J. Basic. Appl. Sci. Res., 3(7)911-914, 2013
© 2013, TextRoad Publication

ISSN 2090-4304
Journal of Basic and Applied Scientific Research
www.textroad.com

EEG Feature Classification Using Enhanced CSSD and Combination of Classifiers

Atiq A. Tahir[1], Imran Usman[2], Zahid Mehmood[1], M. Farhan Shafiq[2], Tariq Bashir[1], Mahmood Ashraf Khan[2]

[1]Department of Electrical Engineering, COMSATS Institute of Information Technology, Park Road, ChakShahzad, Islamabad, Pakistan..
[2]Center for Advance Studies in Telecommunications, COMSATS Institute of Information Technology, Park Road, ChakShahzad, Islamabad, Pakistan.

ABSTRACT

This work presents a Modified Common Spatial Subspace Decomposition (CSSD) for electroencephalogram (EEG) feature extraction targeting the discrimination between left or right hand movement.In the preparation of self-paced key tapping for potential application in brain computer interface (BCI). Three classifiers, Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA) and Mehalanobis Discriminant Analysis (MDA), are used to test the accuracy of predicted results. Using data set IV of BCI Competition II, we achieved an accuracy of 95% as compared to earlier reported 86%.

KEYWORDS: Electroencephalography (EEG), Brain Computer Interface (BCI), Common Spatial Subspace Decomposition (CSSD), Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA).

I. INTRODUCTION

Brain Computer Interface provides a new communication paradigm between the human and machine. Its major utilization is for patients with acute motor disability to aid them in communicating with others in daily life using only their mental activities [1]. Since, electroencephalography (EEG) combines high temporal resolution, simple acquisition and low cost, it's recordings are mostly used in current BCI systems. Single-trial EEG associated with hand movement offers high classification accuracy, short response time, low rejection rate and a simple experimental approach thus, presenting an efficient method for use in practical scenarios [2].

Related Work

Wang et. al. [3] used common spatial subspace decomposition along with Fisher discriminant analysis to extract features from multi-channel EEG. A perceptron neural network was trained to classify the selected features. The classification accuracy achieved through this algorithm is 84% in BCI Competition II when applied to the data set IV of BCI Competition II. Dave et. al.[4]has explored parametric modeling strategies along with linear discriminant analysis to classify left/right self-paced typing exercise. They extended the autoregressive (AR) model with exogenous input model for EEG feature extraction. Data from six subjects was analyzed in this study,reporting an accuracy of 79.1±3.9% across subjects showing better results in comparison to AR method yielding a classification accuracy of 52.8±4.8%. Bashashatiet. al.[5] performed similar type of experiments consisting of 3-state self-paced BCI which was capable of detecting two different brain states (left and right hand movements). They compared the performance of BCI system using two different inputs i.e. monopolar and bipolar electrode setting. The reported average performance of the system in detecting motor activity is 54.7% with 70.25% accuracy in discriminating right and left hand movements using data from two healthy subjects. Lotteet. al. [6] has further improved this classification accuracy to 85 - 86% using the same dataset as in [3]. Their proposed algorithm is based on an inverse model that uses Fuzzy Region of Interest (FuRIA) and Support Vector Machine (SVM) classifier.

This work proposes an algorithm to classify single-trial EEG using CSSD for feature extraction and three classifiers LDA, QDA and MDA along with SVM giving an accuracy of 93%. In addition, significant reduction in complexity is achieved as compared to previous techniques.

[*]**Corresponding Author:** Imran Usman, Center for Advance Studies in Telecommunications, COMSATS Institute of Information Technology, Park Road, ChakShahzad, Islamabad, Pakistan. imran_usman@comsats.edu.pk

II. PROPOSED ENHANCED EEG CLASSIFICATION

Baseline EEG activity varies as a result of brain stimuli. Literature shows significant evidence of suppression or enhancement of certain frequency components in EEG signal prior to finger movement by a subject [8]. Such a suppression or enhancement in EEG activity is called Event Related Desynchronization (ERD) or Event Related Synchronization (ERS) respectively. These event-related potentials are highly frequency band dependent, not in-phase to the event, and different scalp sites can display ERD or ERS simultaneously. ERD has been used for the classification of a variety of tasks in BCI applications.

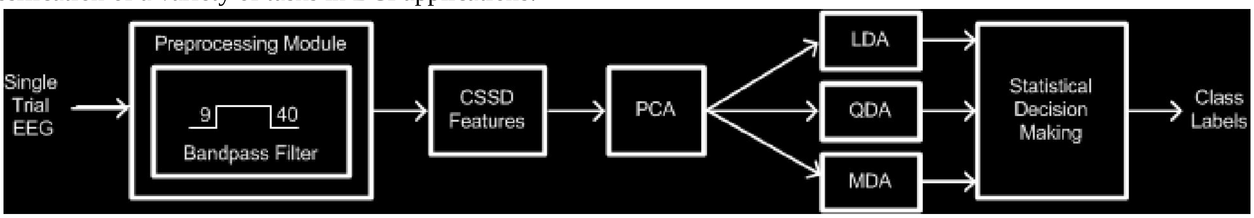

Figure1.Block diagram of the proposed technique

Basic block diagram of the proposed algorithm is shown in figure1. To extract pre-movement ERD activity from single-trial EEG signal, the EEG signal is first filtered using a zero phase, 4^{th} order Chebyshevb and pass filter with cutoff frequencies at 9 and 40 Hz to emphasize the rhythms related to self-paced finger movements. This band limited signal is then passed through a spatial filter designed using CSSD proposed by Wang et. al.[3]. After this, filtration variances of output signal components contain the most discriminative information with respect to left or right hand finger movements. The method of designing spatial filter using CSSD is discussed in forthcoming paragraphs.

