Adaptation of Common Spatial Patterns based on Mental Fatigue for Motor-Imagery BCI

Upasana Talukdar[a], Shyamanta M Hazarika[b], John Q. Gan[c]

[a] *Biomimetic & Cognitive Robotics Lab,Dept. of Computer Science & Engineering,Tezpur University, India*
[b] *Mechatronics & Robotics Lab, Dept. of Mechanical Engineering, Indian Institute of Technology, Guwahati, India*
[c] *School of Computer Science and Electronic Engineering, University of Essex, United Kingdom*

Abstract

Common Spatial Pattern (CSP) is the most popular method in motor imagery (MI) based Brain-Computer Interfaces (BCI) for extracting features from electroencephalogram (EEG) signals. Due to the non-stationary nature of EEG signals, the CSP computed on the training data may not be optimal for the evaluation data. One of the major causes of such non-stationarity is the change in user's cognitive state due to fatigue, frustration, low arousal level etc. This paper proposes an adaptive scheme for the CSP based on the mental fatigue of the user. The proposed method uses Linear Discriminant Analysis (LDA) active learning to adapt the CSP. Breaking ties criterion is used for selecting samples from the evaluation data. The separability of MI EEG features extracted with the proposed adaptive CSP has been compared with that of conventional CSP in terms of three separability metrics: Davies Bouldin Index (DBI), Fisher's Score (FS) and Dunn's Index (DI). Experimental results show significantly higher separability of features extracted with adaptive CSP as compared to that with conventional CSP.

Keywords: Brain-Computer Interface, Motor Imagery, Electroencephelogram, Common Spatial Patterns, Adaptation, Mental Fatigue

1. Introduction

Brain-Computer Interface (BCI) provides a direct communication between the brain and external devices by translating brain activity into control commands. Among different kinds of brain

*Corresponding author

Email addresses: upasanat123@gmail.com (Upasana Talukdar), s.m.hazarika@iitg.ernet.in (Shyamanta M Hazarika), jqgan@essex.ac.uk (John Q. Gan)

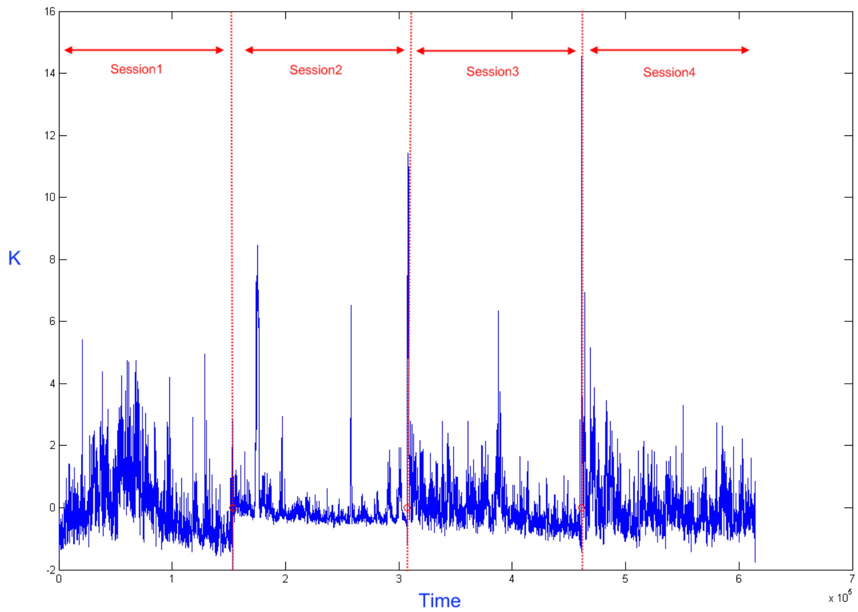

Figure 1: First order reflection coefficients (K) features, extracted from the same subject over four consecutive sessions. Picture taken from Bamdadian [2].

signals that exist, electroencephalogram (EEG) is the most widely used in BCI. This is due to its non-invasiveness, high portability and accessible cost. Imagination of movements called as motor imagery (MI) induce changes in EEG signals. Motor imagery refers to the imagination of self movements without actually performing the movement. This has attracted much attention in the research of EEG-based BCI. Impairment of motor function due to accident, spinal cord injuries or stroke may lead to physical disabilities and to this end EEG MI BCI gives promising solutions.

A major predicament on the reliability of EEG MI BCI is the non-stationary nature of EEG signals. Figure 1 shows the 1st order reflection coefficients (K) features of the EEG signals from the same subject over 4 sessions. It is seen from the figure, that the coefficients as EEG features are shifted in time between the sessions. Shenoy et al. [1] depicted systematic evidence of statistical differences in data recorded during offine and online sessions. This differences across sessions or within sessions leads to the shift in the feature space, which may deteriorate the classification performance.

One of the major causes of such non-stationarity is the cognitive factors like attention, concentration, workload, fatigue, etc.[3]. Repeated accomplishment of MI may make a person feel fatigued. Much cognitive effort is needed to maintain his/her vigilance level while performing MI. Rozand et al. [4] portray in their study that a prolonged session of MI induces mental fatigue. In another study by Myrden et al. [5], it was shown that subjects accomplishing MI tasks often feel fatigued. This change in the fatigue level may cause signal fluctuations which in turn may degrade the separability of signal features. This would eventually deteriorate the performance of EEG MI BCI. Recent work shows that high fatigue level decreases the separability of MI EEG features [6]. Herein lies the motivation for enhancing the separability of MI EEG features based on the fatigue level of the user. The ability to *detect* and *adapt* to the cognitive states like fatigue of a user would enable MI EEG BCI to interact with the user in a much more meaningful way [7].

Adaptation of BCI can be achieved in two ways, one which modifies the user's mental state to enhance the performance and the other which includes modification of the components of BCI to intensify the system's performance. The latter carries out adaptation in different levels of the BCI system such as feature extraction, classification or post processing. The literature has reported different methods to adapt the parameters employed in different classification methods [8, 9, 10, 11]. Besides adaptation of classifiers, adaptation of feature extractors has been also paid much attention [12, 13, 14, 10]. Adaptation can also be carried out in post-processing [11]. Myrden et al. [15] proposed an adaptive BCI based on fatigue, frustration and attention level where adaptation was achieved by resampling the training set and retraining the classifier; training data was resampled based on its proximity to the test data.

In general, to attain an efficient EEG MI BCI, it is substantial to maintain the separability of MI EEG features as high as possible, which can be achieved by using appropriate feature extractors. The literature has reported different methods for extraction of EEG features with high separability [16, 17, 18, 19, 20]. Gaur et al. reported data driven techniques for EEG signal analysis [21, 22, 23, 24, 25, 26].

