




Selection of Relevant Electrodes Based on Temporal Similarity for Classification of Motor Imagery Tasks

Jyoti Singh Kirar^(✉) , Ayesha Choudhary , and R.K. Agrawal 

Jawaharlal Nehru University, New Delhi, India
kirarjyoti@gmail.com, ayesha978@gmail.com, rkajnu@gmail.com

Abstract. Selection of relevant electrodes is of prime importance for developing efficient motor imagery Brain Computer Interface devices. In this paper, we propose a novel spectral clustering based on temporal similarity of electrodes to select a reduced set of relevant electrodes for classification of motor imagery tasks. Further, Stationary common spatial pattern method in conjunction with Composite kernel Support Vector Machine is utilized to develop a decision model. Experimental results demonstrate improvement in classification accuracy in comparison to variants of the common spatial pattern method on publicly available datasets. Friedman statistical test shows that the proposed method significantly outperformed the variants of the common spatial pattern method.

Keywords: Motor imagery · Brain computer interface · Common spatial pattern · Spectral clustering

1 Introduction

Brain computer interfaces (BCI) enable people with neurological disorders to establish communication and repair lost motor functions by transforming the brain signals into device commands. Non-invasive nature, low measurement cost and high resolution of EEG based BCI has favored its wide use for analysis of brain signals as compared to other modalities. EEG based BCI systems associated with motor imagery has received particular attention that involve visualizing movement of a specific motor part of the body [2]. Motor imagery BCIs use brain dynamics originating in primary sensorimotor area called sensorimotor rhythms (*mu* and *beta* rhythms), induced by execution or imagination of hand or leg movement, to translate EEG signals into device commands [2]. During motor imagination or execution, the amplitude of sensorimotor rhythms reduces, which is known as Event-Related Desynchronization (ERD). Increment in the amplitude of sensory-motor rhythms just after the motor imagination or execution is called as Event-Related Synchronization (ERS) [8].

Common spatial pattern (CSP) [2] is a well known spatial filtering method that finds subject specific data dependent spatial filters which maximizes variance of one class and minimizes variance of the other class simultaneously [2]. Extraction of information using the CSP is challenging as this method suffers from the small sample size (SSS) problem [4], i.e. the quantity of electrodes is high and the number of task specific EEG samples is less. In such circumstances, Eigen decomposition of the covariance matrix, whose dimension is number of electrodes \times number of electrodes, is highly computational and may lead to imprecision.

To reduce the number of electrodes, the research work [3] has utilized neurological information to segregate the electrodes placed on the brain on the basis of various anatomical areas of brain cortex and select all electrodes of relevant brain areas. However, this method utilizes topographical information based stationary division of electrodes. Further, this approach either uses all electrodes of a chosen cluster or removes all electrodes of a minimum significant cluster. There is a possibility that few of the chosen electrodes may not be pertinent or might be repetitive and few of the removed electrodes of the not chosen brain region might be applicable for recognizing two motor imagery tasks. Hence, in this work, the electrodes are partitioned into many clusters such that each cluster contains a set of similar electrodes to reduce the number of electrodes. For this we use the spectral graph clustering method [7], where the optimal number of clusters is decided using the Davies Bouldin (DB) index criterion [5]. Features are extracted from each cluster using stationary CSP (SCSP) method [6] and composite kernel based support vector machine (CKSVM) is utilised to build a decision model.

The major contributions of this paper include: (i) To select a subject specific reduced subset of relevant electrodes; (ii) To evaluate and compare the performance of the proposed method and variants of CSP on the publicly available datasets; (iii) Friedman test is used to demonstrate that the proposed method significantly outperforms variants of the CSP method. Rest of the article is organized as follows. Section 2 discusses the proposed work. Experimental results are discussed in Sect. 3 and finally, Sect. 4 concludes the article and provides future directions.

