Error Correction Output Codding Coupled with the CSP for Motor Imagery BCI Systems

Soroosh Shahtalebi, *Student Member, IEEE* and Arash Mohammadi, *Member, IEEE*
Concordia Institute for Information System Engineering (CIISE), Concordia University,
Montreal, QC, Canada H3G 1M8 Emails: {s_shahta, arashmoh}@encs.concordia.ca

***Abstract*—Motivated by the fact that modeling and representation of multi-class signal patterns plays a critical role in Electroencephalogram (EEG)-based brain computer interface (BCI) systems, the paper proposes the coupling of error correction output coding (ECOC) with the common spatial pattern (CSP) analysis. Referred to as the ECO-CSP framework, the ECOC approach is applied to EEG motor imagery classification problem. A BCI system designed to operate in real world conditions, must be able to discriminate multiple tasks and activities. This fact, expresses the urge to develop/implement classifiers intrinsically designed for multi-class problems. One of such techniques which is well regarded in other fields but has not yet been applied to EEG-based classification is the ECOC. The paper addresses this gap.The BCI Competition IV-2a dataset is used to evaluate the performance of the proposed ECO-CSP framework. Our results show that ECO-CSP achieve similar performance in comparison to the state-of-the-art algorithms but is extensively simpler with significantly less computational overhead making it a practical alternative for real-time EEG motor imagery classification tasks.**

***Index Terms*: Brain-computer interface (BCI), Common spatial patterns, Electroencephalogram (EEG), Error correction output coding, Motor Imagery.**

I. Introduction

The human brain is the most intriguing signal processing system in existence due to its ability to extract/fuse information from several streaming signal modalities adaptively and in real-time. This fact has inspired extensive research on development of brain computer interfaces (BCI) [1] which allow users to communicate with outside world using their brain waves. The BCI systems have several practical applications of engineering importance such as rehabilitation/assistive systems [2]–[4], and controlling a wheelchair or neuro-prosthesis for disabled individuals [5]. The BCI is a key member of human-in-the-loop Cyber-physical systems (CPSs) [1], a new class of systems promoting innovative research that further augment human's interaction with physical world. The performance of such artificial interfaces, however, rarely matches that of humans rendering their practical applications considerably limited.

A BCI system typically consists of multiple modules closing the communication link including: (i) A brain imaging modality to record brain activities, and; (ii) A signal processing module to process the brain signals and produce proper control outputs. Electroencephalogram (EEG) is the most commonly used measurement modality for monitoring brain activities. The EEG data is simultaneously collected from a multitude of channels at a high temporal resolution, yielding high dimensional matrices for representation of brain activities. In addition to its unsurpassed temporal resolution, EEG is wearable, and more affordable than other neuro-imaging techniques, therefore, is a prime choice for any type of practical BCI. Processing of EEG signals typically consists of two main steps: feature generation and feature translation (classification). While the latter tries to make sense of the previously extracted features, the former aims at extracting relevant information from raw signals and avoid the so-called "curse of dimensionality". Several motor-related EEG signal modalities have been investigated in the literature among which typically Sensorimotor Activities [6] is considered as the leading developed modality. Frontal and parietal cortices exhibit rhythmic activity in the 8-12 Hz and 13-30 Hz ranges, respectively called μ and β rhythms. In case of a voluntary movement these rhythms fade out, a phenomenon referred to as event-related desynchronization (ERD). Once the movement is over, these rhythms emerge again and produce an event-related synchronization (ERS).

The Common Spatial Patterns (CSP) [7], [8] is an effective tool for discriminating imagery movements based on EEG signals which is commonly used to detect abnormalities in EEG signals. The CSP algorithm introduces spatial filters for multi-channel EEG recordings to better locate and extract ERD and ERS waveforms. Consequently, the CSP methodology enhances EEG channels containing higher weights for the ERD and the ERS. Several recent works have shown that performance of the CSP for discriminating motor imagery (MI) tasks is superior in comparison to its counterparts. Therefore, the CSP has been extended and improved from different aspects to enhance its classification performance, e.g., filter bank common spatial patterns (FBCSP) [9], regularized common Spatial patterns (RCSP) [10]–[12], and separable common spatio-spectral patterns (SCSSP) [13]. Improved performance of the such algorithms, however, comes at cost of high computation which is the main motivation of this work.

