This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2015.2466079, IEEE Transactions on Neural Systems and Rehabilitation Engineering

A Boosting-based Spatial-Spectral Model for Stroke Patients' EEG Analysis in Rehabilitation Training

Ye Liu, Hao Zhang, Min Chen and Liqing Zhang*

***Abstract*—Studies have shown that Motor Imagery Electroencephalogram (EEG) based Brain-Computer Interface (BCI) system can be used as a rehabilitation tool for stroke patients. Efficient classification of EEG from stroke patients is a fundamental in the BCI-based stroke rehabilitation systems. One of the most successful algorithms for EEG classification is the Common Spatial Patterns (CSP). However, studies have reported that the performance of CSP heavily relies on its operational frequency band and channels configuration. To the best of our knowledge, there is no agreed clinical conclusion about motor imagery patterns of stroke patients. In this case, it is not available to obtain the active channels and frequency bands related to brain activities of stroke patients beforehand. Hence, for using the CSP algorithm, we usually set a relatively broad frequency range and channels, or try to find a subject-related frequency bands and channels. To address this problem, we propose an adaptive boosting algorithm to perform autonomous selection of key channels and frequency band. In the proposed method, the spatial-spectral configurations are divided into multiple preconditions, and a new heuristic supervisor of stochastic gradient boost strategy is utilized to train weak classifiers under these preconditions. Extensive experiment comparisons have been performed on three datasets including two benchmark datasets from the famous BCI competition III and BCI competition IV as well as one self-acquired dataset from stroke patients. Results show that our algorithm yields relatively higher classification accuracies compared with seven state-of-the-art approaches. In addition, the spatial patterns (spatial weights) and spectral patterns (bandpass filters) determined by the algorithm can also be used for further analysis of the data, e.g., for brain source localization and physiological knowledge exploration.**

***Index Terms*—Brain Computer Interface (BCI), Stroke Rehabilitation, Boosting, Spatial-Spectral Analysis.**

I. Introduction

Brain-computer interface (BCI) provides a direct communication pathway between a human brain and an external device [1], [2]. Thus, BCI allows stroke patients to use their brain signals for communication and control [3], [4]. In recent years, BCI has been widely used in restoring motor functions in stroke patients, as a rehabilitation tool [3], [5].

Among assorts of different brain diffused signals, electroencephalogram (EEG) is one of the most studied brain signals in BCI due to its low cost and noninvasive acquisitions. Motor imagery EEG has been widely performed recently because of its discriminative property and inexpensive acquisitions. A novel application in motor imagery study is combining BCI with clinical rehabilitation training therapies for strokes [5], [6]. In order to control external devices using EEG-based rehabilitation systems, it is a fundamental to provide reliable recognition accuracies of patients' motor imagination of their movements. Effort has been made to explore suitable signal processing and classification algorithms in [7]–[9].

Among various approaches developed for EEG signals, Common Spatial Patterns (CSP) has been proved to be one of the most effective algorithms [10]. In CSP, it tries to classify two classes of EEG where the variance of one class is maximized while the variance of the other class is minimized [8]. However, studies have shown that the performance of CSP algorithm heavily depends on its operational frequency bands and channels configuration [3], [11]. To the best of our knowledge, there is no agreed clinical conclusion about motor imagery patterns of stroke patients. Stroke patients have significantly impaired motor imagery cognitive process due to stroke [12]. Moreover, some studies have shown that event-related desynchronization (ERD) in stroke patients, which is often used as a neural marker representing cortex excitability [13], is significantly lower than that in healthy subjects [12], [14]. In this case, it is not available to obtain the active channels and frequency bands related to brain activities of stroke patients beforehand. We usually set a relatively broad frequency range and channels or try to select a subject-specific frequency range and channels when applying CSP on strokes.

To address the problem of manually selecting the operational subject-specific frequency band and channels group, several approaches have been proposed. Yang *et al.* proposed a novel channel selection method by measuring the inconsistencies from the outputs of the multiple classifiers [15], while Chin *et al.* proposed DCA approach and DCR approach to select subject-specific discriminative channels by iteratively adding or removing channels based on the classification accuracies [16]. For optimization of the spectral filter, several novel approaches, namely, Common Spatio-Spectral Pattern (CSSP) [17], Common Sparse Spectral Spatial Pattern (CSSSP) [18], Iterative Spatio-Spectral Patterns Learning (ISSPL) [19], Filter Bank Common Spatial Pattern (FBCSP) [20] were proposed.

Despite various studies and recent advances, it is still a challenging and open issue to extract optimal spatial-spectral filters, especially for stroke patients. Moreover, since stroke patients are potential audiences of BCI, it is highly desirable to design an efficient and robust discriminative model suitable for both normal people and stroke patients. For

Y. Liu and L. Zhang are with Key Laboratory of Shanghai Education Commission for Intelligent Interaction and Cognitive Engineering, Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai, 200240, China. *Corresponding author (zhang-lq@cs.sjtu.edu.cn).

H. Zhang is with The Robotics Institute, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, 15213, USA.

M. Chen is with Credit Operations Department, Citibank (China) Co. Ltd, Shanghai, 200120, China.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2015.2466079, IEEE Transactions on Neural Systems and Rehabilitation Engineering

this purpose, in this paper, an adaptive boosting algorithm, termed Common Spatial-Spectral Boosting Pattern (CSSBP), is proposed to promote the performance of decoding EEG patterns from stroke patients by a simultaneous optimization of the frequency filter and spatial filter. It aims to model the usually predetermined spatial-spectral configurations into variable preconditions. Under these preconditions, we introduce a new heuristic supervisor of stochastic gradient boost strategy for training the weak classifiers. In this process, the most important channel groups and frequency bands related to brain activities are produced by our algorithm. To the best of our knowledge, this is the first study of combing CSP with boosting strategy to solve a challenge problem of classifying and decoding motor imagery EEG of stroke patients.

The remainder of this paper is organized as follows. A detailed formulation of CSSBP is illuminated in Section II. Section III briefly describes the experimental arrangement and data acquisition. Section IV displays comparison results among some state-of-the-art algorithms and our proposed CSSBP algorithm. Finally, we give a brief discussion and conclusion about our work in Section V and VI, respectively.

II. Common Spatial-Spectral Boosting Pattern

In this section, we give a detailed description of CSSBP algorithm including the problem modeling and model learning methods. CSSBP comprises four progressive stages of EEG measurements processing: multiple spatial filtering and band-pass filtering, feature extraction using the CSP algorithm, weak classifiers training using CSP features, and pattern recognition using a combination model. Fig. 1 shows the architecture of CSSBP algorithm.

A. Problem Modeling

In BCI-based applications, there are two issues which have to be pre-defined as a default without deliberations: the channels configuration and frequency band for EEG analysis. A universal configuration usually leads to poor performance in practical applications, as different subjects have the variability in their EEG patterns. Therefore, an improved dynamic configuration is desirable and demanding in practical systems.

For the development to follow, we first introduce some notations and the key point of our proposed algorithm. We use $E_{train} = \{x_n, y_n\}_{\mathrm{n=1}}^{N}$ to denote the training dataset, where E_n is the n^{th} sample and y_n is its corresponding label. The objective of CSSBP is that by using a universal set of all possible pre-conditions $\mathcal{V}$ we try to find a subset $\mathcal{W} \subset \mathcal{V}$ which produces a combination model F by combing all sub-models learned under condition $W_k(W_k \in \mathcal{W})$ and minimize the classification error on the dataset E_{train}:

$$\mathcal{W}^* = \arg\min_{\mathcal{W}} \frac{1}{N} |\left\{E_n : F\left(x_n, \mathcal{W}\right) \neq y_n\right\}_{n=1}^{N}| \tag{1}$$

Next, we will formulate three homogeneous problems and then adopt an adaptive boost algorithm to solve the problems.

