IJCTA, 9(36), 2016, pp. 53-65

Common Spatial Pattern Algorithm Based Signal Processing Techniques for Classification of Motor Imagery Movements: A Mini Review

Deepa Beeta Thiyam* and Rajkumar E. R.*

Abstract : Brain-computer interfacing (BCI)is one of the most attractive fields of research in neural and rehabilitation engineering. Although, it still needs lots of improvement for developing a practically reliable BCI system. The major step to improve the BCI performance is by developing a robust signal processing algorithm. Common spatial pattern (CSP) filtering is one of the most effective algorithms for feature extraction in motor imagery based BCI (MI-based BCI). However, it still has a lot of disadvantages and different approaches had been proposed to overcome these disadvantages. In order to develop a robust algorithm, it is necessary to understand the causes of the problem and find a solution. In this paper, we present a mini review of the different CSP based algorithms, findings, its disadvantages and the solutions. The limitations are discussed in details. This can help new researchers to have a full overview of the existing CSP based approaches for classification of MI movements.

Keywords : Motor imagery, Brain-computer interface, Common spatial pattern filter, electroencephalogram.

1. INTRODUCTION

Brain-computer interface (BCI)(2) translates the brain signals into control commands. It has been commonly used in both medical and non-medical applications. In medical applications, it is mainly used a communication channel for paralyzed people with the external world. A person suffering from amyotrophic lateral sclerosis (ALS), multiple sclerosis, cerebral palsy and complete locked-in are unable to move body limbs due to damage in the peripheral nervous system. Here comes the role of BCI that allows this group of people to interact with the external world. It uses the brain signals a control commands for operating assistive devices like a wheelchair, robotic arm and speller without the peripheral nervous system.

There are different ways for acquiring brain signal for BCI application, such as electroencephalography (EEG), positron emission tomography (PET), functional near-infrared spectroscopy (fNIRS) and magnetic resonance imaging (MRI) etc. Among this, EEG is the most appropriate techniques for BCI application because of its non-invasive and easy to use nature. The EEG system acquires the electrical brain signals from the scalp using the surface electrodes(3). Moreover, the control signals of BCI can be categorized into different types. They are: visual evoked potentials (VEPs), slow cortical potentials, P300 evoked potentials and sensorimotor rhythms. VEPs mainly occurred in the visual cortex regions after visual stimulus which is indicated by the sudden raised in the amplitude after the stimulus. Slow cortical potentials represent the change in voltage during the cortical activity. It gives negative value with the increased in neural

* Department of Sensors and Biomedical Technology School of Electronics Engineering VIT University, Vellore-632014 Tamilnadu, India.

activity and positive value with decreased in neural activity. P300 evoked potentials are represented by a positive peak in the amplitude after receiving external stimulus. The peak is observed after 300ms after the stimulus. Sensorimotor rhythms are denoted by the changes in the brain rhythm during the motor imagery (MI) movements. It can be observed in the μ and β rhythm. The change in the rhythmic activity of the brain is called as event-related synchronization and event-related desynchronization (ERS/ERD). BCI can be classified into different types, among them EEG MI-based BCI is commonly used.

2. EEG MI-BASED BCI SYSTEM

The input signals used for this type of BCI system is a motor imagery signal. The EEG signals are acquired during the motor imagery movement. The different types of MI movements are decoded as a control commands for the system. The block diagram of EEG MI-based BCI system is shown in Fig 1.

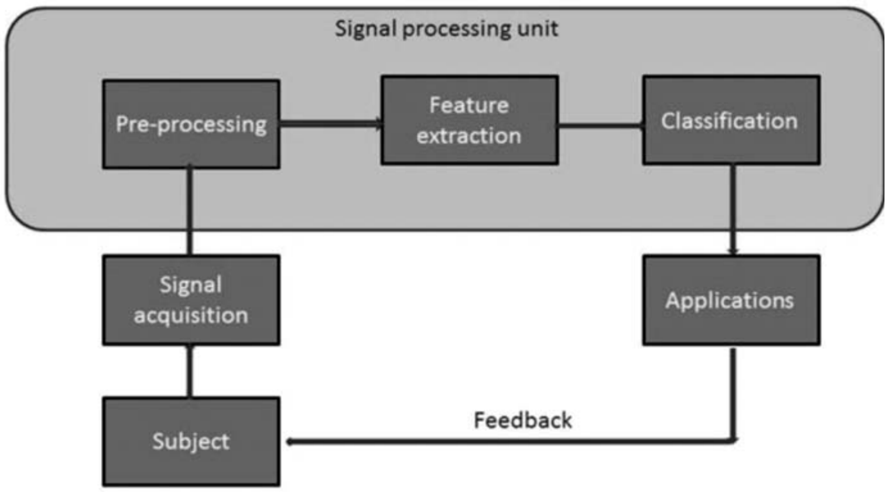

Figure 1: Basic block diagram of BCI system

The block diagram consists of the signal acquisition system like EEG system. The acquired signals are then given to the signal processing unit that consists of the pre-processing, feature extraction and classification. The input signals are band-pass filtered between a particular frequency bands. For MI-based EEG the signals are bandpass filtered between the μ and β rhythms. In this particular band, we can observe the ERS and ERD during the motor imagery process over the corresponding motor cortex area. After band pass filtering the next step consists of the feature extracting. There are various techniques used in BCI for feature extraction such as common spatial pattern filtering (CSP), principal component analysis (PCA), independent component analysis (ICA), common average referencing (CAR), frequency normalization etc. The extracted features are given to the classifier for discriminating into various classes. The commonly used classifiers for BCI are linear discriminant analysis (LDA) and support vector machine (SVM). Finally, the extracted features are used as control commands for directing the assistive devices.