2 Proposed Method: Temporal Similarity Based Clustering in Conjunction with CKCSP (TSC-CKCSP)

The flow diagram of the proposed model is shown in Fig. 1. A brief description of each step is described as follows:

2.1 Spectral Graph Clustering Based Division of Electrodes

Spectral clustering is a graph theoretic approach to obtain clusters using spectral decomposition of a similarity matrix \mathbf{S} . To measure similarity between two

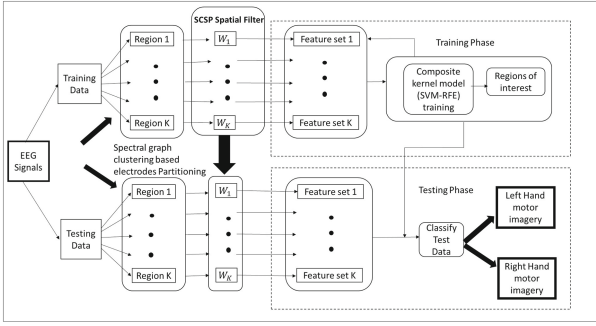


Fig. 1. Flow diagram of proposed model

samples, x_i and x_j , we have used Gaussian kernel which is defined as:

$$S(i, j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (1)$$

where σ is a tuning parameter. Let \mathbf{d} denotes the vector $d = [d_1, d_2, \dots, d_N]$ where $d_i = \sum_{j=1}^N S(i, j)$ measures similarity of i^{th} sample with all other samples. The larger value of d_i signifies more similarity of sample x_i with rest of the samples. Thus, the degree matrix \mathbf{D} is given by:

$$\mathbf{D}(i, j) = \begin{cases} d_i & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

The Laplacian matrix \mathbf{L} and normalized Laplacian matrix $\tilde{\mathbf{L}}$ are defined as:

$$\mathbf{L} = \mathbf{D} - \mathbf{S} \text{ and } \tilde{\mathbf{L}} = \mathbf{D}^{-\frac{1}{2}} \mathbf{L} \mathbf{D}^{-\frac{1}{2}} \quad (3)$$

Eigenvalues and eigenvectors of the normalized Laplacian matrix $\tilde{\mathbf{L}}$ is solved as:

$$\tilde{\mathbf{L}}\gamma = \delta\gamma \quad (4)$$

Let the Eigenvector γ_2 correspond to the second smallest Eigenvalue δ_2 . Binary partitioning of electrodes using Eigenvector γ_2 is done as follows:

$$Electrode_i = \begin{cases} Cluster 1 & \text{if } \gamma_{i,2} < 0 \\ Cluster 2 & \text{otherwise} \end{cases} \quad (5)$$

A given cluster is further recursively partitioned using spectral clustering method. To obtain an optimal number of clusters, a well-known Davies Bouldin (DB) criterion is used, which minimizes the ratio of within-cluster dispersion to between-cluster separation. The within-cluster dispersion for i^{th} cluster, Sw_i and between-cluster separation for cluster i and cluster j , $db_{i,j}$ are computed as:

$$Sw_i = \frac{1}{|C|} \sum_{x \in C_i} \{\|x - z_i\|\} \text{ and } db_{i,j} = \|z_i - z_j\| \quad (6)$$

where z_i is the centroid of the i^{th} cluster. The similarity measure between cluster i and cluster j and the corresponding DB index are given as:

$$M_{i,j} = \frac{Sw_i + Sw_j}{db_{i,j}} \text{ and } DB = \frac{1}{k} \sum_{i=1}^k M_i \quad (7)$$

where $M_i = \max_{i,j} M_{i,j}$ and k is the number of clusters. The number of clusters that provides minimal DB index value is considered optimal. The division of electrodes on both datasets is shown in Fig. 2.