Spatial Channel Selection The channels group is denoted as $\mathcal{C}$, while $\mathcal{U}$ is used to represent a universal set composed of all possible channels subsets such that each subset $\mathcal{U}_k$ in $\mathcal{U}$ satisfies $|\mathcal{U}_k| \leq |\mathcal{C}|$, here we use $|.|$ to represent the size of the corresponding set. Thus for channel selection, we try to find an optimal channel set $\mathcal{S}(\mathcal{S} \subset \mathcal{U})$, where an optimal combination classifier F is constructed by combining weak classifiers trained under kinds of channel set preconditions. Thus we can obtain:

$$F(E_{train}; \mathcal{S}) = \sum_{\mathcal{S}_k \in \mathcal{S}} \alpha_k f_k\left(E_{train}; \mathcal{S}_k\right) \tag{2}$$

where F is the optimal combination model, f_k is k^{th} weak classifier trained with channel set precondition $\mathcal{S}_k$, E_{train} is the training dataset, and α_k is combination parameter.

Frequency Band Selection In general, spectra is not a discrete variable. We simplify spectra as a closed interval, denoted as G, where the elements are all integer points, e.g., G is [8, 30] Hz. We split G into various sub-bands B and $\mathcal{D}$ denotes a universal set composed of all possible sub-bands. The splitting criteria must satisfy:

- Cover: $\cup_{B \in \mathcal{D}} B = G$
- Length: $\forall B = [l, h] \in \mathcal{D}$, $L_{min} \leq h - l \leq L_{max}$, where L_{min} and L_{max} are two constants to determine the length of B.
- Overlap: $\forall B_{min} = [l, l+1] \subset G, \exists B_1, B_2 \in \mathcal{D}, B_{min} \subseteq B_1 \cap B_2$
- Equal: $\forall B_{min} = [l, l+1] \subset G, |\{B : B_{min} \subset B, B \in \mathcal{D}\}| = C$, where C is a constant

These criteria make it possible that the set $\mathcal{D}$ will not underrepresent the original continuous interval. Accordingly, we apply a sliding window strategy to produce all the sub-bands. Four variables involve in the sliding window strategy, they are: the start offset L, the step length S, the sliding window width W and the terminal offset T. In each iteration, the four parameters (L_i, S_i, W_i, T_i) are firstly initialized, where $i = 1 \ldots I$, then the sliding process $Slide(L, S, W, F)$ starts: a band window with width W_i slides from the start point L_i (left edge) with a step length S_i until it reaches the terminal point T_i (right edge) and output sliding windows as sub-bands:

$$BandSet_i = Slide(L_i, S_i, W_i, T_i) \tag{3}$$

By varying the four parameters (L, S, W, T), one universal band set $\mathcal{D}$ including various sub-bands with different start points, widths and terminuses is produced finally:

$$\mathcal{D} = \bigcup_{i=1}^{I} BandSet_i \tag{4}$$

To select the optimal frequency band, we try to seek an optimal band set $\mathcal{B}(\mathcal{B} \subset \mathcal{D})$ which produces an optimal combination classifier F on the training data:

$$F(E_{train}; \mathcal{B}) = \sum_{\mathcal{B}_k \in \mathcal{B}} \alpha_k f_k\left(E_{train}; \mathcal{B}_k\right) \tag{5}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2015.2466079, IEEE Transactions on Neural Systems and Rehabilitation Engineering

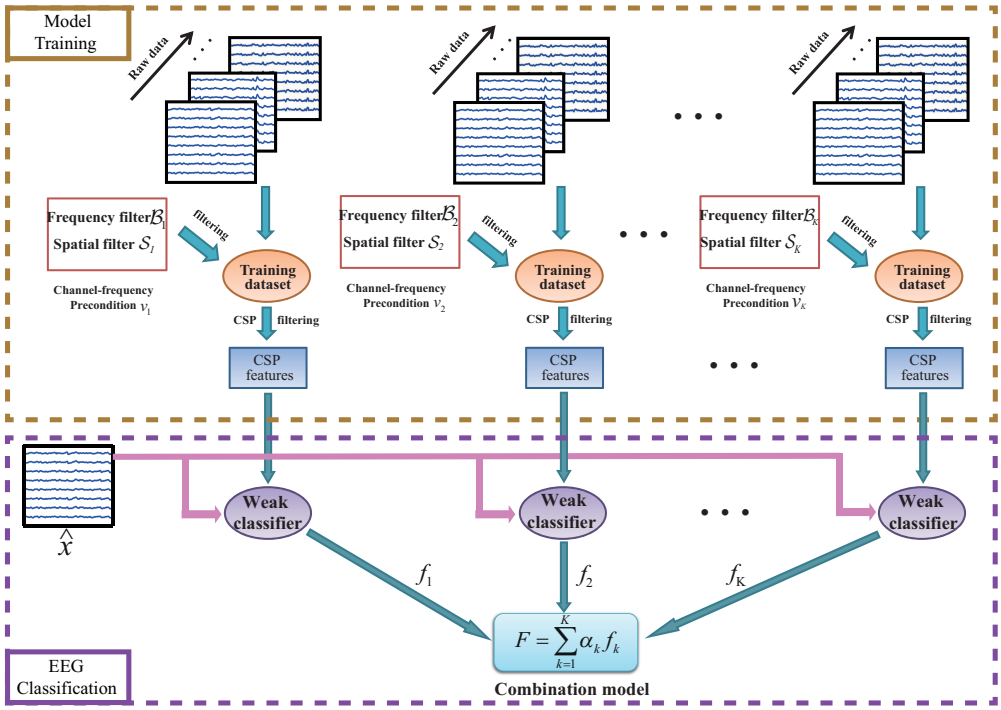

Fig. 1. Architecture of Common Spatial-Spectral Boosting Pattern (CSSBP). Raw EEG data is firstly spatial filtered and bandpass filtered under multiple spatial-spectral preconditions. Afterwards, the CSP algorithm is applied to extract features of the EEG training dataset, by which weak classifiers $\{f_k\}_{k=1}^{K}$ are trained and combined to a weighted commination model F. Finally, a new test sample $\hat{x}$ is classified by this commination model.

where we denote f_k as the k^{th} weak classifier trained by $\mathcal{B}_k$.

Combination To combine the above two models of channel selection and frequency selection, a two-tuple $v_k = (\mathcal{S}_k, \mathcal{B}_k)$ is used to represent a spatial-spectral precondition, and $\mathcal{V}$ is treated as a universal set including all these spatial-spectral preconditions. Finally, the combination function F can be constructed as:

$$F(E_{train}; \mathcal{V}) = \sum_{v_k \in \mathcal{V}} \alpha_k f_k \left(E_{train}; v_k\right) \tag{6}$$

B. Learning Algorithm

In order to effectively solve the problems mentioned above, in this study, we propose an adaptive boosting algorithm to learn the model. This algorithm consists of two main steps of training step and greedy optimization step. We detail the two steps in the following part of this section.

Training step. In this step, weak classifiers are produced under different preconditions. For each spatial-spectral precondition $v_k \in \mathcal{V}$, the training dataset E_{train} are filtered under v_k. Afterwards, the CSP features are obtained by the filtered training dataset E_{train}, by which a weak classifier $f_k(E_{train}; \gamma(v_k))$ is trained, where γ is the model parameter determined by v_k and E_{train}. During this process, a one-to-one relationship between precondition v_k and its related learner f_k is established. Thus we formulate classification error defined in Equation 1 as:

$$\{\alpha, v\}_0^K = \min_{\{\alpha,v\}_0^K} \sum_{n=1}^{N} L(y_n, \sum_{k=0}^{K} \alpha_k f_k(x_n; \gamma(v_k))) \tag{7}$$

where K is the number of weak learners and L is the loss function.