3. METHODS REVIEW

Many reviews have been presented related to EEG signal processing for BCI applications. In this paper, we will focus only on the CSP based feature extraction algorithms. Among the other feature extraction techniques, CSP is the most effective algorithm for classification of two classes. Despite the fact that CSP is not robust to outliers present, a number of approaches had been proposed as an extension CSP. We reviewedall the existing algorithms for MI classification that are based on CSP algorithm. In this paper, we consider only the following parameters: motor imagery EEG signals, CSP based approaches. We have neglected the other approaches which are not based on CSP as well as we did not consider the other control signals like ECOG, P300, evoked potentials etc. This review includes all the conferences and journals paper till September 2016.The summary of different approaches and the parameters used are provided in Table 1.

3.1. Common Spatial pattern filtering

The common spatial pattern (CSP) algorithm was first presented in (4) as a feature classification algorithm. It was initially used for detection of abnormalities component in the clinical EEG (5). It was also used for the classification between normal and abnormal EEG (6). Later, it was used for the discrimination of two class movement from the single trial EEG (7). A similar study had been performed for classification of motor imagery movements from multichannel EEG data (8) *e.g.*, associated with imaginary movement. One-sided hand movement imagination results in EEG changes located at contra- and ipsilateral central areas. We demonstrate that spatial filters for multichannel EEG effectively extract discriminatory information from two popula-tions of single-trial EEG, recorded during left-and right-hand movement imagery. The best classification results for three subjects are 90.8%, 92.7%, and 99.7%. The spatial filters are estimated from a set of data by the method of common spatial patterns and reflect the specific activation of cortical areas. The method performs a weighting of the electrodes according to their importance for the classification task. The high recognition rates and computational simplicity make it a promising method for an EEG-based brain\u2013computer interface. Index Terms\u2014Assistive communication, electroencephalograph (EEG.

The motor imagery process involves the oscillation of brain rhythm in the motor cortex area. To visualize the effect of event-related synchronization and desynchronization, the CSP filters are projected as a source signals to observe the discrimination between the two classes (9).

The main objective of CSP filter is to find the most discriminative spatial filter that can distinguish between the two classes. Let us consider $\mathrm{C}ov_1$ and $\mathrm{C}ov_2$ as the two average covariance matrices of class 1 and class 2. The CSP filter can be solved by the Rayleigh quotient given by

$$\mathrm{J}(w) = \frac{w^{\mathrm{T}}\mathrm{C}ov_1 w}{w^{\mathrm{T}}(\mathrm{C}ov_1 + \mathrm{C}ov_2)w} \tag{1}$$

Where, w is the spatial filter and T represents the transpose. $\mathrm{C}ov_c$, $c \in$ can be obtained using the input EEG signals $\mathrm{X} \in \mathbb{R}^{nxt}$, n represents the no. of channels and denotes the sample point. The average covariance matrix of each class is given by

$$\mathrm{C}ov_c = \frac{1}{2}\sum_{i=1}^{\mathrm{N}} \frac{\mathrm{X}_i \mathrm{X}_i^{\mathrm{T}}}{\mathrm{trace}\,(\mathrm{X}_i \mathrm{X}_i^{\mathrm{T}})}, c \in (1,2) \tag{2}$$

Here, N denotes the total no. of trials for each class. The set of spatial filters, W = [w_1, ... w_p], no. of features to be extracted can be obtained by solving eqn. (1) with a generalized eigenvalue problem

$$\mathrm{C}ov_1 w = \lambda \mathrm{C}ov_2 w \tag{3}$$

where λ represents the eigenvalues which are arranged in descending order. The first three and the last three eigenvalues were selected and the corresponding eigenvectors represent the spatial filters for the two classes. The estimated source signals can be represented as:

$$\mathrm{Y} = \mathrm{W}^{\mathrm{T}}\mathrm{X}$$

The features are determined by taking the log variance of the filtered signals. Inspite of being the most effective method for classification of MI movement, CSP is easily affected by nonstationarities and outliers. Another disadvantage of CSP is that it is not effective for multiclass discrimination. Several methods were proposed based on standard CSP for improving the performance as well as for discriminating multiclass movements. We have classified the different regularization approaches as follows:

3.2. Regularization in estimating covariance matrices

The presence of outliers or nonstationarities leads to the improper estimation of the covariance matrices. Since the solution of CSP directly depend on the covariance matrices. It hinders the classification performance. In order to improve the performance, it is necessary to reduce the nonstationarities from the covariance matrices and can be obtained by proper estimation of covariance matrices. There are several approaches for estimating the robust covariance matrices.

An adaptive spatial filter based approach had been proposed that replace the training data of CSP by the known information. It maximizes the ratio between the variance of the region of interest and the variance of acquired EEG(10). A robust method had been proposed to reduce the effect of outliers in the classification process. This method used the Fast-MCD (Minimum Covariance Determinant) estimator for estimating the class covariance matrices. The scales of the weight matrix of the projected signal are calculated using Median Absolute Deviations (MAD) estimate(10).

Another robust approach is by using a convex set of class covariance matrices instead of using only one covariance matrices for each class. This involves finding the worst covariance matrix from the set and optimizing the objective function based on it. Later, the spatial filter is obtained by applying the eigenvalue decomposition similarly with the standard CSP algorithm(12). The similar approach was presented in(13) using maxmin approach .

The author of (14)uses the average covariance matrices by minimizing the mean squared errors to reduce the outliers present in the trials. Another author applied the regularization process based on generic learning. This is done by introducing two regularization parameter for estimating the covariance matrices by shrinking the covariance matrix towards the identity and generic matrix (15)(16). Another similar approach of CSP regularization is by shrinking to identity matrix using Diagonal loading and the selection of regularization parameter is done automatically using Ledoit and Wolf's method(17).

Sannellli et al. proposed a method to re-weight the outliers trials using Relevant dimensionality reduction (RDE). RDE is mainly performed based on the kernel principal component analysis which separates the trials with the useful information. After re-weighting the components the standard CSP is performed(18). Another approach is by utilizing signal subspace analysis method before performing CSP algorithm. The SSA method can separate between the stationary and nonstationaritiessignals and projected the data to stationary subspace. CSP algorithm is performed after that. This helps to improve the classification accuracy(19).