For MI BCI the most popular feature set is the Common Spatial Pattern (CSP)[27] and most widely used classifier is the Linear Discriminant Analysis (LDA) [28]. CSP is computed based on the training data and hence its performance depends strongly on the quality of the training data. A good training data set has low noise and includes representative MI BCI activities with distinguishable features. However, during high fatigue level, due to non-stationarity of EEG, there may exist a significant change in the observed signal properties / used features between the training and evaluation data, which might make the CSP computed on the training data non-optimal for the evaluation data [6]. This is due to the change in the pattern of the signals during high fatigue level. To address this issue, the training data needs to be updated based on the fatigue level of the user. The literature reports different approaches for adaptation of feature extractors by updating the covariance matrices of CSP. However, although there exist some methods for adapting CSP that address the non-stationary nature of EEG, none of them adapt it based on the cognitive state of the user.

Of late there is a growing interest in designing semi-automatic update of training samples. An interesting solution to update the training data is through active learning which selects the most informative samples from the evaluation data through an iterative process while maintaining the discriminative capabilities as high as possible [29]. Different criteria can be used for selecting the samples so as to update the training data effectively. One such strategy is "breaking ties" [29] which considers the difference between the two highest posterior probabilities obtained from the classifier's output.

The work presented in this paper improves the MI feature separability by integrating the fatigue state of the user into the LDA model based on CSP features. The CSP model computed on the training data is not updated with low fatigue data, while in case of high fatigue data it is updated through updating the training data. Kernel Partial Least Square (KPLS) algorithm [30] is used to monitor the growth of the level of fatigue. Once high fatigue level is detected, the adaptation algorithm updates the training data based on the level of fatigue. Since the CSP model is computed on training data, updating training data leads to the adaptation of the CSP. In this study training data is updated through an active learning approach with LDA using 'Breaking Ties' as sample selection criterion. To the best of our knowledge, despite of the growing interest in active learning, very scarce attention has been paid to designing methods based on this active learning approach for solving the problem of non-stationary EEG data. It is noteworthy that the proposed methodology is different from that of Bhattacharyya et al. [31] who proposed a methodology that amalgamates spatial filtering with tunableQ wavelet transform (TQWT) and carried out their analysis on stereo electroencephalogram (SEEG) signals. SEEG signals are first spatially filtered to isolate the cortical stimulation (CS) artifacts and then, using prior known time-frequency information of the CS artifacts, the method extracted out any remaining significant electro-physiological activity.

In this paper, the separability of the MI EEG features extracted with adaptive CSP (ACSP) has been compared with that of the MI EEG features extracted with conventional CSP (C-CSP). The separability has been analyzed from signal feature perspective in terms of three separability metrics: Davies Bouldin Index (DBI) [32], Fisher's Score (FS) [33] and Dunn's Index (DI) [34]. Lower DBI score implies higher class separability while higher Fisher score and DI value indicates higher class separability. Higher class separability can be considered as better motor imagery performance.

The rest of the paper is organized as follows. The methods are described in Section 2 and Section 3. Section 4 presents experimental results and discusses the findings. Finally, Section 5 concludes the paper.

2. Acquisition of Motor Imagery EEG Data with Mental Fatigue

2.1. Subjects

For this study, EEG data was collected from 11 individuals at University of Essex, England. The subjects were either students or staff of aforesaid university. Before the experiment, the subjects gave their informed consent using a form approved by the Ethics Committee of University of Essex and were paid for their participation. The sample included 5 males and 6 females with a mean age of 29.3 years (SD = 7.4, range = 20-42 years, male mean age = 33 years and female mean age = 26.2 years). The subjects were asked to have a good sleep before the experiment. As reported by them, the mean hours of sleep was found to be 6.8 hours and none of them had any sleep disorder.

2.2. Experimental Protocol

At the beginning of the experiment, the subjects were a) given an orientation to the study; b) asked to read and sign an informed consent document; c) asked to complete a brief demographic questionnaire (name, age, gender, employment status, hours of sleep); d) asked to practice the MI tasks for 5 minutes before EEG data recording.

A short break was provided and thereafter the subjects completed the pre-test self report using Visual Analogue Scale - Fatigue (VAS-F) [35] and Chalder Fatigue Scale (CFS) [36]. The subjects performed 4 different MI tasks during which they need to imagine left hand movement (Class 1), right hand movement (Class 2), both feet movement (Class 3) and tongue movement (Class 4). A session comprises 8 runs with each run lasting 12 mins and producing 80 trials. Each trial begins at t=0 second with a fixation cross that appears on the computer screen along with a short acoustic warning tone followed by a cue at t=2 seconds that appeared either left, right, down or up of the fixation cross indicating the imagination of left hand, right hand, both feet and tongue movement respectively. The subjects accomplished the required task until the cue and the fixation cross disappeared from the screen at t=6 seconds. Thereafter, each trial included a break for 3 seconds. There was no extra breaks between the runs. The paradigm is illustrated in Fig 2. The fatigue state at the end of each run was rated by using a continuous "fatigue scale" which was introduced as a subjective scale with a value from 1 to 5 that extends between two extremes (1 = "Least fatigued" and 5 = "Most fatigued"). The scale chosen by the subject to express the fatigue they were experiencing is taken as the fatigue score of that particular run. Finally, the experiment termination was followed by completing the post-test self report using VAS-F and CFS. Appendix A, Appendix B and Appendix C depicts the VAS-F, CFS and Fatigue Scale respectively.

The subjects carried out the experiment till either they quit due to extreme fatigue or they completed all the 8 runs. Five out of the 11 subjects could not complete the experiment because of high fatigue.

As shown in Table 1 all the participants completed at least 5 runs out of the 8 runs and 6 participants completed all the 8 runs.

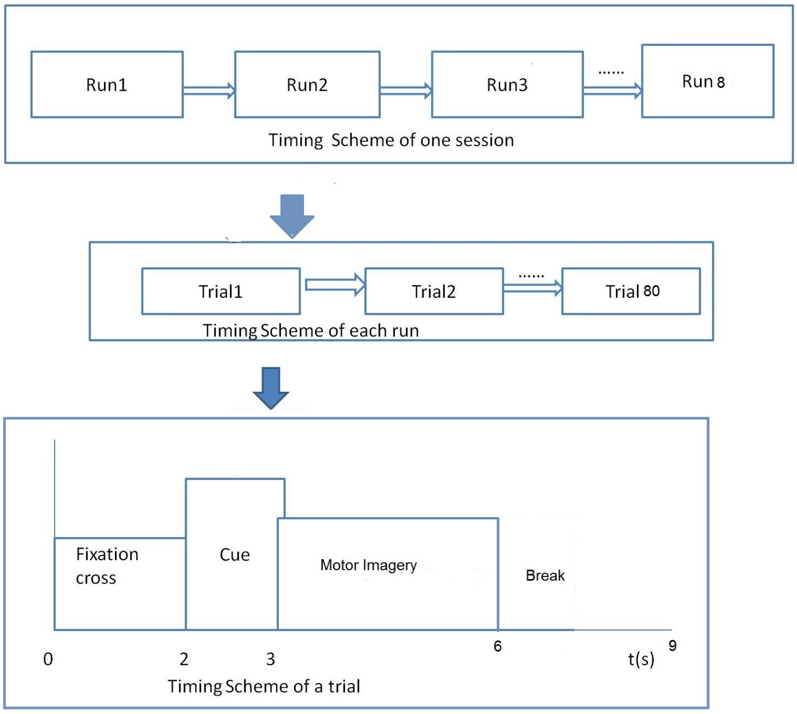

Figure 2: The experimental paradigm

Table 1: Number of participants in each run

Runs	1	2	3	4	5	6	7	8
No. of Participants	11	11	11	11	11	8	7	6

2.3. EEG Recording

The subjects were seated at approximately 80 cm from an LCD screen where the stimuli for the MI task was presented. The Biosemi Active Two System was used to record the EEG data. 64 EEG channels were used to record the data following the 10-20 international montage system [37]. The sampling rate was 256 Hz. The artefacts of the acquired EEG data were removed by EAWICA [38]. The EEG data were then low-pass filtered (40 Hz) and subtracted by common average reference. The EEG data was processed from instant t=0 to t=6 excluding the last 3 seconds break.