Greedy Optimization Step To solve Equation 7, we use a greedy approach [21], [22], detailed as follows:

$$\begin{aligned} F(E_{train}, \gamma, \{\alpha, v\}_0^K) = &\sum_{k=0}^{K-1} \alpha_k f_k(E_{train}; \gamma(v_k)) \\ &+ \alpha_K f_K(E_{train}; \gamma(v_K)) \end{aligned} \tag{8}$$

Equation 8 can be transformed as a simple recursion formula: $F_k(E_{train}) = F_{k-1}(E_{train}) + \alpha_k f_k(E_{train}; \gamma(v_k))$. To determine the two parameters f_k and α_k, we suppose $F_{k-1}(E_{train})$ is known so we get:

$$\begin{aligned} F_k(E_{train}) &= F_{k-1}(E_{train}) \\ &+ \arg\min_f \sum_{n=1}^{N} L\left(y_n, [F_{k-1}(x_n) + \alpha_k f_k\left(x_n; \gamma(v_k)\right)]\right) \end{aligned} \tag{9}$$

Here we use a steepest gradient descent [22] to solve the problem in Equation 9. Given the pseudo-residuals:

$$\begin{aligned} r_{\pi(n)k} &= -\nabla_F L(y_{\pi(n)}, F(x_{\pi(n)})) \\ &= -\left[\frac{\partial L(y_{\pi(n)}, F(x_{\pi(n)}))}{\partial F(x_{\pi(n)})}\right]_{F(x_{\pi(n)})=F_{k-1}(x_{\pi(n)})} \end{aligned} \tag{10}$$

where $\{\pi(n)\}_{n=1}^{\hat{N}}$ is the first $\hat{N}$ members of a random permutation of $\{n\}_{n=1}^{N}$. Then, a new set $\left\{(x_{\pi(n)}, r_{\pi(n)k})\right\}_{n=1}^{\hat{N}}$,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2015.2466079, IEEE Transactions on Neural Systems and Rehabilitation Engineering

which implies a stochastically-partly best descent step direction, is generated and utilized to learn the parameter $\gamma(v_k)$:

$$\gamma_k = arg\min_{\gamma,\rho} \sum_{n=1}^{\hat{N}} \left[r_{\pi(n)k} - \rho f(x_{\pi(n)}; \gamma_k(v_k)) \right] \quad (11)$$

In the training step, we establish a one-to-one relationship between γ_k and v_k. In this case, v_k is determined whenever γ_k is obtained. We would like to remark that we train each weak classifier f_k under a random subset $\{\pi(n)\}_{n=1}^{\hat{N}}$, instead of the whole training data $\{n\}_{n=1}^{N}$. We apply this stochastic gradient [21] to incorporate randomness in the stagewise iteration for improving performances. Different from the original stochastic gradient which uses a completely random strategy, in our study we use a "*Resample*" heuristic for generating stochastic sequences. During the iteration process, we maintain a self-adjusted training data pool $\mathcal{P}$ at background, as detailed in Algorithm 1. We adjust the training data pool by adding copies of the incorrect-classified samples, and the number of copies is determined by the local classification error. This strategy not only conjoins randomness brought by stochastic gradient, but also increases the probability that incorrect classified samples get selected but decreases the probability that correct classified ones being chosen. The distribution of the samples in $\mathcal{P}$ will not change too much until the combination model F has got a strong description ability about the training data. Algorithm 1 details the whole process of the resample heuristic.

Algorithm 1 Resample Heuristic Algorithm for Stochastic Subset Selection

Initialize the training data pool $\mathcal{P}_0 = E_{train} = \{x_n, y_n\}_{n=1}^{N}$;
for $k = 1$ to K **do**
  Generate a random permutation $\{\pi(n)\}_{n=1}^{|\mathcal{P}_{k-1}|} = randperm(\{n\}_{n=1}^{|\mathcal{P}_{k-1}|})$;
  Select the first $\hat{N}$ elements $\{\pi(n)\}_{n=1}^{\hat{N}}$ as $\{x_{\pi(n)}, y_{\pi(n)}\}_{n=1}^{\hat{N}}$ from $\mathcal{P}_0$;
  Use these $\{\pi(n)\}_{n=1}^{\hat{N}}$ elements to optimize new learner f_k and its related parameters in Algorithm 2;
  Use current local optimal classifier F_k to split the original training set $E_{train} = \{x_n, y_n\}_{n=1}^{N}$ into two parts $T_{true} = \{x_n, y_n\}_{n:y_n = F_{k(x_n)}}$ and $T_{false} = \{x_n, y_n\}_{n:y_n \neq F_{k(x_n)}}$;
  Re-adjust the training data pool:
  for each $(x_n, y_n) \in T_{false}$ **do**
    Select out all $(x_n, y_n) \in \mathcal{P}_{k-1}$ as $\{x_{n(m)}, y_{n(m)}\}_{m=1}^{M}$;
    Copy $\{x_{n(m)}, y_{n(m)}\}_{m=1}^{M}$ with $d(d \geq 1)$ times so that we get total $(d+1)M$ duplicated samples;
    Return these $(d+1)M$ samples into $\mathcal{P}_{k-1}$ and we get a new adjusted pool $\mathcal{P}_k$;
  end for
end for

With $\gamma_k(v_k)$, the combination coefficient α_k is obtained:

$$\alpha_k = \arg\min_{\alpha} \sum_{n=1}^{N} L(y_n, F_{k-1}(x_n) + \alpha f_k(x_n; \gamma_k(v_k))) \quad (12)$$

Algorithm 2 details the whole process of the proposed method.

C. Parameter Estimation

In this section, we discuss how to estimate some important parameters. The iteration time K is determined by using the early stopping strategy [23]. For $\hat{N}$ which is the size of the training set used in model training, we can find that more randomness will be incorporated if the ratio $\hat{N}/N$ is decreased. However, if we increase this ratio, more samples are brought into the model to train a more robust local weak learner f_k. In this study, we set $\hat{N}/N \approx 0.7$ and a relatively satisfactory performance is achieved with short training time. For d, the copies of incorrect-classified samples when adjusting $\mathcal{P}$, it is determined by the the local classification error $e = |T_{false}|/N$:

$$d = max\left(1, \lfloor \frac{1-e}{e+\epsilon} \rfloor\right) \quad (13)$$

where ϵ is an accommodation coefficient. Note that e is always smaller than 0.5 and will decrease during the iteration so that a larger penalty will be given on samples that are incorrect classified by stronger classifiers. In terms of the loss function L, we simply choose the squared error loss for convenience.

Algorithm 2 The Framework of Common Spatial-Spectral Boosting Pattern (CSSBP) Algorithm

Input: $\{x_n, y_n\}_{n=1}^{N}$: EEG training set; $L(y, x)$: The loss function; K: The capacity of the optimal precondition set (number of weak learners); $\mathcal{V}$: A universal set including all possible preconditions;
Output: F: The optimal combination classifier; $\{f_k\}_{k=1}^{K}$: The weak learners; $\{\alpha_k\}_{k=1}^{K}$: The weights of weak learners; $\{v_k\}_{k=1}^{K}$: The preconditions under which weak learners are trained.

Feed $\{x_n, y_n\}_{n=1}^{N}$ and $\mathcal{V}$ into a classifier using CSP to extract features to produce a family of weak learners $\mathcal{F}$, so that a one-to-one mapping is established: $\mathcal{F} \leftrightarrow \mathcal{V}$;
Initialize $\mathcal{P}_0$, $F_0(E_{train}) = argmin_{\alpha} \sum_{n=1}^{N} L(y_n, \alpha)$;
for $k = 1$ to K **do**
  Optimalize $f_k(E_{train}; \gamma(v_k))$ described in Equation 11;
  Optimalize α_k as described in Equation 12;
  Update $\mathcal{P}_k$ as in Algorithm 1 and $F_k(E_{train}) = F_{k-1}(E_{train}) + \alpha_k f_k(E_{train}; \gamma(v_k))$;
end for
for each $f_k(E_{train}; \gamma(v_k))$, use the mapping $\mathcal{F} \leftrightarrow \mathcal{V}$ to find its corresponded precondition v_k;
return $F, \{f_k\}_{k=1}^{K}, \{\alpha_k\}_{k=1}^{K}, \{v_k\}_{k=1}^{K}$;

III. Experimental Configuration

A. Data Acquisition

In order to evaluate the effectiveness and robustness of the proposed algorithm, three different motor imagery-based datasets, which were assembled from two famous BCI competitions and one self-collected stroke patients' EEG, were used for simulation. Details of these datasets were described as:

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2015.2466079, IEEE Transactions on Neural Systems and Rehabilitation Engineering

1) Dataset I: Dataset I was dataset IVa from the famous BCI competition III [24] where five subjects ('aa', 'al', 'av', 'aw' and 'ay') participated in a motor imagery-based experiment. All these subjects had to imagine right hand or right foot movement. EEG was recorded using 118 electrodes, band-pass filtered between 0.05 and 200 Hz, and down-sampled to 100 Hz. A time segment from 500 to 2500 ms was extracted for analysis. All the 280 trials for each subject were divided into a training dataset and test dataset. In detail, the training dataset for each subject was 168, 224, 84, 56 and 28, respectively.