The composite CSP (CCSP) uses the other subject information in obtaining the spatial filters for different movements. It employed two methods: the first one is by adding the covariance matrices with a fewer sample from the other subject and the second one includes selecting the covariance matrices of the subject which have similar prediction with the subject of concern by using kl divergence(20).A similar approach had been proposed that uses the other subject data to robustify the covariance matrix. Here, an algorithm has been proposed to select the best subject and the selection of the regularization parameter is done by using the crossvalidation method (21)(16).

The SSA and CSP method have the disadvantage of misclassification for the multiclass data. To overcome this problem a principled way of obtaining the stationary subspace was obtained. In this approach, the stationary subspace is obtained by finding the KL divergence between the average distribution of the group of subjects and the individual epoch(22)(23).Most of the methods described above uses heuristic approach but the author of (24) proposed an automatic feature fusion method. This method is based on multiple kernel learning that selects the important feature from a group of a feature from different subject or sessions.

3.3. Regularization in the objective function

In this approaches, a penalty term is incorporated in the CSP objective function to increase the robustness of the algorithm. The penalty term can be other additional EOG signals, other subject data, previous session data and other normalization parameters etc. The approaches that use regularization in the objective function are discussed below:

The author of (25) proposed an invariant CSP (iCSP) by adding a penalty term in the denominator of the CSP objective function. This penalty is obtained from the additional EOG signal to reduce the effect of eye movement. A different approach for extract the stationary subspace, a stationary CSP (sCSP) was proposed in (26) . The stationary subspace is obtained by including a penalty term in the CSP objective function. The penalty term represents the non-stationarities present and was calculated by comparing each trial covariance matrices with the average covariance matrix of each class.

Other approach proposed for regularization of CSP includes incorporating the Tikhonov and weighted Tikhonov regularization as the penalty term in the CSP objective function(16). Another similar approach uses the global spatial filter from all subjects and the subject specific filter as a regularization parameter in the CSP objective function. The disadvantages of this method are that it takes time in computation and needs more training data(27).

Another regularization algorithm called CSP-L1 was proposed that uses L1 norm instead of L2 in the standard CSP objective function. The difficulty in optimization is solved by using an iterative algorithm(28). A similar approach (L1-SVD-CSP) was presented that uses L1-norm for the estimation of spatial filters by performing singular value decomposition(29)one of the popular feature extraction strategies is Common Spatial Patterns (CSP. The extension of CSP-L1 was proposed recently in (30) that included the noise model as a penalty term in the objective function of the CSP-L1 method to increase the robustness . The CSP-L1 and L1-SVD-CSP both are still dense. Therefore, to overcome this problem a sparse CSP-L1 (sp-CSPL1) was proposed that used L1 norm regularization two times. Initially, it was to generate sparsity and later is to induce robustness(31). The author of (32) proposed a more generalized CSP using L_p norm.The comprehensive CSP (CCSP) was proposed in (33) that used the correlation of the temporal information of the unlabelled trial as a regularizer for reducing the nonstationarities in the CSP objective function.

3.4. Regularization in the estimation of covariance matrices and objective function

The authors of (34)and (35)combine both the estimation of covariance matrices and incorporation of penalty term in one step. The other methods describe in (20),(21) , (17) and (27) failed if the variation of pattern between the subject is large. To deal with this problem a different approach had been proposed in (34). In this approach, instead of extracting the similar subject information the authors consider the nonstationarities information from other subjects. Later this information is incorporated in the objective function as a penalty term.

The stationary CSP combine the extraction of stationary subspace and CSP in one objective function, unlike the other stationary method that used two steps for obtaining the spatial filter. This method does not employ other subjects data(35).

3.5. Spatio-spectral filtering

This approach of improving the performance of CSP is by finding both the spatial and spectral filter because the performance of the standard CSP filter also depends on the selection of the operational frequency. The operational frequency changes from subject to subject. The improper selection of operational frequencies may lead to reduce in the classification accuracy. In order to overcome this problem, many approaches have been proposed which enables the system to select manually the subject-specific frequency band.

One of the approaches known as Common Spatio-Spectral Patterns (CSSP) is presented in (36)non-stationary, and contaminated with artifacts that can deteriorate discrimination/classification methods. In this paper, we extend the common spatial pattern (CSP. It can be considered as the extension of the standard CSP algorithm. In this approach, the frequency filter for each channel is also included along with the spatial filter. This is mainly done by including the one delay tap as a new channel. The limitation of this approach is that it has limited flexibility for frequency filter. To overcome this problem Common Sparse Spectral-Spatial Patterns (CSSSPs) has been proposed (37). In this method, the FIR filter has been obtained which increased the flexibility but this system needs tuning parameter and involves complex optimization process. To overcome this problem a spectrally weighted CSP (SPEC-CSP) was proposed that ease the optimization of the spectral and spatial filter. The optimization is done separately (38)(39). The iterative spatio-spectral pattern learning (ISSPL) improves the SPEC-CSP by optimizing the entire temporal and spatial filter in a common objective function (40)machine learning is carried out in two consecutive stages: feature extraction and feature classification. Feature extraction has focused on automatic learning of spatial filters, with little or no attention being paid to optimization of parameters for temporal filters that still require time-consuming, ad hoc manual tuning. In this paper, we present a new algorithm termed iterative spatio-spectral patterns learning (ISSPL.

Instead of optimizing a single arbitrary FIR filter within the CSP algorithm, SBCSP uses a Gabor filter bank that decomposes the EEG measurements into multiple sub-bands. Spatial filters that use the CSP algorithm are then employed on each of these sub-bands. After obtaining subband scores, recursive band elimination or a classification algorithm is employed to fuse the sub-band score. Another classification algorithm is then employed to classify the fused sub-band score. Although SBCSP can use different sub-band score fusion techniques and classification algorithms, only the results from the use of the Support Vector Machine (SVM) to fuse the sub-band score as well as to perform classification are presented in (41)motor imagery is considered as one of the most effective ways. Different imagery activities can be classified based on the changes in mu and/or beta rhythms and their spatial distributions. However, the change in these rhythmic patterns varies from one subject to another. This causes an unavoidable time-consuming fine-tuning process in building a BCI for every subject. To address this issue, we propose a new method called sub-band common spatial pattern (SBCSP.