In this paper, we propose to couple error-correction output codding coupled with CSP, refers to as the ECO-CSP framework which utilizes the ECOC classifiers during the last stage to classify EEG signals into different MI classes. The ECOC is one of the most widely used classification algorithms in other application domains such as text-recognition, and primary deals with multi-class problems by reducing the original task into a series of binary sub-classification problems. To the

ISBN 978-0-9928626-7-1 © EURASIP 2017

best of our knowledge, the ECOC has not yet been applied to EEG classification. The proposed ECO-CSP framework addresses this gap. In brief, the ECO-CSP builds spatial filters by means of the CSP and extracts appropriate features which are then provided to a ECOC classifier to discriminate the imagery tasks on unseen EEG recordings. The rest of the paper is organized as follows: Section II formulates the EEG classification problem. The proposed ECO-CSP is developed in Section III. Simulation results are provided in Section IV. Finally Section V concludes the paper.

II. Problem Formulation

Throughout the paper, the following notations are used: non-bold letter x denotes a scalar variable, lowercase bold letter $\boldsymbol{x}$ represents a vector, and capital bold letter $\boldsymbol{X}$ denotes a matrix. The real domain is represented by $\mathbb{R}$. The transpose and trace of a matrix $\boldsymbol{X}$ are, respectively, denoted by $\boldsymbol{X}^T$, and $\text{Tr}(\boldsymbol{X})$.

We consider supervised learning from EEG signals based on the available set of EEG epochs (trials) denoted by $\boldsymbol{X}_i \in \mathbb{R}^{N_{\text{ch}} \times N_{\text{t}}}$, for $(1 \leq i \leq N_{\text{Trial}})$, where N_{Trial} is the total number of trails used for processing; N_{ch} is the number of EEG channels (electrodes), and; N_{t} is the number of time samples collected from each electrode in one trial. The training dataset is denoted by $\{(\boldsymbol{X}_i, L_i)\}$, for $(1 \leq j \leq N_{\text{Trial}})$, where L_i represents the label corresponding to the ith trial, e.g., $\mathcal{L}_i$ could be "MI of right hand", "MI of left foot", or "MI of left hand". Before processing EEG signals for classifying MI tasks, typically, a pre-processing step is applied. In this stage, initially the power line interference is removed by applying a notch filter. Then, bandpass filtering is applied to extract 0.5-100 Hz frequency contents. The pre-processing step is commonly followed by constructing second-order statistics of the EEG epochs, i.e., computing the sample covariance matrix corresponding to each trial $\boldsymbol{X}_i$ as follows

$$\boldsymbol{C}_i = \frac{1}{N_{\text{t}} - 1}\big(\boldsymbol{X}_i - \boldsymbol{\mu}_i\big)\big(\boldsymbol{X}_i - \boldsymbol{\mu}_i\big)^T, \tag{1}$$

where $\boldsymbol{\mu}_i$ is the column-wise sum of $\boldsymbol{X}_i$. Since $\boldsymbol{X}_i$ is obtained from bandpass filtering of an EEG signal, all classes have zero mean, i.e., $\boldsymbol{\mu}_i = \boldsymbol{0}$, for $(1 \leq i \leq N_{\text{Trial}})$. Therefore, the discriminant information contained in the second-order statistics of the data can be represented, instead, by the normalized spatial covariance matrix given by

$$\boldsymbol{C}_i = \frac{\boldsymbol{X}_i \boldsymbol{X}_i^T}{\text{Tr}(\boldsymbol{X}_i \boldsymbol{X}_i^T)}. \tag{2}$$

The CSP features [7] are then extracted from the normalized spatial covariance matrices and provided as input to the classifier for performing the MI classification task. It is worth mentioning that, principle component analysis (PCA) and independent component analysis (ICA), which are commonly used pre-processing techniques in other application domains, typically fail to improve the classification performance in BCI systems. The advantage of the CSP compared to the ICA and the PCA may, to a large extend, be attributed to incorporation of label information, i.e., while CSP exploits the information contained in labels in a supervised manner, ICA and PCA are unsupervised methods. This completes a brief presentation of the problem at hand. Next, we present the proposed ECO-CSP framework.

III. The ECO-CSP Framework

In this section, we present the proposed ECO-CSP framework. As stated previously, Reference [14] proved the fact that during MI tasks, the energy in the μ-band (8-13 Hz) decreases and energy in the β-band (13-30 Hz) increases. Motivated by this observation and to decrease the computational cost, in the proposed ECO-CSP, we consider only the aforementioned two frequency bands (unlike the FBCSP method which deploys nine bandpass filters). In addition, since the number of preliminary features would decrease by factor of $2/9$ for each trial, feature selection method based on mutual information is no longer required further reducing the computation complexity of the proposed ECO-CSP. Finally, the ECO-CSP utilizes a new classification methodology, the ECOC, which differentiates it from the existing CSP-based approaches. In the following two sub-sections we elaborate on different aspects of the ECO-CSP framework.