2) Dataset II: Dataset II was dataset IIa from the famous BCI competition IV [25] where four motor imagery based tasks of left-hand, right-hand, foot, and tongue were involved. 9 healthy subjects (labeled S1-S9 respectively) participated in this experiment and had to complete two sessions conducted on different days. In each session, there were 6 runs and each run was composed of 12 trials for each task. Totally, one session was composed of 288 trials. EEG data was collected from 22 electrodes, bandpass-filtered between 0.5Hz and 100Hz, and sampled at 250Hz.

3) Dataset III: This dataset was collected from five unilaterally paralyzed stroke patients diseased in two months who performed motor imagination of their disabled (left or right) arm in a BCI combined with Functional Electrical Stimulation (BCI-FES) rehabilitation training system for 24 times over two months (three times per week). EEG data was recorded by a 16-channel (FC3, FCZ, FC4, C1-C6, CZ, CP3, CPZ, CP4, P3, PZ and P4) g.USBamp amplifier at a sampling rate of 256 Hz. Patients had to complete basic motor imagery related tasks for five sessions and each session comprised forty trials and lasted for 240 s. We extracted a time segment starting from 0.5s to 4.5s after the visual cue for analysis. All the trials in the training model course were divided into a training set with 120 trials and a testing set with 80 trials.

B. Data Preprocessing

In this step, EEG trial was bandpass filtered in a specific frequency band involving in motor imagery. For healthy people, the most contributed frequency bands were α rhythm (8-13 Hz) and β rhythm (14-30 Hz) [11]. Thus Dataset I-II were bandpass filtered in the frequency range of 8-30 Hz. However, the spectral characteristics of stroke patients were not available beforehand [26], [27]. Thus raw data in Dataset III was filtered in a general band ranged from 5 to 40 Hz.

C. Feature Extraction

For comparison, seven state-of-the-art algorithms, which were Power Spectral Density (PSD) [28], Phase Synchrony Rate (SR) [7], the original CSP, regularized CSP (RCSP) [9], the sub-band CSP (SBCSP) [29], CSSP and CSSSP, were utilized for feature extraction. For PSD, we applied a fast Fourier transform to compute power spectral density values. For RCSP, Weighted Tikhonov Regularization (α) was chosen, as suggested by the study [9]. It has to be pointed out that all the model parameters (α for RCSP, τ for CSSP and C for CSSSP) were chosen on the training set using a 5-fold cross validation procedure. Note that for the 4-class motor imagery classification task in Dataset II, the CSP-based methods employed the one-versus-rest (OVR) strategy. After feature extraction, a Fisher score strategy [30] was used to select relevant features, as more features cannot improve the training accuracy [30].

D. Classification and Validation

A linear support vector machine (SVM) [31], which obtains top-level performance in many applications, was conducted for classification. A 5-fold cross-validation was used to choose suitable SVM parameters for prediction.

IV. Results

In this section, we presented the experimental results on the three aforementioned datasets. To begin with, we applied the proposed method on the first two datasets, whose motor imagery related spatial and spectral characteristics were known, and evaluated its performance in comparison with that of seven popular algorithms. Finally, the proposed method was applied on the last dataset, where the discriminative spatial and spectral properties were not specifically identified.

A. Results on Dataset IVa from BCI competition III

We firstly assessed the effectiveness of our proposed method on the widely-used benchmark Dataset IVa from BCI competition III. The results were mainly given in two aspects of illustration of the classification accuracy and visualization of the obtained spatial-spectral filters.

(1) Classification accuracy. Fig. 2 shows the classification results of the test dataets between CSSBP and the other competing methods. Left part in Fig. 2 presents the classification accuracies of all the algorithms while right part gives the details of the box plots. The feature dimensionality of each method varied and was determined according to the training performance, since the training performance can not be improved by more features [30]. Results showed that CSSBP outperformed other competing algorithms for all the subjects, as it reached both the highest median and mean accuracy. e.g., CSSBP outperformed CSP by about 7% in mean classification accuracy and by almost 6% in median classification accuracy.

(2) Spatial and spectral filters. Besides the superior classification performance, we tried to observe the underlying cortical activity patterns by visualizing the spatial patterns (spatial weights) and spectral patterns (spectral weights) from limited training dataset.

(1) *Spatial-spectral weights.* For computing the spatial weight for each channel, we used a quantitative vector $L = \sum_{\mathcal{S}_i \in \mathcal{S}} \alpha_i \mathcal{S}_i$, where $\mathcal{S}$ is the channel sets and α is their weights. The spectral weight were calculated similarly and projected onto the frequency bands [8, 30]. Finally, we sorted the importance of contributed channels and active frequency bands by weight values.

(2) *Peak amplitude of CSSBP filtered EEG.* Besides the spatial and spectral patterns, we also studied the temporal patterns of the filtered EEG. We first filtered the original EEG under all the spatial-spectral preconditions where $v_k \in \mathcal{V}$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2015.2466079, IEEE Transactions on Neural Systems and Rehabilitation Engineering

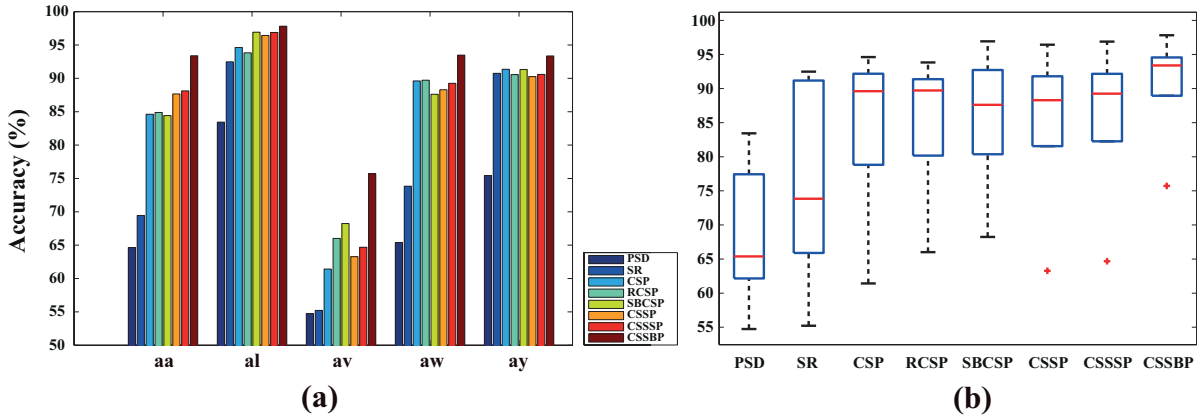

Fig. 2. (a) Experimental results on the test accuracies obtained for each subject in BCI competition III Dataset IVa for all the competing algorithms PSD, SR, CSP, RCSP, SBCSP, CSSP, CSSSP and our proposed algorithm CSSBP. (b) Boxplots of all the eight algorithms.

represents the k^{th} one. The spatial filters of CSP was then obtained on the EEG datasets, yielding relevant components for right hand and right foot imagination movement. We selected the first two spatial components and projected them to the sensor space. Afterwards, the peak amplitude of the CSP filtered EEG signals E_k for each channel $\mathcal{C}_i \in \mathcal{C}$ was computed and denoted by $P_{k\mathcal{C}_i}$. We averaged the $P_{k\mathcal{C}_i}$ over all conditions $v_k \in \mathcal{V}$ evaluated by $\hat{P_{\mathcal{C}_i}} = \frac{1}{|\mathcal{V}|}\sum_{v_k \in \mathcal{V}} \alpha_k P_{k\mathcal{C}_i}$ where α_k denoted the corresponding weight for the k^{th} condition, and visualized them through the 2-D topoplot map.