In CSSP and CSSSP, only the temporal filter optimizes for each subject without considering the effect of different frequency information. This leads to the idea of obtaining the frequency information using the wavelet transformation. This approach uses a Daubechies mother wavelet for decomposition of the EEG signals into low frequency and high frequency. Theextracted signal is the time varying series of EEG from each band. The spatial filter was obtained from the extracted signal using CSP algorithm (42). The author of (43) proposed frequency weighted method based CSP where optimization of the spectral and spatial was done separately. The weighting of the spectrum is done in the frequency domain as well as it directly uses the straight forward method.

Another approach called local temporal CSP (LTCSP) is proposed by considering the temporal information together with the spatial information in obtaining the variance (44). But, similar like CSP this method does not consider the within-class variance. Additionally, this method involves selection of the weights parameter manually which is not easy. To overcome this problem, two approaches: discriminative LTCSP (dLTCSP) and adaptive LTCSP (aLTCSP) were proposed (45).

In FBCSP, the first step is to filter the EEG signals using multiple bands of frequencies and obtaining the spatial filters for each band using the CSP algorithm. The filterbank used for this approach is zero-phase Chebyshev Type II Infinite Impulse Response (IIR) filters. It is followed by feature extraction and the classification algorithm. The main advantages of FBCSP are that it can use any feature selection and classification algorithm as well as it will select only the effective features. Selection of only the effective filters makes the algorithm reduce the computational complexity (46). But FBCSP was proposed only for classification of two classes and it involves manual selection of time segment. An extension of FBCSP is the discriminative filter bank common spatial pattern (DFBCSP) in which the frequency band is selected by obtaining fishers ratio according to the subject instead of using fixed frequency band (47). Another approach called as Sliding Window Discriminative CSP (SWDCSP) for selection of frequency band was proposed in (48). In this approach, it extracts the feature of the entire overlapping frequency band and extracts the features using affinity propagation methods. Optimal Spatio-Spectral Filter Network (OSSFN) approach allow selecting the frequency band with maximal mutual information (49). Another extension of FBCSP includes finding the spatial filter for different time segment and different frequency band and selecting the appropriate feature by using mutual information(50)(51)(52). The author of (53) proposed a

Bayesian Spatio-Spectral Filter Optimization (BSSFO) for obtaining a different weight for the different frequency band . In spite of the improvement BSSFO still did not give the optimal solution. An extension of BSSFO was proposed recently that include Laplacian filtering and update the important features (54).

Most of the above approaches determined the subject-specific frequency band manually. In order to select the subject-specific frequency band automatically, a new approach called sparse filter band common spatial pattern (SFBCSP) was proposed in (55). It was done by using a set of a frequency band that overlapped and obtaining the spatial filters for each band using CSP. The feature selection is done using sparse regression method. The FBCSP do not have the ability to discriminate the feature power obtained from different frequency band. Recently, a separable common spatio-spectral patterns (SCSSP) algorithm was proposed that can discriminate the feature power from different frequency band. The extractions of features were done by analyzing both the spatial and spectral characteristics of the observed signals (56).

Another generalized CSP methods: FCSP and FCSSP based on Fisher discriminant criterion was presented in (57). In this approach, the CSP function is replaced by more random function that was obtained using the spatially filter signal. The discrete Fourier transform was applied to take the frequency information into account. The overall objective function was presented similarly as a Fishers discriminant function.

Spatio-temporally regularized common spatial patterns(STR-CSP) that have only a single objective function to optimize both the spatial and temporal filter simultaneously. Moreover, the L2 norm was also included as a penalty term in the objective function. Thus, this makes the method to reduce the problem of overfitting(58)CAA. Common Spatial Patterns (CSP.

Another different approach of extracting both spatial and spectral filter for classification of motor imagery EEG signals was proposed in (59)common spatial pattern (CSP. It combined fisher wavelet packet decomposition (WPD) together with CSP.

3.6. Adaptation of features

The EEG signal varies from subject to subject as well as from session to session. In order to get the higher performance, BCI needs to be calibrated beforehand. The calibration process consumes a lot of time. The effective BCI system should have less calibration time with effective performance ability. To address this problem the offline study by using previous session data for the same patient has been used for the calibration process. It is done by finding the distance and clustering them into different classes. This helps to reduce the time for calibration for the next sessions (60). This method is further extended to the online study by learning the adaptive classifier using only the meaningful features (61).

3.7. Sparse filtering

Most of the BCI system involves using a large number of electrodes which may be unwanted for practical BCI. The standard CSP algorithm considers all the electrodes in obtaining the spatial filters. A group of researchers proposed different approaches for reducing the number of channels or to select only the channels that contribute discriminative information. However, most of the discriminative information is collected by the electrodes near the sensory motor cortex; it is also true that some of the surrounding electrodes also contribute the useful information.

Gugeret. al proposed a method for real-time application that weights the electrodes. Additionally, reduce the noise present in each electrode by using the correlation between the two adjacent electrodes (62). The author of (63) proposed a method for reducing the channel but obtaining the weight from the CSP algorithm. However, eliminating the channels also leads to neglecting some of the useful information which hinders the classification performance. To improve this method another author proposed a method for finding the minimal set of electrodes to be kept not to hinder the classification performance. This is done by incorporating an L_1-norm as a regularization parameter in the CSP objective function (64)(65).

Arvanehet. al. proposed a sparsity common spatial pattern (SCSP) filter for finding the minimum number of channels to obtained the similar performance obtained while using all the channels. This is done by including the Lp norm as a regularization parameter. This method assists in choosing the least number of the channel without compromising the accuracy (66)(67). Another approach for implementation of sparsity is proposed by using zero-norm optimization and by using Recursive Feature Elimination (RFE) but the zero norm approach cannot decrease the number of channels less than 25 which affects the performance. The author further suggested for combination of this approach with RFE method (68).