3. Methods for EEG Analysis

3.1. Common Spatial Patterns (CSP)

CSP is a two-class supervised method that aims to maximise the feature variation of one class while minimising the feature variation of the other [14]. Hence, CSP aims at achieving optimal discrimination of the features. The construction of CSP takes as input a number of trials of N channel EEG recorded at T consecutive time points and produces an ordered list of distinctive spatial patterns forming a projection matrix W as output [39, 27]. For an EEG trial, the time series is projected onto these discriminating patterns. In brief, the spatially filtered signal Z of a single trial EEG E is obtained by Eq 1.

$$Z = WE \tag{1}$$

where, E is the N×T matrix containing EEG signals, W is the CSP projection matrix, the columns of W^{-1} are the common spatial patterns.

The first and last m rows of Z are then used to compute the variance of the values obtained from this projection. The computed variances are used as CSP features as given in Eq 2.

$$X^p = log(\frac{var(Z^p)}{\sum_{i=1}^{m} var(Z^i) + \sum_{i=N-m-1}^{N} var(Z^i)}), p = 1, ..., m, N-m-1, ..., N \tag{2}$$

where *var()* is the variance calculator and Z^p is the pth row of Z. The logarithmic transformation is used so as to make the distribution of X^p close to Gaussian.

Detailed formulations of CSP can be found in [39]. In the current study m is set to 2.

3.2. Linear Discriminant Analysis (LDA)

Inspite of being a simple classification method, LDA has been proved to be able to achieve performance comparable to other non-linear approaches like SVM or ANN in BCI[40]. Its major advantage is that it requires low computation cost. Other classifiers like KNN or Random Forest are computationally expensive. In case of KNN, it is computationally expensive to find the k nearest neighbours when the dataset is very large, while in case of Random Forest the main drawback is the complexity and the requirement of more computational resources.

LDA classifier is mainly based on the decision function defined as follows:

$$f(x) = W^T x + \omega_0 \tag{3}$$

where x is the sample that need to be discriminated, W is a weight vector while ω_0 is a threshold. The values of the weight vector and the threshold are identified by employing Fisher's criterion on the training data [41]. The classification process is based on the separation by the hyperplane as described in Eq. 4.

$$if\ f(x) > 0\, then\, x \in C_1\, otherwise\, x \in C_2. \tag{4}$$

where C_1 and C_2 are two different classes.

An LDA can only discriminate two classes. To discriminant more than two classes multi-class LDA has been proposed [41]. Multi-class LDA uses different discriminative functions, one for each class [41]. In such case, the discriminant function $f_i(x)$ classifies an unseen sample y as follows:

$$if\ f_i(x) > 0\, then\, x \in C_i. \tag{5}$$

The final label of x will be

$$Label = arg \max_{i \in N} f_i(x) \tag{6}$$

This approach is called as *one against rest*[41].

3.3. Experimental Methodology

In this paper, we update the training data based on the fatigue level of the user. This is achieved in two stages:

1. Track the growth of mental fatigue during MI
2. Based on the mental fatigue state, a new supervised adaptive method is proposed to update CSP.

3.3.1. EEG-based Estimation and Tracking of Mental Fatigue during MI

EEG-based estimation of mental fatigue has been carried out in two stages: a. Track the growth of fatigue during MI and b. Based on the fatigue scores obtained, categorise each run of the experiment into two levels of fatigue: low and high. Detailed analysis can be found in Talukdar et al. [6]. The tracking of growth of fatigue has been carried out on 6 subjects who completed all the 8 runs of the experiment. Spectral power and spectral entropy were employed as features and were estimated in four different frequency bands: δ (0.1-3.5 Hz), θ (4-7.5 Hz), α (8-12 Hz) and β (13-35 Hz) from five different areas of the scalp: frontal (F1, F3, F5, F7,Fz, F2, F4, F6, AFz, AF3, AF4 and FPz), parietal (P1, P3, P5, P7, Pz, P2, P4, P6, POz, PO3, PO4 and CPz), temporal (FT8, T8, TP8, FT7, T7 and TP7), central (C1, C3, C5, Cz, C2, C4,C6) and occipital (O1, O2, Oz). The features exhibiting significant rise during the last run as compared to that during the first run were identified as optimum features and elected for tracking the growth of mental fatigue. The KPLS was used for monitoring the growth of fatigue. The first and the last runs were used as the foundation for analysis of mental fatigue with the first run as active and last run as fatigue state. The input to the KPLS was the optimum feature vector and the two classes, active and fatigue state. The KPLS then provided a fatigue score for each trial as output. The mean of the fatigue scores of each run is taken as the fatigue score of that particular run. The KPLS model was validated with respect to the subjective scores obtained through fatigue scale.

On the basis of the fatigue score obtained through KPLS, each run was classified as either low or high fatigue level using the K-means algorithm. MI EEG class separability was estimated in terms of DBI, DI and FS during each level of fatigue.

3.3.2. Conventional CSP

To evaluate MI EEG separability CSP was used for extracting features. This analysis employed all the 64 electrodes to compute the CSP projection matrix. One of the approach to analyse MI EEG data is to split the data into different time windows and select the optimal temporal segment. Optimal temporal segment refers to the segment that contains the most discriminative information based on a predefined criterion. This study uses optimal spatio-temporal filtering to extract the optimal spatial-temporal patterns and was carried out by employing ADSWIN [27] which is an adaptive sliding window approach that automatically segments EEG trials and then selects the best segments to produce the optimal spatio-temporal patterns. ADSWIN is portrayed as an enhancement to classic sliding window methodology which

- $i$ncreases the class separability,
- $d$ynamically adapts the two parameters window size and overlapping region on which a sliding window approach relies.
- $c$an be applicable to online learning algorithm.