Fig. 3 gives an illustration of spatial, spectral and temporal information of EEG recorded from the five subjects. Left part of Fig. 3 details the spatial-spectral weights obtained by CSSBP while right part presents 2-D topoplot maps of peak amplitudes of CSSBP filtered EEG in each channel. For comparison, Fig. 4 shows the spatial filters constructed by CSP. From these pictures, we observed that for almost all the subjects, the most contributed channels related to motor imagery obtained by CSSBP were concentrated at the motor cortex areas, as also expected from the study [9], [11]. In detail, channels contributed to right foot imagination were focused on central cortical area while important channels for right hand imagination were centered on left cortical area. The same phenomenon was also observed in the pictures given by CSP. As for spectral characteristics, α rhythm and β rhythm contributed more to motor imagery. With a close view to these spectral filters, there was a slight difference in the weights between higher band and lower band, which suggested that there was a diversity in spectral patterns between subjects.

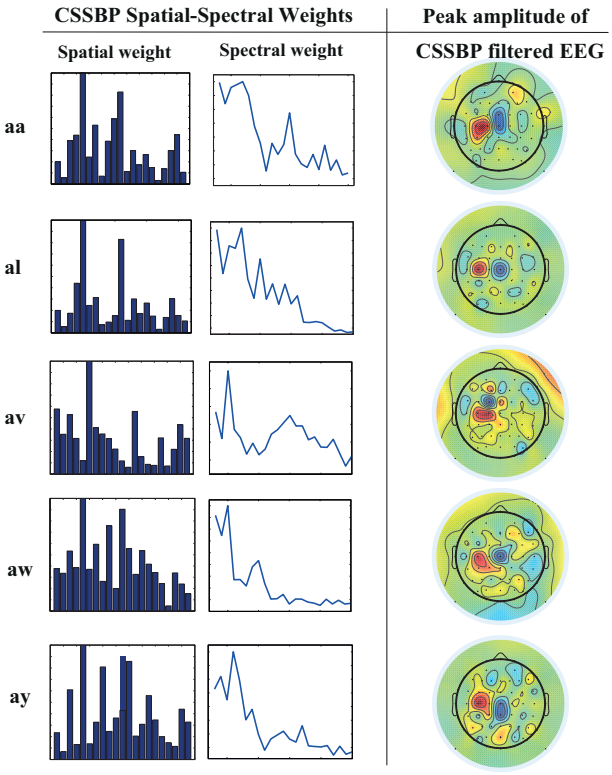

Fig. 3. Spatial, spectral and temporal information of EEG obtained by CSSBP for each subject in BCI competition III Dataset IVa. Left part: the spatial-spectral weights obtained by CSSBP. Note that x axis in 'spatial weight' histogram represents the 21 chosen channels over the motor cortex (left to right : FC3, FC5, FC1, C5, C3, C1, CP5, CP3, CP1, FCZ, CZ, CPZ, CP2, CP4, CP6, C2, C4, C6, FC2, FC4 and FC6) while y-axis describes the normalized weights. x axis in 'spectral weight' subfigure shows the frequency band [8, 30] Hz while y-axis displays the normalized weights. Right part: 2-D topoplot maps of peak amplitude of CSSBP filtered EEG in each channel.

B. Results on Dataset IIa from BCI competition IV

As is known to all, EEG classification performance heavily depends on the selection of a time interval. Thus in this dataset, based on Ang *et al.*'s work [32], a time segment from 0.5s to 2.5s after onset of the visual cue was chosen for analysis. The BCI Competition IV with dataset IIa aimed at evaluating the classification performance of all the methods based on the session-to-session transfer rate. Here we selected session 1 and session 2 for evaluation by using continuous classification output for each sample [25]. We used a sliding window strategy to output the class labels (Window size: 2s; overlapping: 240 points). We used a metric of Cohen's Kappa value that measured the agreement between two estimators [33], as also done in BCI Competition IV.

In Fig. 5, we showed the best performance in terms of Cohen's kappa values of the all the methods. Note that the dimensionality of feature vector for each algorithm was de-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2015.2466079, IEEE Transactions on Neural Systems and Rehabilitation Engineering

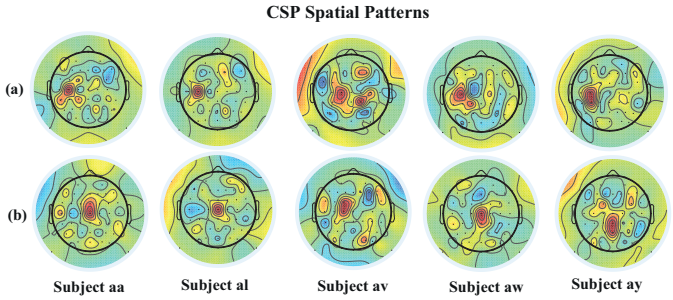

Fig. 4. Spatial weights for the two most discriminative filters constructed by CSP for all the subjects in BCI competition III Dataset IVa. (a) Spatial filters for right hand imagination; (b) Spatial filters for right foot imagination.

termined according to the training performance. The proposed method outperformed the competing methods, and showed the highest average Kappa value. The statistical analysis showed significant difference between CSP and the proposed method in a Mann-Whitney U test. For all the cases, the proposed method outperformed CSP in a 95% confidence level.

With a closer look at the results, we found that the classification accuracies of the subjects with poor CSP accuracies were significantly improved by CSSBP, e.g., S5 and S6. That might be due to the intrinsic characteristics of subject-specific EEG: the subject S3 and S8 with high CSP accuracy might be have clean and noiseless data, whereas the accuracy of S5 was close to chance level, thus there was more room for improvement for subject S5 compared to S3 and S8.

C. Results on Dataset collected from stroke patients

The above experimental results evidenced the superior classification performance of CSSBP when applied on healthy persons with available spatial and spectral characteristics. In this section, CSSBP was utilized to explore the unavailable motor imagery patterns of stroke patients.

(1) Classification accuracy. In this dataset, by varying feature dimensionality, we calculated the classification accuracy for each day. Afterwards, we averaged all the accuracies in all the eight weeks to represent the mean accuracy of the corresponding method in terms of each patient. In Fig. 6, we showed the the averaged classification accuracies for all the methods when varying feature dimensionality.

Obviously CSSBP yielded relatively higher classification accuracies compared with the seven competing approaches. As feature dimensionality increased, the mean classification accuracies changed a lot. Moreover, the classification performance was greatly improved in terms of the patients with poor CSP performances. We used a Mann-Whitney U test to test the statistical significance between CSSBP and the other methods. The results showed that CSSBP achieved significant improvement in classification performance at a 95% confidence level. We can observe that after two months' rehabilitation, the classification accuracies by CSSBP for almost all the subjects could even exceed 70%.

(2) Spatial and spectral filters. Toward understanding the merits of CSSBP, the spatial, spectral and temporal characteristics of all the patients on day 30 were learned and

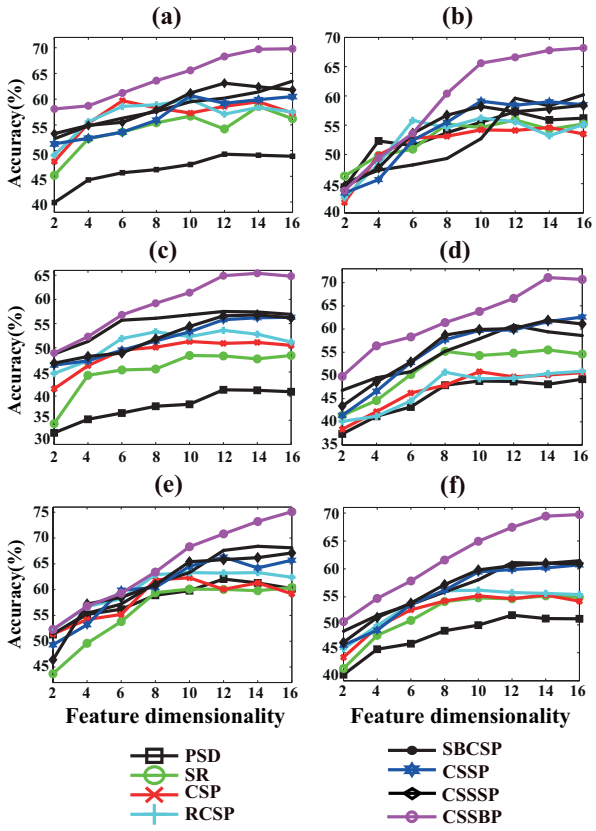

Fig. 6. The mean accuracies obtained for each subject in Dataset III under different feature dimensionalities. (a)-(e) Patient 1-5; (f) Group mean.

visualized in Fig. 7. Left part of Fig. 7 details the spatial-spectral weights obtained by CSSBP while right part presents 2-D topoplot maps of peak amplitudes of CSSBP filtered EEG in each channel. For comparison, Fig. 8 shows the spatial filters obtained by CSP. From these pictures, we can see that CSP could not extract discriminative spatial filters, which to the contrary appeared as messy. Differently, CSSBP filters were more discriminative and physiologically relevant. To explore the spatial characteristics, for patients with lesion in right side (except Patient 2), the important channels contributed to right movement imagination were located at C3; as to left movement imagination, the significant channels were concentrated in larger areas (C4, FC4 and P4), which was also reported in some other studies [34], [35]. In terms of spectral characteristics, we can conclude that the frequency information was subject-specific, e.g., higher bands (28-35 Hz) contributed more to classification for Patient 2 while active frequencies for Patient 1 scattered at relatively wide-ranged bands (15-30 Hz). Similar reports can be found in [36].