Table 1
Summary of all the CSP based regularization approaches

	Approaches	*Parameters*
CSP	Regularization in estimating covariance matrices	Variance between region of interest and acquired EEG signals(9), using Fast MCD method (10), using convex set of covariance matrices (11), maxmin (12), minimum mean squared error (13), Identity and generic matrix (14)(15), diagonal loading (16), pruning (17), stationary subspace (18),other subjects data (19)(20), kl divergence(21)(22), multiple kernel(23)
	Regularization in the objective function	Additional EOG signal (24), Difference of the trial covariance and average covariance matrix (25), Tikhonov and weighted Tikhonov (15),other subjects (26) L1 norm (27-30), Lp norm (31), correlation of temporal information (32)
	Regularization in the estimation of covariance matrices and objective function	Other subjects data (33)(34)
	Spectral filtering	Both spectral and spatial filter (35-38)(42), temporal and spatial filter(39), multiple freq sub bands (40), Wavelet(41), local temporal information (43)(44), frequency bands (45-47), Frequency bands and mutual information (48-51), Bayesian approach (52)(53), Frequency band overlapping (54), Fisher discriminant criterion for spatio-spectral filtering (55)(56), L2 norm regularization (57), Wavelet packet decomposition (WPD)(58)
	Adaptation of features	Previous session data (59)(60)
	Sparse filtering	Correlation between two electrodes(61), weights(62), L1 norm(63)(64), Lp norm(65)(66), zero norm and Recursive Feature Elimination (RFE) (67)
	Divergence-based CSP	Kl divergence(68)(69), β divergence (70)(71), β-div with regularization (70)(72), Bhattacharya and γ divergence (73)
	Multiclass CSP	mCSP (74), information theory (75), source separation (76) beamforming(77)(78), multiclass FBCSP(79), JAD(80), Bayesian method(81), harmonic mean(82), JAD and mutual information(83), probabilities(84), Bayesian errors(85), time-frequency domain reconstruction(86).

3.8. Divergence-based CSP

Divergence is the most effective method in the field of information theory for finding the dissimilarities between the two distributions. The standard CSP only separates the variance between two class and it does not consider the within class information. This makes the CSP not robust in the presence of non-stationarities. A group of researchers proposed divergence based CSP approach. The authors of (69) proposed a new objective function that maximizes the variance between two class at the same time minimizes the variance within the class based on divergence. The within class variance is obtained by finding the KL divergence between the individual trial and the average of all the trial of that particular class (69)(70).

Beta divergence method had been used for estimating the robust covariance matrices in (71).The same group proposed a novel robust method to obtained the CSP solution by maximizing the beta-divergence between the two average class covariance matrices (72). They also proposed a regularized divergence based CSP method. The penalty terms used for this approach were the divergence between the between session, across subject and multisubject(71). The extended work is to consider both the robustness and nonstationarities problem together presented in (73). The new objective function had been proposed by considering the sum of beta-divergence between individual trials of both the classes. The penalty term is obtained by summation of beta-divergence between the trial covariance and the average covariance for each class. The penalty term reduces the nonstationarities present in the signals. Recently, a group of researcher proposed a divergence based CSP method using Bhattacharya distance and Gamma divergence (74).

3.9. CSP Multiclass

The standard CSP is mainly for the discrimination of two classes. The extension of CSP for classification of multiclass was proposed in (75). This approach considered the multiclass problem as a binary problem for obtaining the spatial filter for each class. The multiclass CSP is a heuristics process whereas the approach has been proposed that has a theoretical proof. This approach uses the information theory for selection of relevant features. In this approach, a feature has been extracted based on the mutual information between the extracted features and the class labels (76)the question of optimality of CSP in terms of the minimal achievable classification error remains unsolved. Second, CSP has been initially proposed for two-class paradigms. Extensions to multiclass paradigms have been suggested, but are based on heuristics. We address these shortcomings in the framework of information theoretic feature extraction (ITFE.

The authors of (77)CSSP, CSSSP used the idea of source separation and CSP together for discriminating multiclass MI movements. The problem of overfitting is very common in CSP. To overcome this problem, a method was proposed that uses beamforming method. This helps to extract the sources that generate only under the region of interest that had been defined priory(10)the method of Common Spatial Patterns (CSP(78)such as common spatial patterns (CSPs(79). Based on the FBCSP method which was presented in (46)discriminative patterns can be extracted from the electroencephalogram (EEG, an extended method for classification of multiclass MI movement was proposed in (80). A group of researcher formulated the CSP method in terms of joint approximate diagonalization (JAD) for separation of non-stationarities for multiclass MI discrimination (81). Other author uses Bayesian method for extracting multiclass information (82). The author of (83)however, are only suitable for the two-class paradigm. In this paper, we address this limitation under the framework of Kullback-Leibler (KL formulated CSP based on KL divergence. The spatial filter was obtained by maximizing the harmonic mean between the distributions using symmetric KL divergence.

Another approach is the extension of stationary CSP (sCSP) proposed in (26) for the multiclass problem. The effective sCSP is only for two classes to solve the multiclass problem, sCSP was used together with joint approximate diagonalization of the transformation matrix (84). The authors of (85) proposed a multiclass discriminating algorithm that is based on probabilities to solve the issue of overfitting. It clearly presents the overfitting issues and the steps for solving it. It can also be reduced to standard CSP and RCSP in the absence of noise signals. The Bayesian learning method was proposed in (86)to show the correlation between Bayes error and Rayleigh quotient which has the similar representation with CSP.

A different CSP based approach for classification of multiclass from the single channel was proposed in (87)a portable few- or single-channel BCI system has become necessary. Most recent BCI studies have demonstrated that the common spatial pattern (CSP. In this approach, a multichannel input was constructed using a single channel signal. This is done by converting the single channel data into time-frequency domain and reconstructing the multichannel input for performing CSP algorithm.

4. CONCLUSIONS

This paper presents the mini review of CSP-based signal processing techniques used for classification of motor imagery signal. We have considered only CSP based approaches and MI EEG signals. In this paper, the brief overview of the CSP based methods and its extension for classification of motor imagery signals presented. Moreover, we have discussed the limitation of the available approaches in details and the solution to overcome these problems. This will provide a strong literature by providing the current state-of-art for CSP-based approaches.