A. *ECOC-based EEG Modeling*

The ECO-CSP is developed for multi-class (more than two) EEG classification problems where the conventional binary classifiers such as support vector machines (SVM) or linear discriminant analysis (LDA) are not applicable. To address this issue, one can extend the above mentioned classical binary classifiers to multi-class settings, or adapt current binary classifiers to become applicable to multi-class scenarios, e.g., famous techniques in this category are "one vs. one" and "one vs. all". On the other hand, the ECOC is among few techniques which is developed intrinsically for multi-class classification. Reference [15] compared the performance of decision trees with that of the ECOC over a number of datasets and proposed ECOC as the general solution for multi-class classification. A later work [16] solidified the background of ECOC and proposed routines to better deploy this method for multi-class classification problems. Reference [17] successfully applied a combination of ECOC and Bayesian techniques for text classification problem. Since then, ECOC has been successfully and extensively applied [18] in different application domains other than for EEG classification in BCI systems. The paper addresses this gap.

The ECO-CSP uses a binary *coding matrix* consisting of N_{c} bit-vectors of length N_{Bits}. The set of all bit-vectors (the coding matrix) is denoted by $\boldsymbol{\mathcal{C}}$, where the ith row $\boldsymbol{\mathcal{C}}_i$ is a unique bit-vectors (referred to as a codeword) corresponding to class i, for $(1 \leq i \leq N_{\text{c}})$. In other words, for an EEG classification problem with N_{c} MI classes, codewords are of length N_{Bits}, denoted by $\boldsymbol{\Lambda} = \{\boldsymbol{\gamma}^{(1)}, \boldsymbol{\gamma}^{(2)}, \ldots, \boldsymbol{\gamma}^{(N_{\text{Bits}})}\}$, where total of N_{Bits} classifiers ($\boldsymbol{\gamma}^{(j)}$, for $(1 \leq j \leq N_{\text{Bits}})$) are constructed to decide on whether their corresponding bit is zero or one. Based on the EEG classification problem, we consider a multi-class setting consisting of $(3 \leq N_{\text{c}} \leq 7)$ different classes, and construct

ISBN 978-0-9928626-7-1 © EURASIP 2017

Classes	Codeword							
Right Hand IM	1	1	1	1	1	1	1	
Left Hand IM	0	0	0	0	1	1	1	Class 2
Foot IM	0	0	1	1	0	0	1	
Tongue IM	0	1	0	1	0	1	0	
				Classifier 4				

Fig. 1. The ECOC example with 7 bits for classifying 4 EEG classes.

codes of length $N_{\text{Bits}} = 2^{N_c - 1} - 1$ computed based on the following procedure: (i) All the entries in the first row are ones; (ii) The second row consists of 2^{N_c-2} zeros followed by $2^{N_c-2} - 1$ ones; (iii) The third row consists of 2^{N_c-3} zeros, followed by 2^{N_c-3} ones, followed by 2^{N_c-3} zeros, followed by $2^{N_c-3} - 1$ ones, and; (iv) The ith row consists of alternating runs of 2^{N_c-i} zeros and ones.

The ECO-CSP constructs an individual binary classifier for each column of the coding matrix $\boldsymbol{\mathcal{C}}$. Classifier j, for $(1 \leq j \leq N_{\text{Bits}})$, has positive instances for each class i when $\mathcal{C}_{(i,j)} = 1$. In other words, each classifier predicts whether or not a given trial belongs to a fixed subset of classes, resulting into two super-sets $\boldsymbol{S}^{(0)}$ and $\boldsymbol{S}^{(1)}$. Training the ECO-CSP classifier comprises of learning a set $\Lambda = \{\gamma^{(1)}, \ldots, \gamma^{(N_{\text{Bits}})}\}$ of independent binary classifiers. Based on this learned Λ, the correct class of an unlabeled trial $\boldsymbol{X}_i$ is hypothesized as follows: Evaluate each independent classifier based on $\boldsymbol{X}_i$ resulting in generation of $\Lambda(\boldsymbol{X}_i) = \{\gamma^{(1)}(\boldsymbol{X}_i), \ldots, \gamma^{(N_{\text{Bits}})}(\boldsymbol{X}_i)\}$. Most likely, the generated bit-vector $\Lambda(\boldsymbol{X}_i)$ will not be a row of $\boldsymbol{\mathcal{C}}$, but it will certainly be closer (in Hamming distance Δ) to some rows than to others. Trial $\boldsymbol{X}_i$ is categorized as follows