CSP was employed to extract features of each segment obtained through segmentation. DBI is used as a cost function to identify the optimal segment. The EEG segment with minimum DBI is then selected as optimal EEG segment. CSP projection matrix is computed based on the selected optimal time segment to create a spatio-temporal profile. Our motivation to use such a filter is driven by two factors [27]: a. constructing a reduced representation of the original time series of training data and b. consolidating spatial analysis with temporal study, to extract spatio-temporal patterns. It is feasible to process the whole trial, but extraction of optimal time segment has the benefit in computing results with much shorter time segments. Detailed formulation of the method can be found in [27]. The training data was used to build up the optimal spatio-temporal filter. ADSWIN requires two parameters, default segment length (ω_d) and sliding window overlapping region (Δ). It adapts these two parameters to generate the optimal spatio-temporal filter.

3.3.3. Adaptive Common Spatial Patterns (ACSP)

This paper proposes a new supervised adaptive method to update the conventional CSP named as ACSP, which adapts the CSP based on the fatigue state of the user. Class separability of MI EEG features decreases during high fatigue level as compared to low fatigue level [6]. Hence, the proposed approach adapts the CSP during high fatigue level. Algorithm 1 depicts the proposed ACSP through LDA active learning approach. The algorithm starts with the analysis of fatigue during MI. Based on the fatigue scores obtained for each run during the analysis, each run was categorised as either low or high fatigue level. During low fatigue level, no adaptation is carried out, while during high fatigue level the algorithm updates the training data by selecting samples from the evaluation data using 'Breaking Ties' criterion. The updated training data is then used to compute the CSP projection matrix. The process is repeated till it meets the convergence criterion as described by Eq. 7.

$$|f2(k) - f2(k-1)| < \phi, \tag{7}$$

where ϕ is a pre-determined positive constant, k is the kth iteration and

$$f2(k) = \sum_{i=1}^{n} |\frac{(w_i^{(k)})^T S_L^{(k)} (w_i^{(k)})}{(w_i^{(k)})^T S_R^{(k)} (w_i^{(k)})}|, \tag{8}$$

where $w_i^{(k)}$ is the ith column of the CSP projection matrix $W^(k)$, $S_L^{(k)}$ and $S_R^{(k)}$ are computed as follows

$$S_L^{(k)} = \mathfrak{R}^1 - \mathfrak{R}^2, S_R^{(k)} = \mathfrak{R}^1 + \mathfrak{R}^2 \tag{9}$$

with $\mathfrak{R}^1$ and $\mathfrak{R}^2$ being the covariance matrices of Class 1 and Class 2 resectively.

Eq. 8 computes the sum of Rayleigh coefficients. Since the advantage of CSP feature extraction has been demonstrated in the paradigm of maximization of Rayleigh coefficients, the improvement in sum of Rayleigh coefficients leads to better class separability of the features [42]. Hence, this study uses Eq.7 as the convergence criterion of the adaptive scheme, i.e., when the difference between the sum of Rayleigh Coefficients of two consecutive iteration is less than a certain threshold, the algorithm terminates.

Breaking Ties

Breaking ties [29] is an active learning approach to select samples in order to update the training set. In case of binary Support Vector Machine (SVM) with probability outputs P(a) and P(b) for class a and class b, the breaking ties method computes the difference (f1) between P(a) and P(b) and selects the sample with minimum f1. The sample with minimum f1 is considered to be most uncertain, and hence when added to training set, it may improve the classification performance. This study uses LDA scores instead of SVM probabilities to select trials from high fatigue data (H) to update training data (T). The score is computed as in Eq. 10. The method to compute breaking ties is portrayed in Algorithm 2.

$$u = \omega \times x - b \tag{10}$$

Algorithm 1: ACSP Using LDA Active Learning

Input: D=EEG samples with 8 runs, T=initial training samples (first two runs), Y=labels of T, E=evaluation samples (last 6 runs), N_s=number of samples to be added at each iteration, H_s=samples to be added at every iteration, Y_s=predicted labels of H_s

Output: T=final training set, C=final CSP projection matrix

1. Perform fatigue analysis on D to compute fatigue score of each run.
2. Categorize E into two levels based on the fatigue scores: Low (L) and High (H) fatigue level
3. For L, compute CSP Projection matrix (C) on T, extract features, select features.
4. For H,
    (a) Compute CSP Projection matrix (C) on T, extract and select features and train LDA with T and predict labels of H.
    (b) Compute criterion f1_i for each trial i of H by using Algorithm 2.
    (c) Sort the trials of H based on f1 in increasing order.
    (d) Select the first N_s samples from H to form H_s and Y_s.
    (e) Update $T=T+H_s$ and $Y=Y+Y_s$.
    (f) Compute criterion f2 using Eq 8
    (g) Repeat till it meets convergence criterion as in Eq. 7

where x is the sample

$$\omega = (\mu_1 - \mu_2) \times \sigma^{-1} \tag{11}$$

$$b = \omega \times \frac{(n1 \times \mu_1 + n2 \times \mu_2)^T}{n}; \tag{12}$$

$$\sigma = \frac{n1 \times cov1 + n2 \times cov2}{n} \tag{13}$$

n1 and *n2* are the number of samples of class 1 and class 2 respectively, n is the total number of samples, $\mu1$ and $\mu2$ are mean of class 1 and class 2 respectively whereas *cov1* and *cov2* are covariance matrices of class 1 and class 2.

3.3.4. MI EEG Class Separability Evaluation

The separability of MI EEG features extracted by C-CSP and ACSP has been evaluated from the signal-feature perspective using three class separability metrics DBI, DI and FS. Lower value of DBI and higher value of DI and FS indicates higher MI EEG class separability. The higher

Algorithm 2: Trial selection criterion-Breaking Ties

Input: H=high fatigue data, m = number of trials of H, i=1

Output: f1 =[f1_1, f1_2,......, f1_m,

1. Repeat the following steps till i==m
2. Compute the LDA score of each class for trial H_i using Eq10
3. Identify the two classes with the two highest LDA scores, u_{max} and u_{max2}
4. Compute $f1_i$ as difference between u_{max} and u_{max2}
5. i=i+1

the separability of extracted features, the better would be the classification accuracy [43]. Unlike using classifiers, no pre-training is required for assessing the separability using DBI, FS and DI. They are independent of the number of groupings and grouping algorithm used [44] and hence are simple, feasible and time saving [45].

Davies Bouldin Index (DBI)

Davies Bouldin Index (DBI) is computed as follows:

$$M_{ij} = \{\sum_{k=1}^{n} |\mu_{ik} - \mu_{jk}|^q\}^{1/q} \tag{14}$$

$$S_i = \{\frac{1}{T_i}\sum_{l=1}^{T_i}\sum_{k=1}^{n} |x_{lk} - \mu_{ik}|^q\}^{1/q} \tag{15}$$

$$RI_{ij} = \frac{(S_i + S_j)}{M_{ij}} \tag{16}$$

$$DBI = \frac{1}{m}\sum_{i=1}^{m}(\max_{j\neq i} RI_{ij}) \tag{17}$$

where M_{ij} is a measure of separation between class i and class j, S_i is a measure of scatter within class i, μ_i is the centroid of class i, x_l is a feature vector of size n, T_i is the number of feature vectors in class i, m is the number of classes, and the value of q is usually 2. In this study, q is initialized to be 2. Since DBI is the ratio of within-class scatter to between-class distance, a smaller DBI value implies better class separation.