(3) Gradual changes of EEG patterns over time. Due to brain reorganization in motor areas of stroke patients, EEG patterns may have gradually changed during rehabilitation [11], [37]. In order to observe the gradual changes of EEG patterns in spatial and spectral domains over time, we chose the raw EEG of seven days, day 1, 10, 20, 30, 40, 50 and 60, to represent the different phases during rehabilitation, and CSSBP was utilized to extract the most contributed channels

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2015.2466079, IEEE Transactions on Neural Systems and Rehabilitation Engineering

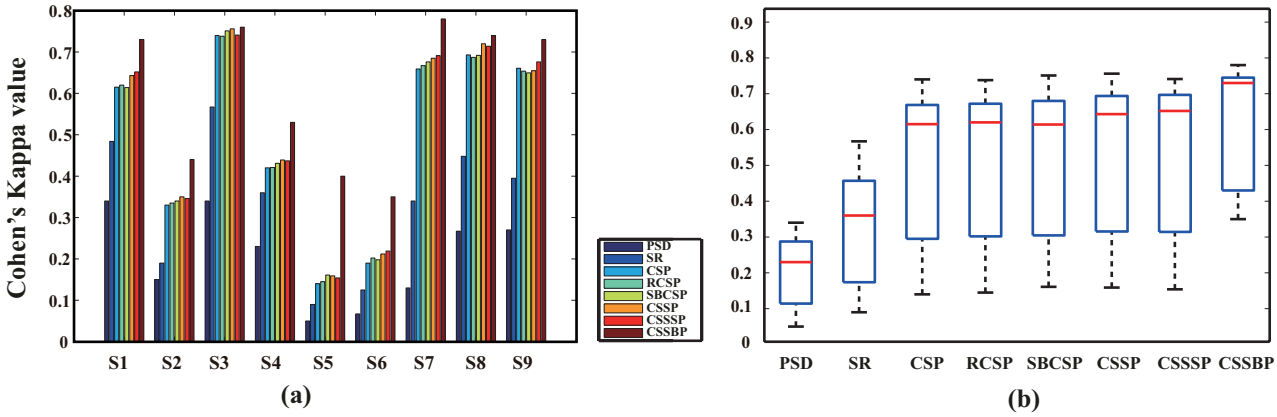

Fig. 5. Classification performances of the seven competing methods and the proposed method. (a) Experimental results on the test accuracies obtained for each subject in BCI Competition IV dataset IIa. (b) Boxplots of all the eight algorithms.

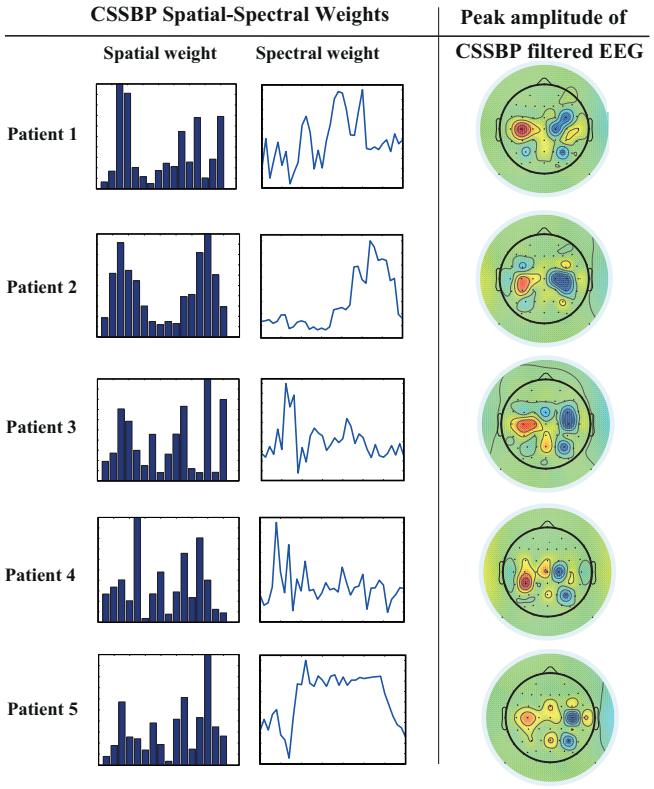

Fig. 7. Spatial, spectral and temporal information of EEG obtained by CSSBP for each patient in Dataset III on day 30. Left part: the spatial-spectral weights obtained by CSSBP. Note that x axis in 'spatial weight' histogram represents all the channels over the motor cortex (left to right : FC3, C5, C3, C1, CP3, P3, FCZ, CZ, CPZ, PZ, P4, CP4, C2, C4, C6 and FC4) while y-axis describes the normalized weights. x axis in 'spectral weight' subfigure shows the frequency band 5-40 Hz while y-axis displays the normalized weights. Right part: 2-D topoplot maps of peak amplitude of CSSBP filtered EEG in each channel.

group and frequency bands of each subject. Fig. 9 illustrates the chosen 7 days' weights of channels and frequency band 5-40 Hz for Patient 2 with lesion in the left side.

Obviously we found that larger regions were activated in the affected hemisphere during rehabilitation. The importance of sensorimotor area on the affected hemisphere increased over time while that of the unaffected hemisphere held, that is,

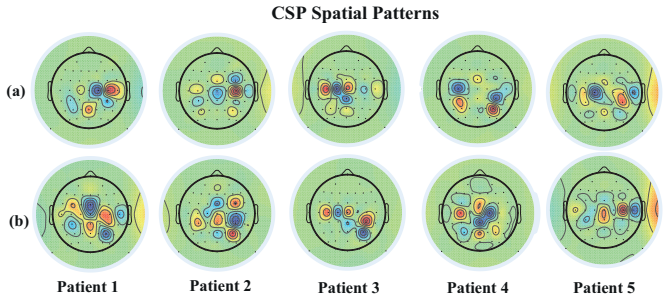

Fig. 8. Spatial weights for the two most discriminative filters obtained by CSP for all the patients in Dataset III on day 30. (a) Spatial filters for left hand motor imagery; (b) Spatial filters for right hand motor imagery.

e.g., at the beginning, central parietal area (CP3) played a compensatory role but the importance decreased over time. Active areas gradually shifted back to central lobes (C3). Similar phenomenon was also reported in some other studies [11], [38]. Considering the spectral characteristics, a higher frequency band 25-38 Hz contributed largely to motor imagery, but the importance decentralized and distributed to lower bands partly over time. This dynamic band accentuation implied that active rhythms might be changing during rehabilitation, as also reported in [27], [38].

V. Discussion

This study presented a spatial-spectral boosting algorithm, namely CSSBP, in which the usually predetermined spatial-spectral configurations were divided into variable preconditions, and then introduced a new heuristic supervisor of stochastic gradient boost strategy for training weak classifiers under these preconditions.