$$\boldsymbol{\Phi}(\boldsymbol{X}_i) = \text{argmin}_i \Delta\big(\boldsymbol{\mathcal{C}}_i, \Lambda(\boldsymbol{X}_i)\big), \tag{3}$$

where $\Delta(\boldsymbol{a}, \boldsymbol{b})$ is the number of bits in which vectors a and b differ. In other words, when classifying a new signal $\boldsymbol{X}_i$, we compute the Hamming distance between $\boldsymbol{\Lambda}(\boldsymbol{X}_i)$ and all available codewords $\boldsymbol{\mathcal{C}}_i$. We assign the new signal to class i if it has the minimum distance among other classes.

Illustrative Example: Fig. 1 shows an illustrative example for the task of classifying new EEG trials into $N_c = 4$ categories, i.e., {Right hand IM, Left hand IM, Feet IM, Tongue IM}. Seven classifiers are trained in terms of our running example, i.e., $N_{\text{Bits}} = 7$. For instance, in Fig. 1, the 4th classifier is responsible to distinguish trials whose label is "Right hand MI", "Tongue IM", or "Feet MI" with those whose label is "Left hand MI". In other words, the two super-sets for the 4th classifier are as follows: $\boldsymbol{S}^{(1)} = $ {Right hand MI, Tongue IM, Feet MI} and $\boldsymbol{S}^{(0)} = $ {Left hand MI}. Algorithm 1 outlines different steps of ECOC-based modeling of an EEG trial.

B. CSP-based Feature Extraction/Classification

Features required for classificatin/traning of N_{Bits} binary classifiers are obtained by extracting CSP from EEG epochs.

Algorithm 1 ECOC-BASED EEG MODELING

Input: EEG Trials: $\{\boldsymbol{X}_i\}_{i=1}^{N_{\text{Trial}}}$; Labels: $\{h_i\}_{i=1}^{N_{\text{Trial}}}$ with N_c distinct MI classes, and; Number of classifiers: N_{Bits}.

Output: The coding matrix: $\boldsymbol{\mathcal{C}}$, and; N_{Bits} trained binary classifiers $\{\boldsymbol{\gamma}^{(1)}, \ldots, \boldsymbol{\gamma}^{(N_{\text{Bits}})}\}$.

S1. *Codebook Generation*: Generate the $(N_c \times N_{\text{Bits}})$ binary coding matrix $\boldsymbol{\mathcal{C}}$.

S2. *Classifier Design*:

for $1 \leq j \leq N_{\texttt{Bits}}$ **do**

  Construct two super-sets, $\boldsymbol{S}^{(j,0)}$ and $\boldsymbol{S}^{(j,1)}$ where $\boldsymbol{S}^{(j,1)}$ consists of all labels h_i for which $\mathcal{C}_{ij} = 1$, and $\boldsymbol{S}^{(j,0)}$ is the complement set.

  Construct a binary classifier $b^{(j)}$ to distinguish $\boldsymbol{S}^{(j,0)}$ from $\boldsymbol{S}^{(j,1)}$.

end for

It is worth mentioning that each classifier $\gamma^{(j)}$, for $(1 \leq j \leq N_{\text{Bits}})$, uses its specific features, i.e., features are classifier specific. Intuitively speaking, the CSP methodology uses a linear transformation matrix to project multi-channel EEG signals into a lower-dimensional spatial subspace for the N_{Bits} classifiers representing the code book $\boldsymbol{\mathcal{C}}$. The projection matrix is used to maximize the variance of two-class signals by simultaneous diagonalization of the normalized spatial covariance matrix of EEG signals corresponding to the two class.