Dunn's Index(DI)

Dunn's index is a measure to evaluate class separability and is defined as follows:

$$DI = \frac{\underset{1\leq i<j\leq m}{min} \delta(\mu_i, \mu_j)}{\underset{1\leq k\leq m}{max} \Delta_k} \tag{18}$$

where m is the number of classes, $\delta(\mu_i, \mu_j)$ is the distance between the centroids of class i and j and Δ_k computes the intra-class distance, i.e., distance of all the points from their centroid μ.

$$\Delta_k = \frac{\sum_{x\in C_i} d(x, \mu)}{|C_i|} \tag{19}$$

where C_i represents class i. $|C_i|$ is the number of samples of a particular class. A larger value of DI implies better class separability.

Fisher Score (FS)

Fisher Score is computed as follows:

$$FS = \frac{|\mu_1 - \mu_2|^2}{|\sigma_1^2 - \sigma_2^2|} \tag{20}$$

where μ_1 and μ_2 represent the centroids of class 1 and 2 while σ_1 and σ_2 represent the variance of class 1 and class 2 respectively. Higher value of FS implies better class separability.

4. Results and Discussions

4.1. Subjective Evaluation of Fatigue

Subjective evaluation of fatigue through self-report using VAS-F and CFS reveals that MI produces considerable amount of fatigue. The results as shown in Fig. 3 portray the growth of fatigue level after the completion of MI experiment as compared to that at the beginning. Friedman statistical test was conducted on the averaged subjective scores obtained before the beginning and after the completion of the experiment and the results are shown in Table 2 which show a significant rise in the averaged subjective scores on mental fatigue after the experiment (post-task) as compared to that at the beginning of the experiment (pre-task).

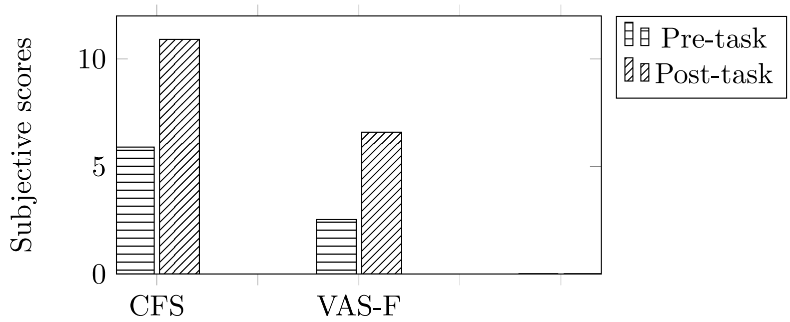

Figure 3: Comparison of subjective scores on mental fatigue averaged before and after the experiment.

Table 2: Statistical test on the subjective scores

CFS	VAS-F
p=0.001565402 (p<0.05)	p=0.001565402 (p<0.05)

4.2. EEG-based Estimation of Mental Fatigue

The growth of fatigue was then tracked using the KPLS algorithm which comprises of two key steps: i. KPLS model selection and ii. KPLS model prediction. During KPLS model selection, the optimum number of KPLS components (KPLS latent vectors) that provides the maximum classification accuracy with LDA has been identified. For estimating the optimum number of components, the first and the last runs were used as training as well as testing data, with the first run as active state and the last run as fatigue state. The KPLS components were evaluated in the range of 1 to 10. During KPLS model prediction, the KPLS score of each trial of all the runs was predicted. The KPLS takes as input the optimum feature vector as discussed in Section 3.3.1 and two class labels, active (-1) and fatigue (+1), and predicts KPLS scores as output. In our work [6], the analysis of EEG spectral power and spectral entropy from different frequency bands in different areas of the scalp showed that spectral power increases during the last run as compared to that of the first run. δ, θ and α power from frontal region, α power from parietal lobe, δ, θ and α power from temporal lobe and θ power from occipital lobe have been considered as optimal feature vector since they show significant increase during the last run as compared to that of the first run; while β power from all the lobes shows insignificant change between the first and last runs. The KPLS scores are the projection of explanatory variables onto the KPLS regression coefficients. The KPLS score of each trial is interpreted as fatigue score of that trial. Negative scores indicates low fatigue

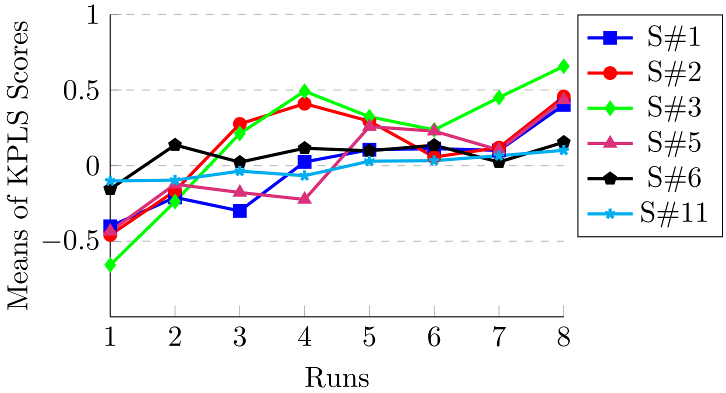

Figure 4: Means of KPLS scores for each run with 6 subjects

level and positive scores high fatigue level. The mean of the KPLS scores of each run was then computed, which is interpreted as the fatigue score of that particular run. The growth of fatigue based on the KPLS model prediction is shown in Fig. 4. The figure shows the orderly progression from active to fatigue state with some intermittent reversals. This is because fatigue may not have a perfect monotonic increase over time and sometimes waxing and waning behaviours can be observed.

The KPLS model was then validated with respect to the subjective scores obtained through the ”fatigue scale”. The trends for fatigue scores estimated through KPLS model and fatigue scores obtained through subjective scores were computed using the Centered Moving Average algorithm. The results portray that the KPLS model in case of 5 subjects (out of 6) showed strong correlation (>0.85) with the fatigue scores obtained through fatigue scale [6]. This is because in case of subject 11 no correlation is found since the subjective score was same throughout the whole experiment while the fatigue score based on KPLS model shows orderly progression towards fatigue state. Talukdar et. al. [6] has discussed the detailed formulation for fatigue analysis. Further analysis has been carried out on the 5 subjects that showed strong correlation between fatigue scores estimated through KPLS model and the fatigue scores obtained through fatigue scale.

Based on these computed fatigue scores, each run was then either categorised as low or high fatigue state by using the K-means algorithm [6]. Table 3 shows the runs that comes under low and high fatigue level based on the computed fatigue score. The first column shows the subject id while the second and the third columns show the runs categorized as low and high fatigue level

Table 3: Runs categorized as low or high fatigue level

Subject	Low fatigue	High fatigue
#1	3,4	5,6,7,8
#2	6,7	3,4,5,8
#3	3,5,6	4,7,8
#5	3,4,7	5,6,8
#6	3,7	4,5,6,8

respectively.