5. REFERENCES

1. Arvaneh M, Guan C, Ang KK, Quek C. Optimizing the channel selection and classification accuracy in EEG-based BCI. Biomed Eng IEEE Trans. IEEE; 2011;58(6):1865–73.
2. Vidal J. Towards direct brain-computer communication. Annu Rev Biophys ans Bioeng. 1973;2(1):157–80.
3. Berger H. Über das elektrenkephalogramm des menschen II. J für Psychol und Neurol. 1930;40:160–79.
4. Fukunaga, Keinosuke and WLK. Application of the Karhunen-Loeve expansion to feature selection and ordering. IEEE Trans Comput. 1970;19(4):311–8.
5. Koles ZJ. The quantitative extraction and topographic mapping of the abnormal components in the clinical EEG. Electroencephalogr Clin Neurophysiol. 1991;79(6):440–7.
6. Z.J. Koles, J.C. Lind PF-H. Spatial patterns in the background EEG underlying mental disease in man. Electroencephalogr Clin Neurophysiol. 1994;91(5):319–28.
7. Müller-Gerking J, Pfurtscheller G, Flyvbjerg H. Designing optimal spatial filters for single-trial EEG classification in a movement task. Clin Neurophysiol. 1999;110(5):787–98.
8. Ramoser H, Müller-Gerking J, Pfurtscheller G. Optimal Spatial Filtering of Single Trial EEG During Imagined Hand Movement. IEEE Trans Rehabil Eng. 2000;8(4).
9. Blankertz B, Tomioka R, Lemm S, Kawanabe M, Müller KR. Optimizing spatial filters for robust EEG single-trial analysis. IEEE Signal Process Mag. 2008;
10. Grosse-Wentrup M, Gramann K, Buss M. Adaptive spatial filters with predefined region of interest for EEG based brain-computer-interfaces. Adv Neural Inf Process Syst. 2007;537–44.
11. Analysis of Mathematical attack on RSA algorithm using factoring a prime number by Division algorithm method"A-ICTE sponsored National Conference Recent Research Trends in Network Security NCRTNS-2012 organized by Srinivas Institute of Technology, Valachil, Farangipet Post, Mangalore ,19th and 20th April 2012.
12. Kawanabe M, Vidaurre C, Scholler S, Müller KR. Robust common spatial filters with a maxmin approach. Proceedings of the 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society: Engineering the Future of Biomedicine, EMBC 2009. 2009. p. 2470–3.
13. Kawanabe M, Samek W, Uller K-R, Vidaurre C, Kawanabe M, Samek W, et al. Robust Common Spatial Filters with a Maxmin Approach. Neural Comput. 2014;26:349–76.
14. Kawanabe M, Vidaurre C. Improving BCI performance by modified common spatial patterns with robustly averaged covariance matrices. IFMBE Proceedings. 2009. p. 279–82.
15. Lu H, Plataniotis KN, Venetsanopoulos AN. Regularized common spatial patterns with generic learning for EEG signal classification. Proceedings of the 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society: Engineering the Future of Biomedicine, EMBC 2009. 2009. p. 6599–602.
16. Lu H, Eng HL, Guan C, Plataniotis KN, Venetsanopoulos AN. Regularized common spatial pattern with aggregation for EEG Classification in small-sample setting. IEEE Trans Biomed Eng. 2010;57(12):2936–46.
17. Lotte F, Guan C. Regularizing common spatial patterns to improve BCI designs: Unified theory and new algorithms. IEEE Trans Biomed Eng. 2011;58(2):355–62.
18. Sannelli C, Braun M, Müller KR. Improving BCI performance by task-related trial pruning. Neural Networks. 2009;22(9):1295–304.
19. Von Bünau P, Meinecke FC, Scholler S, Müller KR. Finding Stationary brain sources in EEG data. 2010 Annu Int Conf IEEE Eng Med Biol Soc EMBC'10. 2010;2810–3.
20. Kang H, Nam Y, Choi S. Composite common spatial pattern for subject-to-subject transfer. IEEE Signal Process Lett. 2009;16(8):683–6.

21. Lotte F, Guan C. Learning from other subjects helps reducing brain-computer interface calibration time. ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings. 2010. p. 614–7.

22. Analysis of Mathematical attack on RSA algorithm using factoring a prime number by Division algorithm method"A-ICTE sponsored National Conference Recent Research Trends in Network Security NCRTNS-2012 organized by Srinivas Institute of Technology, Valachil, Farangipet Post, Mangalore ,19th and 20th April 2012.

23. Samek W, Muller KR, Kawanabe M, Vidaurre C. Brain-computer interfacing in discriminative and stationary subspaces. Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS. 2012. p. 2873–6.

24. Samek W, Binder A, Muller KR. Multiple kernel learning for brain-computer interfacing. Conference proceedings : . Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Annual Conference. 2013. p. 7048–51.

25. Blankertz B, Kawanabe M, Tomioka R, Hohlefeld F, Müller K, Nikulin V V. Invariant common spatial patterns: Alleviating nonstationarities in brain-computer interfacing. Advances in neural information processing systems. 2007. p. 113–20.

26. Wojcikiewicz W, Vidaurre C, Kawanabe M. Stationary common spatial patterns: Towards robust classification of non-stationary EEG signals. ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings. 2011. p. 577–80.

27. Devlaminck D, Wyns B, Grosse-Wentrup M, Otte G, Santens P. Multisubject learning for common spatial patterns in motor-imagery BCI. Comput Intell Neurosci. 2011;2011(217987):1–9.

28. Wang H, Tang Q, Zheng W. L1-norm-based common spatial patterns. IEEE Trans Biomed Eng. 2012;59(3):653–62.

29. Li P, Xu P, Zhang R, Guo L, Yao D. L1 norm based common spatial patterns decomposition for scalp EEG BCI. Biomed Eng Online. 2013;12:77.

30. Wang H, Li X. Regularized Filters for L1-Norm-Based Common Spatial Patterns. IEEE Trans Neural Syst Rehabil Eng. 2016;24(2):201–11.