More specifically, for trial $\boldsymbol{X}_i$, classifier $\gamma^{(j)}$, for $(1 \leq j \leq N_{\text{Bits}})$, first derives the normalized spatial covariance matrix denoted by $\boldsymbol{C}_i^{(j)}$ based on Eq. (2). As the goal of the CSP is to discriminate two classes of data, we define $\bar{\boldsymbol{C}}^{(j,0)}$ and $\bar{\boldsymbol{C}}^{(j,1)}$ as the average of spatial covariance matrices of different trials belonging to each super-set ($\boldsymbol{S}^{(j,0)}$ and $\boldsymbol{S}^{(j,1)}$). For example, in terms of our running example, $\bar{\boldsymbol{C}}^{(4,0)}$ is the average of covariance matrices belonging to $\boldsymbol{S}^{(0)} = $ {Lefthand MI} while $\bar{\boldsymbol{C}}^{(4,1)}$ is the average of covariance matrices belonging to $\boldsymbol{S}^{(1)} = $ {Right hand IM, Feet IM, Tongue IM}. Based on the computed average covariance matrices ($\bar{\boldsymbol{C}}^{(j,0)}$ and $\bar{\boldsymbol{C}}^{(j,1)}$), the composite spatial covariance matrix denoted by $\boldsymbol{C}^{(j,c)}$ is computed as follows

$$\boldsymbol{C}^{(j,c)} = \bar{\boldsymbol{C}}^{(j,0)} + \bar{\boldsymbol{C}}^{(j,1)}. \tag{4}$$

The next step is to perform eigenvalue decomposition on the composite covariance matrix as

$$\boldsymbol{C}^{(j,c)} = \boldsymbol{U}^{(j,c)} \boldsymbol{\lambda}^{(j,c)} [\boldsymbol{U}^{(j,c)}]^T, \tag{5}$$

where $\boldsymbol{U}^{(j,c)}$ is the matrix of eigenvectors associated with the composite covariance, and $\boldsymbol{\lambda}^{(j,c)}$ is the diagonal matrix of its corresponding eigenvalues. The next step in ECO-CSP is to apply a whitening transform denoted by $\boldsymbol{P}^{(j)}$ on $\boldsymbol{U}^{(j,c)}$ as

$$\boldsymbol{P}^{(j)} = \sqrt{\big[\boldsymbol{\lambda}^{(j,c)}\big]^{-1}} [\boldsymbol{U}^{(j,c)}]^T. \tag{6}$$

Intuitively speaking, the whitening operator equalizes the variance in the space spanned by $\boldsymbol{U}^{(j,c)}$, i.e., all the eigenvalues of $\boldsymbol{P}^{(j)} \boldsymbol{C}^{(j,c)} [\boldsymbol{P}^{(j)}]^T$ are equal to one. Using the whitening matrix defined in Eq. (6), the average covariance matrices ($\bar{\boldsymbol{C}}^{(j,0)}$ and $\bar{\boldsymbol{C}}^{(j,1)}$) are transformed as follows

$$\boldsymbol{S}^{(j,0)} = \boldsymbol{P}^{(j)}\bar{\boldsymbol{C}}^{(j,0)}[\boldsymbol{P}^{(j)}]^T \tag{7}$$

$$\text{and} \quad \boldsymbol{S}^{(j,1)} = \boldsymbol{P}^{(j)}\bar{\boldsymbol{C}}^{(j,1)}[\boldsymbol{P}^{(j)}]^T, \tag{8}$$

therefore, $\boldsymbol{S}^{(j,0)}$ and $\boldsymbol{S}^{(j,1)}$ share common eigenvectors denoted by $\boldsymbol{B}^{(j)}$, i.e.,

$$\boldsymbol{S}^{(j,0)} = \boldsymbol{B}^{(j)}\boldsymbol{\lambda}^{(j,0)}[\boldsymbol{B}^{(j)}]^T \tag{9}$$

$$\text{and} \quad \boldsymbol{S}^{(j,1)} = \boldsymbol{B}^{(j)}\boldsymbol{\lambda}^{(j,1)}[\boldsymbol{B}^{(j)}]^T, \tag{10}$$

with $\boldsymbol{\lambda}^{(j,0)} + \boldsymbol{\lambda}^{(j,1)} = \boldsymbol{I}$, where $\boldsymbol{I}$ denotes an identity matrix of appropriate dimension. Since the sum of two corresponding eigenvalues is always one, the eigenvector with largest eigenvalue corresponding to $\bar{\boldsymbol{S}}^{(j,0)}$ has the smallest eigenvalue for $\bar{\boldsymbol{S}}^{(j,1)}$ and vice versa. This property makes the eigenvectors $\boldsymbol{B}^{(j)}$ useful for classification of the two distributions. The projection of whitened EEG signals onto the first and last eigenvectors in $\boldsymbol{B}^{(j)}$ (i.e., the eigenvectors corresponding to the largest and smallest eigenvalues) will provide feature vectors that are optimal for discriminating two populations of EEG signals in the least square sense. Therefore, the ECO-CSP projection matrix corresponding to classifier $\boldsymbol{\gamma}^{(j)}$, for ($1 \leq j \leq N_{\text{Bits}}$), is

