were previously designed, the modulation filtering technique is fully automated. Also, it is noteworthy that the modulation filters were empirically designed based on the training dataset only, thus the test dataset comprised really unseen data. However, the so-called 'Best results' were based on individual choices of filters that were only *a posteriori* found as being 'best', by combining the best individual results from Test 2 and Test 3. Thus, if we were on a real BCI competition, only Test 2 results would be considered as valid results, since in that case all subjects underwent the same modulation-filtering scheme. Regarding the overlapping bands used for feature extraction, which is another novelty we introduced in this study, the procedure was

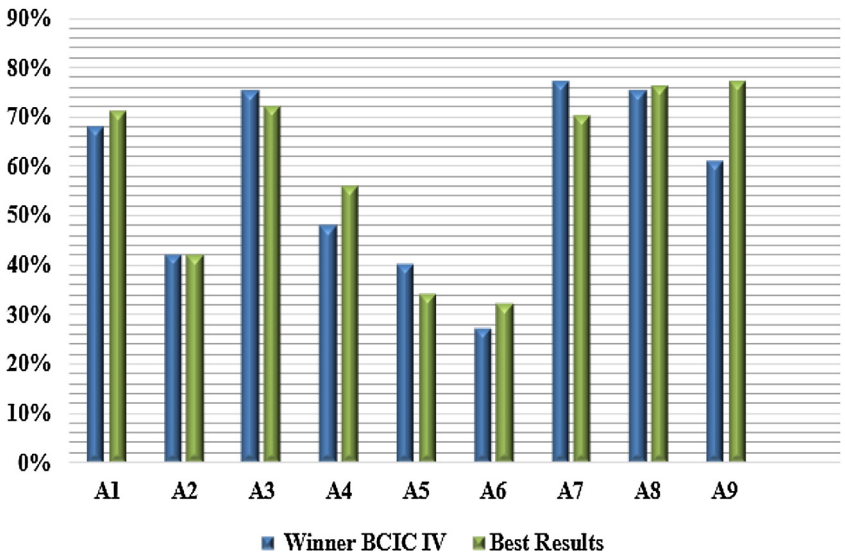

Fig. 8. Our best results compared to the results of the BCIC IV winner (kappa 1.0 = 100 %).

Table 3
Confusion matrices of subjects A06, A09 and all of them. CI: Confidence Interval (binomial distribution).

	Subject A06Kappa = 0.32415 % CI = 0.0772				Subject A09Kappa = 0.76855 % CI = 0.0585 (*)				All subjectsKappa = 0.58805 % CI = 0.0238			
4. Rec MI task5. True MI task	L	R	F	T	L	R	F	T	L	R	F	T
Left hand (L)	45	9	3	15	65	7	0	0	439	93	50	66
Right hand ®	19	32	6	15	5	46	20	1	87	446	67	48
Feet (F)	13	12	22	25	2	5	62	3	46	37	406	159
Tongue (T)	16	11	2	43	0	3	4	65	43	36	69	500

(*): Statistically significant at 5 % level when compared to the 1st. place in the BCI Competition IV.

similar to the one employed for the modulation spectrogram filters: we tested several combinations of overlapping bands on the training database to finally get the one giving the best test-on-training performance. Again, we did no empirical tests on the test database in order to preserve the principle of really unseen data.

Interestingly, in [17] the winners of BCI IV (dataset 2a) benchmarked their FBCSP algorithm against a traditional CSP algorithm, which achieved a kappa of 0.5 (averaged across the nine subjects). Our Test 2 results, which combined the same modulation-filtering setup for all subjects with Tikhonov regularization, overlapping bands and two-stage classification, outperformed this conventional CSP algorithm by 16 %. This is a more suitable way to frame the real progress achieved by our improvements, since the comparison is made with a traditional CSP algorithm applied to the same dataset.

In respect the two-stage classification, the main novelty introduced here lies is the use of a Naïve Bayes classifier in the second stage, receiving as inputs the six outputs from all the possible combinations of two-class classifiers in a four-class scenario. Usually with multiple-class classification, a matrix combination of outputs is used based on majority voting. However, in our experiments with the training database, we found the Naïve Bayes classifier to make improved four-class decision as it works in a way similar to fuzzy logic when compared to classical logic, since the majority voting classifier ignores the "weights" and compares just the outputs in terms of their absolute values.

BCI systems operating in the motor imagery paradigm (IM-BCI) suffer great limitations, mainly regarding the complexity and difficulty of being properly controlled by the end users. Efforts have been made to overcome such limitations, such as training of subjects and classifiers, alternative visual instructions and improvements in the processing and correction of detected errors. Even with such improvements, there will still be people who will not be able to control an IM-BCI system. About 20 % of users are not capable for IM-BCI control at all, and most of the remaining 80 % can only achieve a moderate level of control [28].