4.3. Conventional CSP (C-CSP)

CSP was employed for extracting features of each segment obtained through ADSWIN. DBI was used as a cost function to identify the optimal segment. The EEG segment with minimum DBI was then selected as optimal EEG segment. CSP projection matrix was computed based on the selected optimal time segment to create a spatio-temporal profile. In this study, we use two different value of ω_d (2.99 seconds, i.e, 767 time points and 3.99 seconds, i.e., 1023 time points) because ω_d lower than 2.99 seconds would be too small to find the optimal number of time points and ω_d higher than 3.99 seconds may create large ω_d. Two different values of Δ (0.5 sec, i.e, 28 time points and 1 sec, i.e, 256 time points) were examined. The values of Δ were chosen in such a way that it maintains even distributions along the trial. Also it was set by keeping in mind that the larger value of Δ would keep more historic information rather than new information. Six different band pass filters were examined (4-7 Hz, 8-13 Hz, 13-30 Hz, 30-40 Hz; 4-9 Hz, 9-15 Hz, 15-30 Hz, 30-40 Hz; 4-9 Hz, 9-16 Hz, 15-30 Hz, 30-40 Hz; 4-9 Hz, 9-16 Hz, 15-32 Hz, 30-40 Hz; 4-9 Hz, 9-16 Hz, 15-32 Hz, 30-40 Hz; 4-9 Hz, 9-16 Hz, 15-32 Hz, 30-40 Hz) and the best two (FB_1: 4-9 Hz, 9-17 Hz, 15-30 Hz, 30-40 Hz and FB_2 :4-9 Hz, 9-16 Hz, 15-32 Hz, 30-40 Hz) were chosen for the study. Based on the different combinations of ω_d and Δ, four segments were then investigated. The spatio-temporal patterns was then used to extract the features and KPLS-mRMR [46] selects the features. The number of selected features was set to 25. Parameters that have been employed for filtering are shown in Table 4. Based on the different combinations of ω_d and Δ, four segments were then investigated for each filter bank as shown in Table 5.

Table 4: Parameters used by ADSWIN

Parameters	Alias	Values
Filter Bank	FB_1 FB_2	4-9 Hz, 9-17 Hz, 15-30 Hz, 30-40 Hz 4-9 Hz, 9-16 Hz, 15-32 Hz, 30-40 Hz
Default segment length	ω_{d1} ω_{d2}	best of the range [3.6 4.5] seconds, i.e, [922 1152] time points best of the range [2.99 3.5] seconds, i.e., [767 896] time points
Overlapping region	Δ_1 Δ_2	0.5 sec (256 time points) 1 sec (128 time points)
No. of selected features		25

Table 5: Segments investigated for analysis of effect of fatigue on motor imagery

Segments	Alias
ω_{d1} and Δ_1	Seg_1
ω_{d1} and Δ_2	Seg_2
ω_{d2} and Δ_1	Seg_3
ω_{d2} and Δ_2	Seg_4

4.4. Adaptive Common Spatial Patterns : ACSP

4.4.1. Training and Testing Sessions:

The first two runs were used as training data to compute the optimal spatio-temporal filter with the parameters as shown in Table 4. The last 6 runs were used as evaluation data. Based on the CSP model computed on the training data, CSP features were extracted for the evaluation data. Evaluation data were categorized as low and high fatigue level by K-means algorithm as described in Section 3.3.1. The training data was then updated with the samples from high fatigue data using Algorithm 1 and the CSP model was adapted using the updated training data.

4.4.2. Results

The MI EEG separability in terms of average DBI, FS and DI across all the 5 subjects with C-CSP and ACSP during high fatigue level are shown in Table 6. The table shows that the average DBI value across 5 subjects is lower with ACSP as compared to that with C-CSP in all the segments while average FS and DI are higher with ACSP as compared to that with C-CSP for all the segments. These collectively indicates that the class separability of MI EEG features

Table 6: Average DBI, FS and DI values across all the 5 subjects for MI EEG features extracted with C-CSP and ACSP

Segment	Bandpass filter	CSP	DBI	p-value	Fisher's Score	p-value	Dunn's index	pvalue
ω_d=1023, Δ=256	FB_1	C-CSP ACSP	17.4 15.2	0.32(×)	0.1 0.35	0.32 (×)	0.09 0.13	0.04✓)
	FB_2	C-CSP ACSP	17.2 16	0.62(×	0.17 0.23	0.32(×)	0.102 0.114	0.04(✓)
ω_d=1023, Δ=128	FB_1	C-CSP ACSP	19 16.2	0.04(✓)	0.18 0.23	0.32 (×)	0.09 0.11	0.32(×)
	FB_2	C-CSP ACSP	17.8 16	0.04(✓)	0.14 0.23	0.04(✓)	0.084 0.114	0.04(✓)
ω_d=767, Δ=256	FB_1	C-CSP ACSP	16.6 16	1 (×)	0.13 0.4	0.04(✓)	0.095 0.12	0.32×)
	FB_2	C-CSP ACSP	19 16	0.04(✓)	0.12 0.18	0.04(✓)	0.089 0.12	0.04(✓
ω_d=767, Δ=128	FB_1	C-CSP ACSP	17.8 16.8	1 (×)	0.12 0.21	0.32(×)	0.089 0.11	(0.32×)
	FB_2	C-CSP ACSP	19.8 17	0.04(✓)	0.11 0.22	0.04(✓)	0.08 0.11	0.04(✓)

extracted with ACSP is higher than that with the C-CSP in terms of all the three evaluation metrics. Friedman statistical test was carried out to examine its statistical difference. The p-value of the test is shown in the 5th colum for DBI, 7th column for FS and 9th column for DI. In the table, the significant difference in class separability between C-CSP and ACSP is indicated by ✓ while × indicates insignificant difference.

The results portray that MI EEG features extracted with ACSP are more separable as compared to those extracted with C-CSP in terms of all the three separability metrics. The results collectively suggest that the identified mental fatigue level can be used as a metric to adapt the CSP so as to improve the MI EEG class separability and hence presents an initial step towards the design of EEG MI BCI that can adapt to the changes in fatigue level. The results show that the adaptive approach when applied to FB_2 performs better as compared to that of FB_1. Also, the adaptive approach when applied to Seg_1 does not produce significant improvement in separability of MI EEG features compared to Seg_2. This is because in the proposed method the class labels of the high fatigue data were estimated first and then the trial from high fatigue data along with its predicted label was added. But due to non-stationarity of EEG, it may be possible that the new trial is mis-classified and the training data is updated based on the erroneous classification which can affect BCI performance. Hence, an unsupervised adaptation of the CSP may hold promise.