$$\boldsymbol{W}^{(j)} = \left[\boldsymbol{P}^{(j)}\right]^T \boldsymbol{B}^{(j)}, \tag{11}$$

which is then used to form the decomposition (mapping) of each trial $\boldsymbol{X}_i$, for ($1 \leq i \leq N_{\text{t}}$) as follows

$$\boldsymbol{Z}_i^{(j)} = \left[\boldsymbol{W}^{(j)}\right]^T \boldsymbol{X}_i. \tag{12}$$

Term $\boldsymbol{Z}_i^{(j)}$ is computed for each direction of imagined movement (each MI class). As the variances of only a small number m of signals are suitable for discrimination analysis, only the first and last m rows of $\boldsymbol{Z}_i^{(j)}$ are used for the construction of the classifier. In other words, matrix $\boldsymbol{Z}_{i,p}^{(j)}$ is constructed from the first and last m rows of matrix $\boldsymbol{Z}_i^{(j)}$ which represents rows of $\boldsymbol{Z}_i^{(j)}$ associated with the largest eigenvalues that maximizes the difference of variance between two super-sets

$$\boldsymbol{f}_{i,p}^{(j)} = \log\left(\frac{\text{var}\big(\boldsymbol{Z}_{i,p}^{(j)}\big)}{\sum\limits_{k=1}^{2m}\text{var}\big(\boldsymbol{Z}_{k,p}^{(j)}\big)}\right), \tag{13}$$

where var$(\cdot)$ denotes the variance operator. Note that, the log-transformation in Eq. (13) is included to approximate normal distribution of the data. This completes development of the proposed ECO-CSP which is summarized in Algorithm 2.

Algorithm 2 ECO-CSP TESTING

Input: Unlabled EEG Trial $\boldsymbol{X}_{\text{Test}}$; The coding matrix: $\boldsymbol{\mathcal{C}}$, and; N_{Bits} trained binary classifiers $\{\boldsymbol{\gamma}^{(1)}, \ldots, \boldsymbol{\gamma}^{(N_{\text{Bits}})}\}$.

Output: The output class $\boldsymbol{\Phi}(\boldsymbol{X}_{\text{Test}})$ of $\boldsymbol{X}_{\text{Test}}$.

S1. Filter $\boldsymbol{X}_{\text{Test}}$ using two bandpass filters to extract μ (8-13 Hz) and β (13-30 Hz) frequency contents.

S2. Extract classifier-specific features (transforming matrices) using Eqs. (12) and (13).

S3. Compute the N_{Bits}-bit codeword ($\boldsymbol{\Lambda}(\boldsymbol{X}_{\text{Test}})$).

S4. Compute the class of the test epoch using Eq. (3).

IV. SIMULATIONS

In this section, performance of the proposed ECO-CSP framework is evaluated based on BCI Competition IV-2a dataset [19] which consists of four classes of motor imagery EEG measurements (Right hand IM, Left hand IM, Feet IM, and Tongue IM) obtained from nine subjects. Signals are recorded at sampling rate of 250Hz using 22 EEG channels and 3 monopolar electrooculogram (EOG) channels (with left mastoid serving as the reference). The original EEG signal recordings are already bandpass filtered (0.5-100Hz) and notch filtered. For each subject, two sessions are recorded (one for training purposes and the other one for evaluation). Each session is performed six times with each run consisting of 48 trials of length 3 seconds. In total and for each subject, 288 trials for training and 288 trials for evaluation are available. In order to measure performance of the proposed ECO-CSP framework and based on the recommendation from BCI competition, kappa coefficient κ is used, i.e., $\kappa = \frac{\text{CCR}-P_{\text{rand}}}{1-P_{\text{rand}}}$, where CCR represents the correct classification rate, and the value of P_{rand} for this dataset is equal to 0.25.