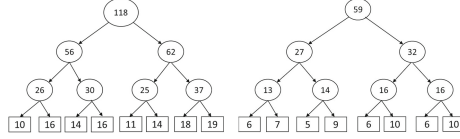


Fig. 2. Temporal Similarity based division of electrodes for Dataset 1 and Dataset 2

2.2 Stationary Common Spatial Patterns (SCSP)

Relevant features from each cluster of electrodes are computed using SCSP, which is evolved by introducing variations in the Rayleigh criterion function of the traditional CSP technique. Let Σ_1 and Σ_2 be the average covariance matrices of motor imagery left and right hand movement (class 1 and class 2), respectively and \mathbf{W} is a spatial filter matrix for a given cluster. The Rayleigh criterion maximization function using SCSP is given as:

$$R(\mathbf{w}) = \frac{\mathbf{w}^T \Sigma_1 \mathbf{w}}{\mathbf{w}^T (\Sigma_1 + \Sigma_2) \mathbf{w} + \beta \mathbf{P}(\mathbf{w})} \quad (8)$$

where $\mathbf{P}(\mathbf{W})$ is the introduced penalty term and β is a constant obtained after cross validation method. Further, the features computed from each cluster are transformed to a high dimensional Hilbert kernel space using a Gaussian kernel function to capture the non-linear relations of the extracted features as:

$$k_l(f_{i,l}, f_{j,l}) = \exp\left(-\frac{\|f_{i,l} - f_{j,l}\|}{2\sigma^2}\right) \quad (9)$$

where $f_{i,l}$ represents features from cluster l for trial i , obtained using SCSP method.

2.3 Feature Selection and Classification

In the proposed method, we have used CKSVM method for classification of motor imagery tasks. CKSVM considers the relevancy of an electrode cluster for recognition of motor imagery tasks.

$$\begin{cases} \max_{\alpha} -\frac{1}{2} \sum_{i,j} \alpha_i \alpha_j \gamma_i \gamma_j \sum_{l=1}^L k_l(f_{i,l}, f_{j,l}) + \sum_i \alpha_i \\ \text{s.t. } \sum_i \alpha_i \gamma_i = 0, 0 \leq \alpha_i \leq C, 1 \leq l \leq R, 1 \leq i \leq N \end{cases} \quad (10)$$

where α_i is the classifier parameter, C is a classifier regularization parameter, N is the number of samples and R is the number of clusters. Recursive Feature Elimination (RFE) is used to order the electrode cluster of a specific brain region by calculating the quadratic norm of each cluster of electrodes. Higher is the estimation of the quadratic norm, more important is that cluster to a motor imagery task. The cluster with lowest value of quadratic normal form is removed in each iteration till maximum classification accuracy is achieved.

3 Experimental Setup and Results

The experimental evaluation of the proposed work is performed on publicly available BCI Competition III dataset 4a (Dataset 1) and BCI competition IV dataset Ia (Dataset 2) [3]. Dataset 1 and Dataset 2 consist of motor imagery data acquired from five and seven healthy subjects respectively. The data captured for each trial belongs to the time window of 0.5–2.5 s after the onset of stimulus. Thus, from each electrode, 200 time units are utilized. Whole data is filtered using a [7–30 Hz] bandpass filter. SCSP penalty parameter $\beta = 0.1$ was used for all the experiments on both datasets, which is determined using cross-validation. Number of spatial patterns was fixed to $r = 1$. The average classification accuracy of the proposed method is reported in terms of 10 fold cross-validation run 10 times and is compared to CSP, SCSP and CKSCSP methods as shown in Tables 1 and 2 for subjects of Dataset 1 and Dataset 2, respectively.

Table 1. Comparison of TSC-CKCSP with existing methods in terms of average classification accuracy for Dataset 1.

Subject	CSP	SCSP	CKSCSP	TSC-CKCSP
aa	75.37	80.45	81.14	82.8052
al	97.73	94.38	98.34	96.91304
av	69.14	69.82	77.56	78.1667
aw	82.27	82.43	86.64	88.5667
ay	82.17	89.33	88.17	90.13
MEAN	81.34	83.28	86.37	87.3163

We can observe the following from Table 1: (i) The proposed method TSC-CKCSP achieves highest average classification accuracy for Dataset 1; (ii) An overall increment of 7.35%, 4.85%, and 1.1% in classification accuracy as compared to CSP, SCSP and CKSCSP is achieved with our proposed method TSC-CKCSP for Dataset 1. Similarly, the following can be observed from Table 2: (i) The proposed method TSC-CKCSP achieves highest average classification accuracy for Dataset 2 (ii) An overall increment of 12.08%, 3.54%, and 3.17% in classification accuracy as compared to CSP, SCSP and CKSCSP is achieved with our proposed method TSC-CKCSP. The spectral graph clustering is a data

Table 2. Comparison of TSC-CKCSP with existing methods in terms of average classification accuracy for Dataset 2.