The study is based on offline analysis that is often over-optimistic regarding its conclusions [15]. However, there is a reason to believe that adaptation of CSP based on fatigue level can improve its performance. Repeated accomplishment of MI often makes the subjects experience mental fatigue which causes shift between the training and evaluation data and hence adapting CSP during high fatigue level holds promise. The relationship between mental fatigue and MI is cyclical and a prolonged MI session induces mental fatigue which in turn deteriorates MI EEG separability [6]. Hence, the effectiveness of adapting CSP based on fatigue state may be more credible in case of online analysis.

5. Conclusion

This study proposes an adaptive scheme using LDA active learning approach to adapt the CSP of MI EEG based on the fatigue state of the user. The CSP is adapted through semi-automatic update of training data. Braking ties criterion is used to select samples from the evaluation data so as to update the training data. The CSPs are then computed based on the updated training data. Adaptation is activated during high fatigue state. The results collectively show that the class separability of MI EEG is significantly improved with ACSP as compared to C-CSP. However, the study is made supervised and offline. Future work would include an online and unsupervised CSP adaptation.

Acknowledgements

Financial support from MHRD as Centre of Excellence on Machine Learning Research and Big Data Analysis is gratefully acknowledged. Assistance received from DST-UKEIRI Project: DST/INT/UK/P-91/2014 is also acknowledged.

References

[1] P. Shenoy, M. Krauledat, B. Blankertz, R. P. Rao, K.-R. Müller, Towards adaptive classification for BCI, Journal of Neural Engineering 3 (1) (2006) R13.

[2] A. Bamdadian, Towards prediction and improvement of EEG-based MI-BCI performance., Ph.D. thesis, National University of Singapore (2014).

[3] A. Myrden, T. Chau, Effects of user mental state on EEG-BCI performance, Name: Frontiers in Human Neuroscience 9 (2015) 308.

[4] V. Rozand, F. Lebon, P. J. Stapley, C. Papaxanthis, R. Lepers, A prolonged motor imagery session alter imagined and actual movement durations: potential implications for neurorehabilitation, Behavioural Brain Research 297 (2016) 67–75.

[5] A. Myrden, T. Chau, Effects of user mental state on EEG-BCI performance, Frontiers in Human Neuroscience 9 (2015) 308.

[6] U. Talukdar, S. M. Hazarika, J. Q. Gan, Motor imagery and mental fatigue : inter-relationship and EEG based estimation, Journal of Computational Neuroscience 46 (1) (2018) 55–77.

[7] A. Nijholt, D. F. Tan, Brain-Computer Interfacing for Intelligent Systems, Intelligent Systems, IEEE 23 (3) (2008) 72–79.

[8] C. Vidaurre, A. Schlogl, R. Cabeza, R. Scherer, G. Pfurtscheller, A fully on-line adaptive BCI, IEEE Transactions on Biomedical Engineering 53 (6) (2006) 1214–1219.

[9] H. Raza, H. Cecotti, Y. Li, G. Prasad, Adaptive learning with covariate shift-detection for motor imagery-based brain–computer interface, Soft Computing 20 (8) (2016) 3085–3096.

[10] A. Chowdhury, H. Raza, Y. K. Meena, A. Dutta, G. Prasad, Online covariate shift detection based adaptive brain-computer interface to trigger hand exoskeleton feedback for neuro-rehabilitation, IEEE Transactions on Cognitive and Developmental Systems.

[11] L. F. Nicolas-Alonso, R. Corralejo, J. Gomez-Pilar, D. Álvarez, R. Hornero, Adaptive semi-supervised classification to reduce intersession non-stationarity in multiclass motor imagery-based brain–computer interfaces, Neurocomputing 159 (2015) 186–196.

[12] A. Bamdadian, C. Guan, K. K. Ang, J. Xu, Online semi-supervised learning with KL distance weighting for motor imagery-based BCI, in: Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2012, pp. 2732–2735.

[13] Y. Li, C. Guan, An extended EM algorithm for joint feature extraction and classification in brain-computer interfaces, Neural Computation 18 (11) (2006) 2730–2761.

[14] X. Song, S.-C. Yoon, Improving Brain-Computer Interface classification using adaptive common spatial patterns, Computers in Biology and Medicine 61 (2015) 150–160.

[15] A. Myrden, T. Chau, Towards psychologically adaptive brain–computer interfaces, Journal of Neural Engineering 13 (6) (2016) 066022.

[16] A. Bhattacharyya, M. Sharma, R. B. Pachori, P. Sircar, U. R. Acharya, A novel approach for automated detection of focal eeg signals using empirical wavelet transform, Neural Computing and Applications 29 (8) (2018) 47–57.

[17] A. Bhattacharyya, R. B. Pachori, A multivariate approach for patient-specific EEG seizure detection using empirical wavelet transform, IEEE Transactions on Biomedical Engineering 64 (9) (2017) 2003–2015.

[18] A. Bhattacharyya, R. Pachori, A. Upadhyay, U. Acharya, Tunable-q wavelet transform based multiscale entropy measure for automated classification of epileptic EEG signals, Applied Sciences 7 (4) (2017) 385.

[19] A. Bhattacharyya, V. Gupta, R. B. Pachori, Automated identification of epileptic seizure EEG signals using empirical wavelet transform based Hilbert marginal spectrum, in: The 22nd International Conference on Digital Signal Processing (DSP), IEEE, 2017, pp. 1–5.

[20] A. Bhattacharyya, L. Singh, R. B. Pachori, Identification of epileptic seizures from scalp eeg signals based on tqwt, in: Machine Intelligence and Signal Analysis, Springer, 2019, pp. 209–221.

[21] P. Gaur, R. B. Pachori, H. Wang, G. Prasad, An empirical mode decomposition based filtering method for classification of motor-imagery eeg signals for enhancing brain-computer interface, in: 2015 International Joint Conference on Neural Networks (IJCNN), IEEE, 2015, pp. 1–7.

[22] 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 Systems with Applications 95 (2018) 201–211.

[23] P. Gaur, R. B. Pachori, H. Wang, G. Prasad, A multivariate empirical mode decomposition based filtering for subject independent bci, in: 2016 27th Irish Signals and Systems Conference (ISSC), IEEE, 2016, pp. 1–7.

[24] P. Gaur, R. B. Pachori, H. Wang, G. Prasad, Enhanced motor imagery classification in eeg-bci using multivariate emd based filtering and csp features, in: International Brain-Computer Interface (BCI) Meeting 2016, 2016.

[25] P. Gaur, R. B. Pachori, H. Wang, G. Prasad, An automatic subject specific intrinsic mode function selection for enhancing two-class eeg based motor imagery-brain computer interface, IEEE Sensors Journal 19 (16) (2019) 69386947.

[26] P. Gaur, G. Kaushik, R. B. Pachori, H. Wang, G. Prasad, Comparison analysis: Single and multichannel emd-based filtering with application to BCI, in: Machine Intelligence and Signal Analysis, Springer, 2019, pp. 107–118.

[27] U. Talukdar, S. M. Hazarika, Designing optimal spatio-temporal filter for single trial EEG based BCI, in: The 3rd International Conference on Advances in Robotics (AIR), ACM, 2017.