During the training stage, at first a segment of 6 seconds is selected from each trial. The segment starts 2 seconds before a cue is presented to the subject and lasts till 4 seconds after that. Then, two bandpass filters (in this experiments we used Chebychev type 2 filters of order 54) are applied to each segment to extract the frequency contents of μ-band (8-13 Hz) and the β-band (13-30 Hz). Finally, a segment of 2 seconds is selected starting from 0.5 second after presenting the cue. It is worth mentioning that the first step is applied to avoid windowing effects that might influence the start and end of the signal segment. The pre-processed segments of 2 seconds (each segment starts 0.5 second after the cue is presented) are then imported to ECO-CSP algorithm (as developed in Section III-B) to train $N_{\text{Bits}} = 7$ binary classifiers. Note that, incorporation of codewords with seven bits provides maximum Hamming distance of 2 bits. In this experiment, commonly used SVM and LDA classifiers are trained. During the testing phase, unseen signals are used consisting of segments of 3 seconds (each segment starts from the presentation of the unknown cue). After pre-processing, segments of length 2 seconds are utilized for spatial filtering and prediction.

Within the training stage, the performance of the trained classifiers is examined via the technique of K-fold cross validation. In this work we deployed 10-fold cross validation to maintain consistency with the works in [9], [13]. Table I represents the 10-fold cross validation of the training stage of the classifiers. which is used to observe the over-fitting or under-fitting of the training stage. The values that are provided in this table, are rather comparable with the ones that are provided in the FBCSP work [9]. Table II provides the performance of our proposed algorithm in Kappa Value

Subjects	Classifiers for each bit in class codewords													
	Classifier1		Classifier2		Classifier3		Classifier4		Classifier5		Classifier6		Classifier7	
	LDA	SVM	LDA	SVM	LDA	SVM	LDA	SVM	LDA	SVM	LDA	SVM	LDA	SVM
Subject 1	78.95	83.16	75.2	74.27	52.75	52.75	79.88	82.69	90.18	91.11	78.95	81.75	83.63	88.77
Subject 2	78.82	77.38	72.56	74.01	62.94	60.05	78.34	78.34	67.75	68.71	82.67	84.12	77.38	78.34
Subject 3	88.32	88.32	70.32	74.21	77.13	75.67	83.45	87.35	76.64	78.1	81.02	80.05	82.97	80.54
Subject 4	71.7	67.39	53	53.08	51.08	51.08	72.18	76.5	73.62	75.06	75.54	76.98	77.94	82.25
Subject 5	74.74	75.2	63.51	62.11	66.78	67.72	73.8	75.2	71.46	73.33	77.08	79.88	78.01	78.01
Subject 6	76014	79.42	67.72	66.78	68.19	66.78	78.48	80.82	65.38	65.38	80.35	80.82	76.61	77.54
Subject 7	85.45	85.94	69.94	68.48	74.79	74.79	78.18	83.52	97.09	98.55	87.39	86.42	84	84.97
Subject 8	83.27	83.78	82.26	82.76	83.27	81.75	89.86	92.4	80.23	77.19	73.17	75.16	87.83	88.34
Subject 9	88.85	89.33	75.76	77.21	68.48	68	80.12	80.12	72.85	73.82	75.27	81.09	77.21	78.18

TABLE I

10-FOLD CROSS VALIDATION RESULTS FOR TRAINING OF 7 CLASSIFIERS FOR 9 SUBJECTS OF BCI COMPETITION IV2A DATASET. THE RESULTS ARE MEASURED IN KAPPA VALUE

Subjects	Different Classifiers	
	LDA	SVM
Subject 1	48.61	48.14
Subject 2	20.83	20.37
Subject 3	50.46	49.07
Subject 4	34.25	31.49
Subject 5	18.05	16.2
Subject 6	8.33	12.5
Subject 7	57.87	63.88
Subject 8	44.44	38.42
Subject 9	37.05	37.96
Average	**35.54**	**35.34**

TABLE II

PERFORMANCE (IN KAPPA VALUE (κ)) OF 2 DIFFERENT CLASSIFIERS FOR 9 SUBJECTS OF BCI COMPETITION IV_{2a} DATASET.

measure over unseen data. The results are superior in comparison to the ones obtained from the application of conventional CSP algorithm. In contrast, our algorithm is using bare CSP analysis with a novel classification method. The proposed method is achieving rather similar results to the recent works that deploy complicated and intensive algorithms.

V. CONCLUSION

Modeling and realizing different signal patterns plays an important role in EEG-based BCI systems. Processing of an EEG signal typically consists of the following major steps: pre-processing, feature extraction, feature selection, and classification. For feature extraction, it has been proved that CSP analysis yields the most compromising results. One of the techniques which is well regarded in other research fields but has not yet been applied to EEG-based BCI, is the ECOC classifier. The paper proposes to couple ECOC with the CSP (referred to as the ECO-CSP framework) where the ECOC classification technique is applied to EEG motor imagery classification problem. The BCI Competition IV-2a dataset is used to evaluate the performance of the proposed framework. Our experiments indicate that the proposed ECO-CSP framework provides similar results when compared to other recently developed algorithms, but has extensively less computational complexity making it a practical alternative for real-time EEG motor imagery classification tasks.