Thus, one of the major concerns in MI-BCI is the extremely wide range of overall performances achieved by individual users of the same system. In fact, herein we got subject A09 achieving a kappa value of 0.77, while subject A06 achieved only 0.32 (Table 2). When it comes to the confusion matrices (Table 3), which bring the results separately for the four MI tasks (left and right hand, feet and tongue), we can also see very different performances according to the subject. While subject A09 achieved a good performance in the MI of the feet (86.11 %, 62/72), subject A06, on the other hand, got a performance close to chance level (30.56 %, 22/72). However, when we look to the results of individual tasks averaged across the nine subjects, the differences between classes got smaller, with MI of tongue reaching the highest accuracy (77.16 %, 500/648) and MI of feet the lowest (62.65 %, 406/648). Moreover, in the same confusion matrix one can see that MI of feet was more confused with MI of tongue (159/648) than with MI of left and right hands.

Finally, it is interesting to note that modulation filtering contributed to make the individual differences in overall performance even larger, since among the subjects with good overall performance, subject A09 was the one who benefited the most with modulation-filtering enhancement: he/she from a kappa of 0.68 in Test 1 (no filtering) to 0.77 in Test3 (two-region modulation filtering), an improvement of 13.2 %. Just the opposite, subject A06 worsened even more his/her already poor performance, going from a kappa of 0.32 to 0.3. It is important to remark that the kappa of subject A09 we obtained with our method is statistically significantly better when compared to the kappa value obtained for the

same subject by the 1st place of BCI competition IV, as measured by confidence interval using the binomial distribution.

5. Conclusions

Conventional EEG signal pre-processing approaches for BCI are based on frequency or spatial filtering techniques [29]. In order to increase the classification performance of motor-imagery tasks, we investigated a novel filtering technique, which is performed in the modulation frequency domain, to improve EEG SNR. Visual inspection of the modulation spectrogram was carried out to delimitate the modulation filtering regions, applying to EEG a similar method successfully used by Tobon and Falk for ECG [8]. Such visual analysis to identify the ideal regions for modulation filtering was revealed as an efficient method to observe the energy concentration of EEG recorded during motor imagery tasks. After EEG signal enhancement, we divided the full-band signal into overlapping sub-bands and applied CSP with Tikhonov regularization (TRCSP) [10,11,17,23] for feature extraction. A two-stage classification scheme was then proposed, which relied on six binary LDA classifiers, whose outputs, in turn, served as inputs to a four-class Naive Bayes classifier. For the four-class MI-BCI data available via the BCI competition IV, experimental results showed our proposed solution outperforming the competition's 1st place winner by 3.5 %. The obtained results are promising and motivate future studies to focus on how to find optimal regions in the modulation spectrogram to be filtered out on a per-participant basis. Data-driven methods could be particularly useful in this case. Such exploration is left for future work.