Extensive experiment comparisons were performed on two datasets containing data acquired from normal persons and one dataset related to unilaterally paralyzed stroke patients. Experimental results showed that our algorithm yielded relatively higher classification accuracies compared with seven state-of-the-art approaches. In particular, with a close observation at the classification results in the three datasets, we found one interesting phenomenon: for the subjects with poor CSP

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2015.2466079, IEEE Transactions on Neural Systems and Rehabilitation Engineering

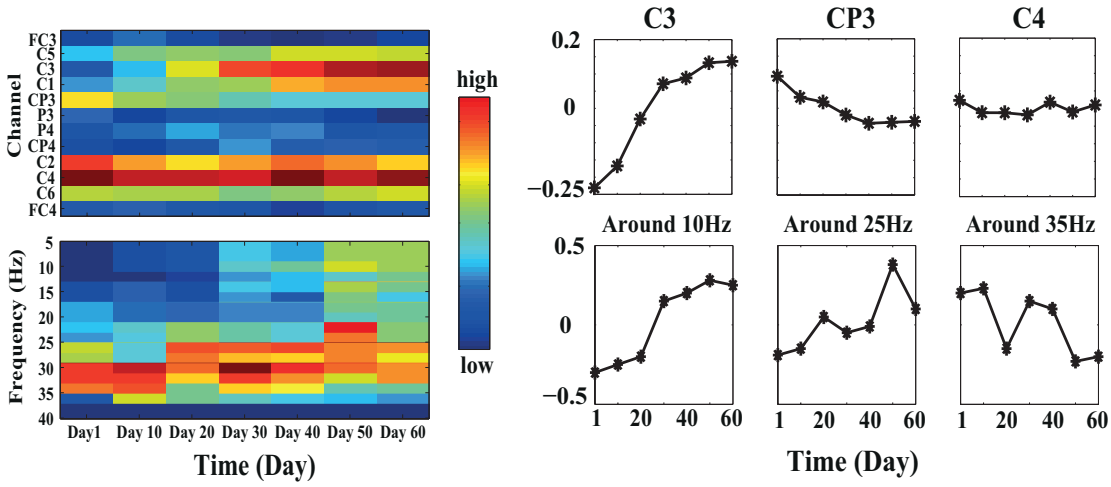

Fig. 9. The weights of the 16 channels and frequency band 5-40 Hz on the chosen 7 days for Patient 2 in Dataset III. All the values have been normalized for comparison. Left part: the weights of channels and frequency bands on the chosen 7 days, where red rectangle represents a larger value and blue one is lower. Right part: the changes of weights subtracting their average values of three chosen channels C3, CP3, C4 and frequencies 10Hz, 25Hz, 35Hz on the chosen 7 days. Note that: (1) weight differences of channel C3 and lower frequencies 10Hz and 25Hz gradually increase while that of channel C4 holds; (2) weight differences of channel CP3 and higher frequency 35Hz present a decreasing tendency.

performances (i.e., CSP error rate more than 30% like av in Dataset I, and almost all the subjects in Dataset III), the proposed CSSBP algorithm significantly outperformed the CSP algorithm. Moreover, compared to Dataset I and II, CSSBP achieved salient improvement in classification performance in Dataset III.

To explore the reasons, it could be the fact that the spatial and spectral information is not available for stroke patients. For stroke patents under recovery, there is a higher chance that the most contributed channels group and frequency bands will change. Moreover, EEG from some unknown channels is contaminated with more noisy and nonstationary signals [39]. Therefore, CSSBP is proposed to seek the operational subject-specific frequency band and channels group for stroke patients. As an adaptive boosting algorithm, CSSBP does not use the spatial-spectral configurations directly but divides them into multi-subsets, and finally calculates each subset's contribution (weight) to classification. This makes sense, since the channels mainly involved in performing motor imagery will get larger weights and finally be selected while the other futile channels with smaller weights will be eliminated after channel selection. Deeply, we have calculated the variance of the channels' weights for each day, and found that the variances maintained at a relatively larger level. We conclude that a larger variance, which indicates a polarized distribution of channel weights, may lead a better performance for channel selection. The best performance achieved by CSSBP in the classification results evidence that channel selection and frequency selection extract extra effective information in strokes' EEG, and they complement each other, both contributing to the classification ability. All these results show that CSSBP could be more successful and robust in capturing the spatial-spectral filters.

Another important issue is the exploration of the neurophysiologic mechanism of injured brain's motor functional reconstruction during rehabilitation. Stroke is associated with deficits in a number of cognitive processes and executive functions, and the neurological plasticity develops with the progression of stroke, which may lead to a changing of the active motor cortex regions and frequency bands. We conducted further investigations on exploring EEG characteristics by visualizing the spatial-spectral filters obtained our proposed method. The results showed that during motor recovery motor imagery EEG patterns of stroke patient were changing. By tracking these gradual changes of the obtained spatial-spectral filters during rehabilitation, we tried to reveal the compensatory mechanism about cortex reorganization. From the experimental results, we concluded that the contribution of the injured sensorimotor area and some other areas such as frontal premotor area and parietal area was changing over time. In detail, parietal lobes and frontal premotor lobes were activated and contributed to motor imagery. However, during rehabilitation, those cortical areas gave back their responsibility to central lobes over time. Similar phenomena were also reported in some other studies [34], [37].

VI. Conclusion

In this paper, a spatial-spectral boosting algorithm was proposed for EEG classification, which divided the channel and frequency configuration into preconditions by using a sliding window strategy and then trained the weak learners through a stochastic gradient boost under these preconditions. The most contributed channel groups and frequency bands related to motor imagery were selected and could be treated as effective instructions for CSP.

We evaluated the effectiveness of our proposed algorithm when compared with the original CSP and some other CSP-based algorithms on three different datasets recorded from diverse populations including the healthy people and stroke patients. The results demonstrated its superior classification performance. Moreover, we also observed that the discriminative frequency bands among subjects were subject-specific, leading to a time-consuming tuning process in order to achieve

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2015.2466079, IEEE Transactions on Neural Systems and Rehabilitation Engineering

the optimal performance. Our approach addressed this problem and can easily achieve satisfactory results.

Besides the excellent classification performance, another merit of CSSBP was that it could learn neurophysiological relevant spatial and spectral filters, which allowed us to explore the neurophysiologic knowledge of brain activity in some special populations when lacking prior knowledge. e.g., when applied on analyzing EEG recorded from stroke patients, the most contributed channels and active frequency band obtained by CSSBP were more significant than the messy CSP filters. This makes sense, since it can better explore the neurophysiologic mechanism of underlying brain's activities.

Finally, we would like to remark that CSSBP is—although very well suited to EEG analysis—a general framework that can be easily used in the other BCI-based paradigms where spatial-spectral filters need to be constructed. e.g., in this study, we can see that CSP is only one special method embedding in our framework for EEG analysis. For the other applications, one can just replace CSP with the other corresponding methods in Algorithm 2 without any change of our framework.

ACKNOWLEDGMENT

The work was supported in part by the National Basic Research Program of China (Grant No. 2015CB856004) and in part by the National Natural Science Foundation of China (Grant Nos. 91120305, 91420302, 61202155). This work is also sponsored by China Scholarship Council (CSC).

REFERENCES

[1] S. Gao, Y. Wang, X. Gao, and B. Hong, "Visual and auditory brain-computer interfaces," 2014.

[2] W. Wu, Z. Chen, X. Gao, Y. Li, E. Brown, and S. Gao, "Probabilistic common spatial patterns for multichannel EEG analysis," 2014.

[3] K. K. Ang, K. S. G. Chua, K. S. Phua, C. Wang, Z. Y. Chin, C. W. K. Kuah, W. Low, and C. Guan, "A randomized controlled trial of EEG-based motor imagery Brain-Computer Interface robotic rehabilitation for stroke," *Clinical EEG and neuroscience*, p. 1550059414522229, 2014.

[4] K. K. Ang, C. Guan, K. S. G. Chua, B. T. Ang, C. W. K. Kuah, C. Wang, K. S. Phua, Z. Y. Chin, and H. Zhang, "A large clinical study on the ability of stroke patients to use an EEG-based motor imagery brain-computer interface," *Clinical EEG and Neuroscience*, vol. 42, no. 4, pp. 253–258, 2011.