[28] V. Mishuhina, X. Jiang, Feature weighting and regularization of common spatial patterns in EEG-based motor imagery BCI, IEEE Signal Processing Letters 25 (6) (2018) 783–787.

[29] E. Pasolli, F. Melgani, D. Tuia, F. Pacifici, W. J. Emery, SVM active learning approach for image classification using spatial information, IEEE Transactions on Geoscience and Remote Sensing 52 (4) (2014) 2217–2233.

[30] R. Rosipal, L. J. Trejo, Kernel partial least squares regression in reproducing kernel hilbert space, Journal of Machine Learning Research 2 (Dec) (2001) 97–123.

[31] A. Bhattacharyya, R. Ranta, S. Le Cam, V. Louis-Dorr, L. Tyvaert, S. Colnat-Coulbois, L. Maillard, R. B. Pachori, A multi-channel approach for cortical stimulation artefact suppression in depth eeg signals using time-frequency and spatial filtering, IEEE Transactions on Biomedical Engineering (7) (2018) 1915–1926.

[32] D. L. Davies, D. W. Bouldin, A cluster separation measure, IEEE Transactions on Pattern Analysis and Machine

Intelligence (2) (1979) 224–227.

[33] R. O. Duda, P. E. Hart, D. G. Stork, Pattern Classification, Wiley, New York, 1973.

[34] J. C. Dunn, A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters, Journal of Cybernetics 3 (3) (1973) 32–57.

[35] K. A. Lee, G. Hicks, G. Nino-Murcia, Validity and reliability of a scale to assess fatigue, Psychiatry Research 36 (3) (1991) 291–298.

[36] M. Cella, T. Chalder, Measuring fatigue in clinical and community settings, Journal of Psychosomatic Research 69 (1) (2010) 17–22.

[37] R. W. Homan, J. Herman, P. Purdy, Cerebral location of international 10–20 system electrode placement, Electroencephalography and Clinical Neurophysiology 66 (4) (1987) 376–382.

[38] N. Mammone, F. C. Morabito, Enhanced automatic wavelet independent component analysis for electroencephalographic artifact removal, Entropy 16 (12) (2014) 6553–6572.

[39] J. Müller-Gerking, G. Pfurtscheller, H. Flyvbjerg, Designing optimal spatial filters for single-trial EEG classification in a movement task, Clinical neurophysiology 110 (5) (1999) 787–798.

[40] B. Blankertz, K.-R. Muller, D. J. Krusienski, G. Schalk, J. R. Wolpaw, A. Schlogl, G. Pfurtscheller, J. R. Millan, M. Schroder, N. Birbaumer, The BCI Competition III: Validating alternative approaches to actual BCI problems, IEEE Transactions on Neural Systems and Rehabilitation Engineering 14 (2) (2006) 153–159.

[41] J. Asensio-Cubero, Multiresolution analysis over graphs for brain computer interfacing, Ph.D. thesis, University of Essex (2014).

[42] Y. Li, C. Guan, A semi-supervised SVM learning algorithm for joint feature extraction and classification in brain computer interfaces, in: The 28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS'06), 2006, pp. 2570–2573.

[43] B. A. S. Hasan, Adaptive methods exploiting the time structure in eeg for self-paced brain-computer interfaces, Ph.D. thesis, University of Essex (2010).

[44] T. Adel, A. Wong, D. Stashuk, A weakly supervised learning approach based on spectral graph-theoretic grouping, arXiv preprint arXiv:1508.00507.

[45] T. Löster, Determining the optimal number of clusters in cluster analysis, in: The 10th International Days of Statistics and Economics,Prague, 2016.

[46] U. Talukdar, S. M. Hazarika, J. Q. Gan, A kernel partial least square based feature selection method, Pattern Recognition 83 (2018) 91–106.

Appendix A. Visual Analogue Scale- Fatigue

Table A.7: Visual Analogue Scale-Fatigue

1	not at all tired	0 1 2 3 4 5 6 7 8 9 10	extremely tired
2	not at all sleepy	0 1 2 3 4 5 6 7 8 9 10	extremely sleepy
3	not at all drowsy	0 1 2 3 4 5 6 7 8 9 10	extremely drowsy
4	not at all fatigued	0 1 2 3 4 5 6 7 8 9 10	extremely fatigued
5	not at all worn out	0 1 2 3 4 5 6 7 8 9 10	extremely worn out
6	not at all energetic	0 1 2 3 4 5 6 7 8 9 10	extremely energetic
7	not at all active	0 1 2 3 4 5 6 7 8 9 10	extremely active
8	not at all vigorous	0 1 2 3 4 5 6 7 8 9 10	extremely vigorous
9	not at all efficient	0 1 2 3 4 5 6 7 8 9 10	extremely efficient
10	not at all lively	0 1 2 3 4 5 6 7 8 9 10	extremely lively
11	not at all bushed	0 1 2 3 4 5 6 7 8 9 10	totally bushed
12	not at all exhausted	0 1 2 3 4 5 6 7 8 9 10	totally exhausted
13	keeping my eyes open is no effort at all	0 1 2 3 4 5 6 7 8 9 10	keeping my eyes open is a tremendous chore
14	moving my body is no effort at all	0 1 2 3 4 5 6 7 8 9 10	moving my body is a tremendous chore
15	concentrating is no effort at all	0 1 2 3 4 5 6 7 8 9 10	concentrating is a tremendous chore
16	carrying on a conversation is no effort at all	0 1 2 3 4 5 6 7 8 9 10	carrying on a conversation is a tremendous chore
17	I have absolutely no desire to close my eyes	0 1 2 3 4 5 6 7 8 9 10	I have a tremendous desire to close my eyes
18	I have absolutely no desire to lie down	0 1 2 3 4 5 6 7 8 9 10	I have a tremendous desire to lie down

Appendix B. Chalder Fatigue Scale

Table B.8: Chalder Fatigue Scale

		less than usual	no more than usual	more than usual	much more than usual
1	do you have problems with tiredness?				
2	do you need to rest more?				
3	do you feel sleepy or drowsy?				
4	do you have problems starting things?				
5	do you lack energy?				
6	do you have less strength in your muscles?				
7	do you feel weak?				
8	do you have difficulties concentrating?				
10	how is your memory?				
		better than usual	no worse than usual	worse than usual	much worse than usual
11	do you find it more difficult to find the right word?				

Appendix C. Fatigue Scale (FS) at the end of each run

Table C.9: Fatigue Scale (FS) at the end of each run

1	Least fatigued	1 2 3 4 5	Most fatigued
2	Least fatigued	1 2 3 4 5	Most fatigued
3	Least fatigued	1 2 3 4 5	Most fatigued
4	Least fatigued	1 2 3 4 5	Most fatigued
5	Least fatigued	1 2 3 4 5	Most fatigued
6	Least fatigued	1 2 3 4 5	Most fatigued
7	Least fatigued	1 2 3 4 5	Most fatigued
8	Least fatigued	1 2 3 4 5	Most fatigued