Acknowledgments

This work was supported by the Sa˜o Paulo Research Foundation (FAPESP, grant #2017/15243-7) through equipment and software and by the Brazilian Government through a graduate-student scholarship from the National Council for the Improvement of Higher Education (CAPES).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] D.M. Brandman, L.R. Hochberg, Brain computer interfaces, in: Neurobionics: The Biomedical Engineering of Neural Prostheses, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2016, pp. 231–263.
[2] D.J. Mcfarland, J.R. Wolpaw, Sensorimotor rhythm-based brain – computer regression improves performance, Neural Syst. Rehabil. Eng. IEEE Trans. 13 (3) (2005) 372–379.
[3] Y. Wang, S. Makeig, Decoding intended movement from human EEG in the posterior parietal cortex, Neuroimage 47 (2009) S103.
[4] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, B. Arnaldi, A review of classification algorithms for EEG-based brain-computer interfaces, J. Neural Eng. 4 (June (2)) (2007).
[5] R. Cassani, T.H. Falk, Spectrotemporal modeling of biomedical signals: theoretical foundation and applications, Encycl. Biomed. Eng. (2018) 144–163.
[6] T.H. Falk, Wai-Yip Chan, Modulation filtering for heart and lung sound separation from breath sound recordings, 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (2008) 1859–1862.
[7] S.M. Schimmel, Theory of modulation frequency analysis and modulation filtering, with applications to hearing devices, in: Ph.D. dissertation, University of Washington, 2007, pp. 1–143.
[8] D.P. Tobon, T.H. Falk, Adaptive spectro-temporal filtering for electrocardiogram signal enhancement, IEEE J. Biomed. Health Inform. 22 (March (2)) (2018) 421–428.
[9] D. Wang, D. Miao, G. Blohm, Multi-class motor imagery EEG decoding for brain-computer interfaces, Front. Neurosci. (October) (2012).
[10] F. Lotte, Cuntai Guan, Regularizing common spatial patterns to improve BCI designs: unified theory and new algorithms, IEEE Trans. Biomed. Eng. 58 (February (2)) (2011) 355–362.
[11] A. Meinel, et al., Tikhonov Regularization Enhances EEG-Based Spatial Filtering for Single Trial Regression to Cite this Version: HAL Id: hal-01655755, 2017.
[12] D.J. Krusienski, et al., A comparison of classification techniques for the P300 speller, J. Neural Eng. 3 (2006) 299–305.
[13] T.H. Falk, S. Stadler, W.B. Kleijn, W.Y. Chan, Noise suppression based on extending a speech-dominated modulation band Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, 2, 2007, pp. 1469–1472.
[14] M. Le, et al., Comparison of Hilbert transform and wavelet methods for the analysis of neuronal synchrony, J. Neurosci. Methods 111 (2001) 83–98.
[15] S. Lemm, C. Schaefer, G. Curio, Probabilistic modeling of sensorimotor μ-rhythms for classification of imaginary hand movements, IEEE Trans. Biomed. Eng. 51 (6) (2003) 1077–1080.
[16] B. Blankertz, et al., Review of the BCI competition IV, Front. Neurosci. 6 (July) (2012) 1–31.
[17] K.K. Ang, Z.Y. Chin, C. Wang, C. Guan, H. Zhang, Filter bank common spatial pattern algorithm on BCI competition IV datasets 2a and 2b, Front. Neurosci. 6 (2012) 39, March.
[18] A. Barachant, S. Bonnet, M. Congedo, C. Jutten, Multiclass brain-computer interface classification by Riemannian geometry, IEEE Trans. Biomed. Eng. 59 (4) (2012) 920–928.
[19] K. Sadatnejad, S. Shiry Ghidary, Kernel learning over the manifold of symmetric positive definite matrices for dimensionality reduction in a BCI application, Neurocomputing 179 (2016) 152–160.
[20] A. Davoudi, S.S. Ghidary, K. Sadatnejad, Dimensionality reduction based on distance preservation to local mean for symmetric positive definite matrices and its application in brain-computer interfaces, J. Neural Eng. 14 (3) (2017) 036019.
[21] P. Gaur, R.B. Pachori, H. Wang, G. Prasad, A multi-class EEG-based BCI classification using multivariate empirical mode decomposition based filtering and Riemannian geometry, Expert Syst. Appl. 95 (2018) 201–211.
[22] S. Guan, K. Zhao, S. Yang, Motor imagery EEG classification based on decision tree framework and Riemannian geometry, Comput. Intell. Neurosci. 2019 (2019), 5627156.
[23] H. Ramoser, J. Müller-Gerking, G. Pfurtscheller, Optimal spatial filtering of single trial EEG during imagined hand movement, IEEE Trans. Rehabil. Eng. 8 (4) (2000) 441–446.
[24] Johannes Müller-Gerking, Gert Pfurtscheller, Henrik Flyvbjerg, Designing optimal spatial filters for single-trial EEG classification in a movement task, Clin. Neurophysiol. 110 (1999) 787–798.
[25] A.E. Hassanien, A. Taher, A. Editors, Intelligent Systems Reference Library 74 Brain-Computer Interfaces Current Trends and Applications, 2019.
[26] G.A. Barreto, R.A. Frota, F.N.S. de Medeiros, On the classification of mental tasks: a performance comparison of neural and statistical approaches, Proceedings of the 2004 14th IEEE Signal Processing Society Workshop Machine Learning for Signal Processing (2004) 529–538.
[27] G.H. Rosenfield, K. Fitzpatrick-Lins, A coefficient of agreement as a measure of thematic classification accuracy, Photogramm. Eng. Remote Sens. 52 (2) (1986) 223–227.
[28] F. Lotte, C. Jeunet, Towards improved BCI based on human learning principles, The 3rd International Winter Conference on Brain-Computer Interface (2015) 1–4.
[29] R. Aler, I.M. Galván, J.M. Valls, Evolving spatial and frequency selection filters for brain-computer interfaces, 2010 IEEE World Congr. Comput. Intell. WCCI 2010 - 2010 IEEE Congr. Evol. Comput. CEC 2010 (2010).