[5] F. Meng, K.-y. Tong, S.-t. Chan, W.-w. Wong, K.-h. Lui, K.-w. Tang, X. Gao, and S. Gao, "BCI-FES training system design and implementation for rehabilitation of stroke patients," in *IJCNN 2008*. IEEE, 2008, pp. 4103–4106.

[6] E. Buch, C. Weber, L. G. Cohen, C. Braun, M. A. Dimyan, T. Ard, J. Mellinger, A. Caria, S. Soekadar, A. Fourkas *et al.*, "Think to move: a neuromagnetic brain-computer interface (BCI) system for chronic stroke," *Stroke*, vol. 39, no. 3, pp. 910–917, 2008.

[7] L. Song, E. Gysels, and E. Gordon, "Phase synchrony rate for the recognition of motor imagery in brain-computer interface," 2006.

[8] H. Ramoser, J. Muller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, pp. 441–446, 2000.

[9] F. Lotte and C. Guan, "Regularizing common spatial patterns to improve BCI designs: unified theory and new algorithms," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 2, pp. 355–362, 2011.

[10] G. Blanchard and B. Blankertz, "BCI competition 2003-data set IIa: spatial patterns of self-controlled brain rhythm modulations," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, pp. 1062–1066, 2004.

[11] W.-K. Tam, K.-y. Tong, F. Meng, and S. Gao, "A minimal set of electrodes for motor imagery BCI to control an assistive device in chronic stroke subjects: a multi-session study," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 19, no. 6, pp. 617–627, 2011.

[12] J. Yan, X. Guo, Z. Jin, J. Sun, L. Shen, and S. Tong, "Cognitive alterations in motor imagery process after left hemispheric ischemic stroke," *PloS one*, vol. 7, no. 8, p. e42922, 2012.

[13] J. Matsumoto, T. Fujiwara, O. Takahashi, M. Liu, A. Kimura, and J. Ushiba, "Research modulation of mu rhythm desynchronization during motor imagery by transcranial direct current stimulation," 2010.

[14] Y. Kasashima, T. Fujiwara, Y. Matsushika, T. Tsuji, K. Hase, J. Ushiyama, J. Ushiba, and M. Liu, "Modulation of event-related desynchronization during motor imagery with transcranial direct current stimulation (tdcs) in patients with chronic hemiparetic stroke," *Experimental brain research*, vol. 221, no. 3, pp. 263–268, 2012.

[15] H. Yang, C. Guan, K. K. Ang, K. S. Phua, and C. Wang, "Selection of effective EEG channels in brain computer interfaces based on inconsistencies of classifiers," in *EMBC 2014*. IEEE, 2014, pp. 672–675.

[16] Z. Y. Chin, K. K. Ang, C. Wang, and C. Guan, "Discriminative channel addition and reduction for filter bank common spatial pattern in motor imagery BCI," in *EMBC 2014*. IEEE, 2014, pp. 1310–1313.

[17] S. Lemm, B. Blankertz, G. Curio, and K.-R. Muller, "Spatio-spectral filters for improving the classification of single trial EEG," *IEEE Transactions on Biomedical Engineering*, vol. 52, no. 9, pp. 1541–1548, 2005.

[18] G. Dornhege, B. Blankertz, M. Krauledat, F. Losch, G. Curio, and K.-R. Muller, "Combined optimization of spatial and temporal filters for improving brain-computer interfacing," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 11, pp. 2274–2281, 2006.

[19] W. Wu, X. Gao, B. Hong, and S. Gao, "Classifying single-trial EEG during motor imagery by iterative spatio-spectral patterns learning (ISSPL)," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 6, pp. 1733–1743, 2008.

[20] K. K. Ang, Z. Y. Chin, H. Zhang, and C. Guan, "Mutual information-based selection of optimal spatial–temporal patterns for single-trial EEG-based BCIs," *Pattern Recognition*, vol. 45, no. 6, pp. 2137–2144, 2012.

[21] J. H. Friedman, "Stochastic gradient boosting," *Computational Statistics & Data Analysis*, vol. 38, no. 4, pp. 367–378, 2002.

[22] ——, "Greedy function approximation: a gradient boosting machine," *Annals of Statistics*, pp. 1189–1232, 2001.

[23] T. Zhang and B. Yu, "Boosting with early stopping: convergence and consistency," *Annals of Statistics*, pp. 1538–1579, 2005.

[24] B. Blankertz, K.-R. Muller, D. J. Krusienski *et al.*, "The BCI competition III: Validating alternative approaches to actual BCI problems," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 14, no. 2, pp. 153–159, 2006.

[25] C. Brunner, R. Leeb, G. Müller-Putz, A. Schlögl, and G. Pfurtscheller, "BCI competition 2008–Graz data set A," *Institute for Knowledge Discovery (Laboratory of Brain-Computer Interfaces), Graz University of Technology*, 2008.

[26] D. J. Leamy, J. Kocijan, K. Domijan, J. Duffin, R. A. Roche, S. Commins, R. Collins, and T. E. Ward, "An exploration of EEG features during recovery following stroke–implications for BCI-mediated neurorehabilitation therapy," *Journal of neuroengineering and rehabilitation*, vol. 11, no. 1, p. 9, 2014.

[27] S. Bermudez i Badia, H. Samaha, A. G. Morgade, and P. F. Verschure, "Exploring the synergies of a hybrid BCI-VR neurorehabilitation system," in *Virtual Rehabilitation (ICVR), 2011 International Conference on*. IEEE, 2011, pp. 1–8.

[28] P. Welch, "The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms," *IEEE Transactions on Audio and Electroacoustics*, vol. 15, no. 2, pp. 70–73, 1967.

[29] Q. Novi, C. Guan, T. H. Dat, and P. Xue, "Sub-band common spatial pattern (SBCSP) for brain-computer interface," in *Neural Engineering, 2007. CNE'07. 3rd International IEEE/EMBS Conference on*. IEEE, 2007, pp. 204–207.

[30] C. M. Bishop and N. M. Nasrabadi, *Pattern recognition and machine learning*. springer New York, 2006, vol. 1.

[31] C.-C. Chang and C.-J. Lin, "LIBSVM: a library for support vector machines," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 2, no. 3, p. 27, 2011.

[32] K. K. Ang, Z. Y. Chin, H. Zhang, and C. Guan, "Filter bank common spatial pattern (FBCSP) in brain-computer interface," in *IJCNN 2008*. IEEE, 2008, pp. 2390–2397.

[33] J. Cohen, "A coefficient of agreement for nominal scales," *Educ. Psychol. Meas.*, vol. 20, no. 1, pp. 37–46, 1960.

[34] A. Feydy, R. Carlier, A. Roby-Brami, B. Bussel, F. Cazalis, L. Pierot, Y. Burnod, and M. Maier, "Longitudinal study of motor recovery after

stroke recruitment and focusing of brain activation," *Stroke*, vol. 33, no. 6, pp. 1610–1617, 2002.

[35] C. Calautti, F. Leroy, J.-Y. Guincestre, and J.-C. Baron, "Dynamics of Motor Network Overactivation After Striatocapsular Stroke A longitudinal PET Study Using a Fixed-Performance Paradigm," *Stroke*, vol. 32, no. 11, pp. 2534–2542, 2001.

[36] S. Shahid, R. Sinha, and G. Prasad, "Mu and beta rhythm modulations in motor imagery related post-stroke EEG: a study under BCI framework for post-stroke rehabilitation," *BMC Neuroscience*, vol. 11, pp. 1–2, 2010.

[37] C. Tangwiriyasakul, V. Mocioiu, M. J. van Putten, and W. L. Rutten, "Classification of motor imagery performance in acute stroke," *Journal of neural engineering*, vol. 11, no. 3, p. 036001, 2014.

[38] Y. Liu, M. Li, H. Zhang, H. Wang, J. Li, J. Jia, Y. Wu, and L. Zhang, "A Tensor-Based Scheme for Stroke Patients' Motor Imagery EEG Analysis in BCI-FES Rehabilitation Training," *Journal of neuroscience methods*, vol. 222, pp. 238–249, 2014.

[39] T. Solis-Escalante, J. Belda-Lois, and G. Müller-Putz, "Cortical Activation Patterns of Cue-Paced Foot Movement in Subacute Stroke Patients," *Biomed Tech*, vol. 58, p. 1, 2013.