31. Li X, Lu X, Wang H. Robust common spatial patterns with sparsity. Biomed Signal Process Control. 2016;26:52–7.

32. Park J, Chung W. Common spatial patterns based on generalized norms. Brain-Computer Interface (BCI), 2013 Int Winter Work. 2013;39–42.

33. Wang H, Xu D. Comprehensive common spatial patterns with temporal structure information of EEG data: Minimizing nontask related EEG component. IEEE Trans Biomed Eng. 2012;59(9):2496–505.

34. Samek W, Meinecke FC, Muller KR. Transferring subspaces between subjects in brain - Computer interfacing. IEEE Trans Biomed Eng. 2013;60(8):2289–98.

35. Samek W, Vidaurre C, Uller K-R, Kawanabe M. Stationary common spatial patterns for brain–computer interfacing. J Neural Eng. 2012;9(2):26013–4.

36. Lemm S, Blankertz B, Curio G, Müller KR. Spatio-spectral filters for improving the classification of single trial EEG. IEEE Trans Biomed Eng. 2005;52(9):1541–8.

37. Dornhege G, Blankertz B, Krauledat M, Losch F, Curio G, Müller KR. Combined optimization of spatial and temporal filters for improving brain-computer interfacing. IEEE Trans Biomed Eng. 2006;

38. Tomioka R, Dornhege G, Nolte G. Optimizing Spectral Filters for Single Trial EEG Classification. Dagm. 2006;414–23.

39. Tomioka R, Dornhege G, Nolte G, Blankertz B, Aihara K, Muller K-R. Spectrally weighted Common Spatial Pattern algorithm for single trial EEG classification. Dept Math Eng Univ Tokyo Tokyo Japan Tech Rep. Citeseer; 2006 p. 25.

40. Wu W, Gao X, Hong B, Gao S. Classifying single-trial EEG during motor imagery by iterative spatio-spectral patterns learning (ISSPL). IEEE Trans Biomed Eng. 2008;55(September 2016):1733–43.

41. Novi Q, Guan C, Dat TH, Xue P. Sub-band common spatial pattern (SBCSP) for brain-computer interface. Proceedings of the 3rd International IEEE EMBS Conference on Neural Engineering. 2007. p. 204–7.

42. Mousavi EA, Maller JJ, Fitzgerald PB, Lithgow BJ. Wavelet Common Spatial Pattern in asynchronous offline brain computer interfaces. Biomed Signal Process Control. 2011;6(2):121–8.

43. Liu G, Huang G, Meng J, Zhu X. A frequency-weighted method combined with common spatial patterns for electroencephalogram classification in brain-computer interface. Biomed Signal Process Control. 2010;

44. Zheng HW and W. Local Temporal Common Spatial Patterns for Robust Single-Trial EEG Classification. IEEE Trans NEURAL Syst Rehabil Eng. 2008;16(2):131–9.

45. Wang H. Discriminant and adaptive extensions to local temporal common spatial patterns. Pattern Recognit Lett [Internet]. Elsevier B.V.; 2013;34(10):1125–9. Available from: http://dx.doi.org/10.1016/j.patrec.2013.03.014

46. Ang KK, Chin ZY, Zhang H, Guan C. Filter Bank Common Spatial Pattern (FBCSP) in Brain-Computer Interface. IEEE Int Jt Conf Neural Networks. 2008;2390–7.

47. Thomas KP, Guan C, Lau CT, Vinod AP, Ang KK. A New Discriminative Common Spatial Pattern Method for Motor Imagery Brain–Computer Interfaces. IEEE Trans Biomed Eng. 2009;56(11):2730–3.

48. Sun G, Hu J, Wu G. A Novel Frequency Band Selection Method for Common Spatial Pattern in Motor Imagery Based Brain Computer Interface. Proc 2010 IEEE Int Jt Conf Neural Networks. 2010;335–40.

49. Zhang H, Chin ZY, Ang KK, Guan C, Wang C. Optimum spatio-spectral filtering network for brain-computer interface. IEEE Trans Neural Networks. 2011;

50. Zhang H, Ang KK, Guan C, Wang C. Spatio-spectral feature selection based on robust mutual information estimate for Brain Computer Interfaces. Conf Proc IEEE Eng Med Biol Soc. 2009;2009:4978–81.

51. Ang KK, Chin ZY, Wang C, Guan C, Zhang H. Filter bank common spatial pattern algorithm on BCI competition IV datasets 2a and 2b. Front Neurosci. 2012;6(March):1–9.

52. Ang KK, Chin ZY, Zhang H, Guan C. Mutual information-based selection of optimal spatial-temporal patterns for single-trial EEG-based BCIs. Pattern Recognit. 2012;

53. Suk H Il, Lee SW. A novel bayesian framework for discriminative feature extraction in brain-computer interfaces. IEEE Trans Pattern Anal Mach Intell. 2013;35(2):286–99.

54. Obayashi M, Geshi T, Kuremoto T, Mabu S. An Optimization of Spatio-Spectral Filter Bank Design for EEG Classification.

55. Zhang Y, Zhou G, Jin J, Wang X, Cichocki A. Optimizing spatial patterns with sparse filter bands for motor-imagery based brain-computer interface. J Neurosci Methods. 2015;255:85–91.

56. Aghaei AS, Mahanta MS, Member S. Separable Common Spatio-Spectral Patterns for Motor Imagery BCI Systems. IEEE Trans Biomed Eng. 2016;63(1):15–29.

57. Fattahi D, Nasihatkon B, Boostani R. A general framework to estimate spatial and spatio-spectral filters for EEG signal classification. Neurocomputing. 2013;119:165–74.

58. Qi F, Li Y, Wu W. Spatio-temporally regularized common spatial patterns (STR-CSP) for single-trial EEG classification. Proceedings of the 33rd Chinese Control Conference, CCC. 2014. p. 7212–5.