Subject	CSP	SCSP	CKSCSP	TSC-CKCSP
ds1a	73.1	81.75	67.05	82.65
ds1b	65.4	59.95	71.45	71.89
ds1c	70.5	75.1	75.35	77.55
ds1d	76.8	89.2	90.3	90.55
ds1e	83.3	90.35	89.55	90.15
ds1f	82.8	86.9	88.3	87.35
ds1g	79	91.45	94.75	94.89
MEAN	75.84	82.1	82.39	85.0043

centric approach, hence provide relevant and subject specific clusters. Thus, the proposed method performs better than existing methods.

Table 3. The Friedman ranking obtained for each method.

Algorithm	Ranking
TSC-CKCSP	1.333
CKSCSP	2.166
SCSP	2.833
CSP	3.66

A non-parametric Friedman statistical test [1], is carried out to find the statistical difference between the proposed method and existing methods at significance value of $\alpha = 0.5$. Table 3 shows the Friedman ranking obtained for each method. P-value computed by Friedman Test is 1.5642 E-6, which signifies that the all methods under comparison are statistical significantly different from each other. Smaller value of Friedman ranking suggests proposed method outperforms variants of CSP.

4 Conclusion

An immense interest has been garnered by Motor imagery BCI due its wide applicability for communication. CSP is a widely used feature extraction technique for motor imagery BCI. However, it suffers from SSS problem due to numerous electrodes and smaller quantity of samples. The proposed method determines a reduced set of relevant electrodes using spectral clustering and CKSVM.

However, the proposed method uses only temporal information of data for division of electrodes and does not consider frequency information for clustering of electrodes. Thus, in future, we will incorporate both spatial, temporal and frequency information for obtaining a reduced and relevant subset of electrodes. Further, RFE in conjunction with SVM utilized in this work is computationally intensive, which requires improvement for real time-applications.

References

1. Iman, R.L., Davenport, J.M.: Approximations of the critical region of the fbietkan statistic. *Commun. Stat.-Theor. Methods* **9**(6), 571–595 (1980)
2. Kirar, J.S., Agrawal, R.: Optimal spatio-spectral variable size subbands filter for motor imagery brain computer interface. *Proc. Comput. Sci.* **84**, 14–21 (2016)
3. Kirar, J.S., Agrawal, R.: Composite kernel support vector machine based performance enhancement of brain computer interface in conjunction with spatial filter. *Biomed. Sig. Process. Control* **33**, 151–160 (2017)
4. Lu, H., Eng, H.L., Guan, C., Plataniotis, K.N., Venetsanopoulos, A.N.: Regularized common spatial pattern with aggregation for EEG classification in small-sample setting. *IEEE Trans. Biomed. Eng.* **57**(12), 2936–2946 (2010)
5. Maulik, U., Bandyopadhyay, S.: Performance evaluation of some clustering algorithms and validity indices. *IEEE Trans. Patt. Anal. Mach. Intell.* **24**(12), 1650–1654 (2002)
6. Samek, W., Vidaurre, C., Müller, K.R., Kawanabe, M.: Stationary common spatial patterns for brain-computer interfacing. *J. Neural Eng.* **9**(2), 026013 (2012)
7. Von Luxburg, U.: A tutorial on spectral clustering. *Stat. Comput.* **17**(4), 395–416 (2007)
8. Wolpaw, J., Wolpaw, E.W.: *Brain-Computer Interfaces: Principles and Practice*. OUP, USA (2012)