REFERENCES

[1] G. Schirner, D. Erdogmus, K. Chowdhury and T. Padir, "The Future of Human-in-the-Loop Cyber-Physical Systems," *Computer*, vol. 46, no. 1, pp. 36-45, 2013.

[2] J.J. Daly and J.R. Wolpaw, "Brain-Computer Interfaces in Neurological Rehabilitation," *Lancet Neurology*, vol. 7, no. 11, pp. 1032-1043, 2008.

[3] R. Xu et al., "A Closed-Loop Brain-Computer Interface Triggering an Active Ankle-Foot Orthosis for Inducing Cortical Neural Plasticity," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 7, pp. 2092-2101, 2014.

[4] Grosse-Wentrup M, Mattia D, Oweiss K., "Using Brain-Computer Interfaces to Induce Neural Plasticity and Restore Function," *Journal of Neural Engineering*, vol. 8, no. 2, 2011.

[5] R. Leeb, L. Tonin, M. Rohm, L. Desideri, T. Carlson and J. d. R. Millan, "Towards Independence: A BCI Telepresence Robot for People with Severe Motor Disabilities," *Proc. IEEE*, vol. 103, pp. 969-982, 2015.

[6] B. J. Edelman, B. Baxter and B. He, "EEG Source Imaging Enhances the Decoding of Complex Right-Hand Motor Imagery Tasks," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 1, pp. 4-14, 2016.

[7] H. Ramoser, J. Muller-Gerking, and G. Pfurtscheller, "Optimal Spatial Filtering of Single Trial EEG during Imagined Hand Movement," *IEEE Trans. Rehab. Eng.*, vol. 8, no. 4, pp. 441–446, 2000.

[8] Z. J. Koles, "The quantitative extraction and topographic mapping of the abnormal components in the clinical eeg," *Electroencephalography and clinical Neurophysiology*, vol. 79, no. 6, pp. 440–447, 1991.

[9] K. K. Ang, Z. Y. Chin, C. Wang, C. Guan, and H. Zhang, "Filter bank Common Spatial Pattern Algorithm on BCI Competition IV datasets 2a and 2b," *Frontiers in neuroscience*, vol. 6, p. 39, 2012.

[10] H. Lu, K.N. Plataniotis, and A.N. Venetsanopoulos, "Regularized Common Spatial Patterns with Generic Learning for EEG Signal Classification," *IEEE Int. Con. Eng. Medicine & Biology*, 2009, pp. 6599-6602.

[11] H. Lu, H.-L. Eng, C. Guan, K. N. Plataniotis, and A. N. Venetsanopoulos, "Regularized common spatial pattern with aggregation for eeg classification in small-sample setting," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 12, pp. 2936–2946, 2010.

[12] 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.

[13] A.S. Aghaei, M.S. Mahanta, and K.N. Plataniotis, "Separable Common Spatio-Spectral Patterns for Motor Imagery BCI Systems," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 1, pp. 15–29, 2016.

[14] G. Pfurtscheller and F. L. Da Silva, "Event-related eeg/meg synchronization and desynchronization: basic principles," *Clinical neurophysiology*, vol. 110, no. 11, pp. 1842–1857, 1999.

[15] T. G. Dietterich and G. Bakiri, "Error-correcting output codes: A general method for improving multiclass inductive learning programs," in *AAAI*, 1991, pp. 572–577.

[16] T. G. Dietterich and G. Bakiri, "Solving multiclass learning problems via error-correcting output codes," *Journal of artificial intelligence research*, vol. 2, pp. 263–286, 1995.

[17] A. Berger, "Error-correcting output coding for text classification," in *IJCAI-99: Workshop on machine learning for information filtering*, 1999.

[18] M. Liu, D. Zhang, S. Chen and H. Xue, "Joint Binary Classifier Learning for ECOC-Based Multi-Class Classification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 11, pp. 2335-2341, Nov. 1 2016.

[19] M. Tangermann, *et al.*, "Review of the BCI Competition IV," *Frontiers in neuroscience*, vol. 6, p. 55, 2012.

ISBN 978-0-9928626-7-1 © EURASIP 2017