59. Yang B, Li H, Wang Q, Zhang Y. Subject-based feature extraction by using fisher WPD-CSP in brain-computer interfaces. Comput Methods Programs Biomed [Internet]. Elsevier Ireland Ltd; 2016;129:21–8. Available from: http://dx.doi.org/10.1016/j.cmpb.2016.02.020

60. Krauledat M, Schro¨der M, Blankertz B M ller K. Reducing calibration time for brain-computer interfaces: A clustering approach. Adv Neural Inf Process Syst. 2007;19:753–60.

61. Krauledat M, Tangermann M, Blankertz B, Müller KR. Towards zero training for brain-computer interfacing. PLoS One. 2008;

62. Guger C, Guger C, Ramoser H, Pfurtscheller G. Real-Time EEG Analysis with Subject-Specific Real-Time EEG Analysis with Subject-Specific Spatial Patterns for a Brain – Computer Interface (BCI). 2001;8(4):447–56.

63. Wang Y, Gao S, Gao X. Common Spatial Pattern Method for Channel Selection in Motor Imagery Based Brain-computer Interface. Proc 2005 IEEE. 2005;5392–5.

64. Farquhar J, Hill NJ, Lal TN, Schölkopf B. REGULARISED CSP FOR SENSOR SELECTION IN BCI.

65. Yong X, Ward RK, Birch GE. Sparse spatial filter optimization for EEG channel reduction in brain-computer interface. ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings. 2008. p. 417–20.

66. Arvaneh M, Guan C, Ang KK, Quek C. Optimizing the channel selection and classification accuracy in EEG-based BCI. IEEE Trans Biomed Eng. 2011;

67. Arvaneh M, Guan C, Ang KK, Quek HC. Spatially sparsed common spatial pattern to improve BCI performance. ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings. 2011. p. 2412–5.

68. Lal TN, Schröder M, Hinterberger T, Weston J, Bogdan M, Birbaumer N, et al. Support vector channel selection in BCI. IEEE Trans Biomed Eng. 2004;51(6):1003–10.

69. Arvaneh M, Guan C, Ang KK, Quek C. Optimizing spatial filters by minimizing within-class dissimilarities in electroencephalogram-based brain-computer interface. IEEE Trans Neural Networks Learn Syst. 2013;24(4):610–9.

70. Arvaneh M, Guan C, Ang KK, Quek C. EEG data space adaptation to reduce intersession nonstationarity in brain-computer interface. Neural Comput. MIT Press; 2013;25(8):2146–71.

71. Samek W, Kawanabe M, Muller KR. Divergence-based framework for common spatial patterns algorithms. IEEE Rev Biomed Eng. Institute of Electrical and Electronics Engineers; 2014;7:50–72.

72. Samek W, Blythe D, Uller K-R, Kawanabe M. Robust Spatial Filtering with Beta Divergence. Adv Neural Inf Process Syst. 2013;1007–15.

73. Samek W, Muller KR. Tackling noise, artifacts and nonstationarity in BCI with robust divergences. 23rd European Signal Processing Conference, EUSIPCO 2015. 2015. p. 2741–5.

74. Brandl S, Muller KR, Samek W. Robust common spatial patterns based on Bhattacharyya distance and Gamma divergence. 3rd International Winter Conference on Brain-Computer Interface, BCI 2015. 2015. p. 1–4.

75. Dornhege G, Blankertz B, Curio G, Müller KR. Boosting bit rates in noninvasive EEG single-trial classifications by feature combination and multiclass paradigms. IEEE Trans Biomed Eng. 2004;

76. Grosse-Wentrup M, Buss M. Multiclass common spatial patterns and information theoretic feature extraction. IEEE Trans Biomed Eng. 2008;55(8):1991–2000.

77. Gouy-Pailler C, Congedo M, Jutten C, Brunner C, Pfurtscheller G. Model-based source separation for multi-class motor imagery. Signal Processing Conference, 2008 16th European IEEE. 2008. p. 1–5.

78. Grosse-Wentrup M, Liefhold C, Gramann K, Buss M. Beamforming in noninvasive brain-computer interfaces. IEEE Trans Biomed Eng. 2009;56(4):1209–19.

79. Sannelli C, Vidaurre C, Müller K-R, Blankertz B. CSP patches: an ensemble of optimized spatial filters. An evaluation study. J Neural Eng. 2011;8(2):025012.

80. Chin ZY, Ang KK, Wang C, Guan C, Zhang H. Multi-class Filter Bank common spatial pattern for four-class motor imagery BCI. Proc 31st Annu Int Conf IEEE Eng Med Biol Soc Eng Futur Biomed EMBC 2009. 2009;138632:571–4.

81. Cedric Gouy-Pailler, Marco Congedo, Clemens Brunner, Christian Jutten and GP. Nonstationary Brain Source Separation for Multiclass Motor Imagery. IEEE Trans Biomed Eng. 57(2):469–78.

82. Kang H, Choi S. Bayesian multi-task learning for common spatial patterns. Proceedings - International Workshop on Pattern Recognition in NeuroImaging, PRNI 2011. 2011. p. 61–4.

83. Wang H. Harmonic mean of Kullback-Leibler divergences for optimizing multi-class EEG spatio-temporal filters. Neural Process Lett. 2012;

84. Security in Wireless Sensor Networks: Key Management Module in EECBKM"Presented in International Conference on World Congress on Computing and Communication Technologies on Feb 27- & 28 and 1st march 2014, on St.Joseph college,Trichy

85. Wu W, Chen Z, Gao X, Li Y, Brown EN, Gao S. Probabilistic Common Spatial Patterns for Multichannel EEG Analysis. IEEE Trans Pattern Anal Mach Intell. 2015;37(3):639–53.

86. Zhang H, Yang H, Guan C. Bayesian learning for spatial filtering in an EEG-based brain–computer interface. Neural Networks Learn … [Internet]. 2013;24(7):1–12. Available from: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6484183\nhttp://www.ncbi.nlm.nih.gov/pubmed/24808520

87. Ge S, Wang R, Yu D. Classification of four-class motor imagery employing single-channel electroencephalography. PLoS One. 2014;9(6):e98019.