Let the trials belonging to class Left (left hand finger movement) be represented by C_L and those belonging to class Right (right hand finger movement) by C_R. For the single-trial EEG sampleX^t (N channels by Tsamples) corresponding to thet^{th} trial in the dataset, the spatially filtered signal Y^t is given by,

$$Y^t = WX^t \tag{1}$$

where, W is the spatial filter designed using CSSD.

If we assume R^t to be the covariance matrix of a single trial, then R_L and R_R matrices, represent the cumulative effect of left and right class signals respectively. R_L and R_R are computed as follows:

$$R_L = \frac{1}{N_1}\sum_{t=1}^{N_1}\frac{R_L^t}{trace(R_L^t)} \tag{2}$$

and,

$$R_R = \frac{1}{N_2}\sum_{t=1}^{N_2}\frac{R_R^t}{trace(R_R^t)} \tag{3}$$

where R_L^t and $R_R^t$are the covariance matrices of t^{th} trial in class left C_L and right C_R, respectively; and N_1 and N_2 are the number of trials in C_L and C_R respectively. R is the cumulative sum of normalized covariance matrices R_L and R_R, which is given by.

$$R=R_L+R_R$$

If L is the matrix representing the Eigen-values of R at its diagonal and U is the matrix of Eigen-vectors of R, then using Eigen-value decomposition we can represent $R=ULU^T$. To normalize the statistics of the data, whitening transform matrix P is formulated as $P=L^{-1/2}U$. Further, the spatial filters for left and right class signals are constructed as $S_L=PR_LP^T$ and $S_R=PR_RP^T$ respectively. The most important property of S_L and S_R is that they share common eigen vectors. If L_L is eigen-value matrix of S_L and L_R is eigen-value matrix for S_R, then simultaneously diagonalizing $S_L$and S_R produces $S_L=BL_LB^T$ and $S_R=BL_RB^T$. Where matrix B represents the common Eigen-vectors. Hence, the spatial filter matrix W can be calculated as $W=B^TP$.

As given in (1), filtered EEG signal Y^t for t^{th} trial can be calculated by multiplying X^t with W. To compute the CSSD features $f^t[m]$, oft^{th} trial and m^{th} channel, log of normalized variances of$Y^t=\{y_L^t\}$, m=1,2…N, from (1), is taken. Normalized variances of channel m can be calculated as:

$$NVar_m^t = \frac{\mathrm{var}(y_m^t)}{\sum_{k=1}^{N}\mathrm{var}(y_k^t)} \tag{4}$$

And

$$f^t[m] = \log(\mathrm{NVar}_m^t) \quad (5)$$

where m=1…..N (number of channels). The N dimensional CSSD feature vector for a trial, t is denoted by $f^t[mt]$, m=1...N.

Equation 5 is applied for evaluating the coefficients for any given trial when mapped on the principal components. Principle component analysis is applied to reduce the dimensionality of the feature vector. For the purpose of classification we have used three classifiers, namely, LDA, QDA, and MDA, separately on the reduced features as shown in figure 1. The results of these classifiers are forwarded to the Statistical Decision Making module, whereby, cumulative probability of belonging to a specific class is used to assign the final class labels as seen in figure1.

III. Results and Discussions

Davies Bouldin Index (DBI) [9] is used to select the optimal window size and window start time. For the current investigation, the black-shaded box in figure2. represents the optimum values. For 28 channels EEG, 28 dimensional CSSD feature vectors are calculated by using equation5 for two classes. To reduce the dimensionality of the feature vector, Principal Componenet Analysis(PCA) is applied to 28 dimensional CSSD feature vectors for 208 trials from the training data which result in 28 principal components sorted according to their corresponding Eigen values. In order to find the number of most discriminating principal components, we apply the Wilcoxon Rank Sum Test [10] for all 28 PCA components. The hypothesis probability values for the first two principal components came out to be statistically significant ($p<0.001$) indicating their high discrimination power. All 28 dimensional CSSD feature vectors of training trials were projected onto these two principal components. These two dimensional co-efficient vectors were then used for classification using SVM classifier. A Gaussian SVM kernel was used in this work and its optimal bandwidth was selected as $\sigma = 6.9$. Bound on the Lagrange multipliers for the SVM Classifier, $L_{max} = 3\times10^7$, have been obtained through a 50 element cross validation dataset. The optimal separation boundary along with the 2 dimensional feature space is shown in figure2. A total of 316 trials were used for training while the remaining 100 trials were used for testing.

Window Start

Widow Size	2	6	10	14	18	22	26	30	34
16	4.62	4.71	4.19	3.90	3.72	3.53	3.34	3.55	3.55
18	4.54	4.30	4.02	3.66	3.54	3.48	3.36	3.52	
20	4.56	4.21	3.82	3.66	3.35	3.44	3.23	3.31	
22	4.20	4.02	3.74	3.55	3.41	3.39	3.26		
24	4.24	3.86	3.61	3.33	3.22	3.25	3.12		
26	4.07	3.67	3.44	3.31	3.25	3.27			
28	3.80	3.79	3.34	3.17	3.10	3.12			
30	3.76	3.48	3.31	3.22	3.13				
32	3.77	3.40	3.20	3.17	**2.92**				
34	3.51	3.39	3.18	3.16					
36	3.43	3.21	3.11	3.01					
38	3.42	3.27	3.09						
40	3.28	3.18	2.96						

Figure2.Davies Bouldin Index for different window sizes. Note that the shaded box represent the values selected.

The proposed technique has been evaluated on test data set, comprising of 100 trials, and achieved an accuracy of 90% which shows a significant improvement from the earlier reported accuracy of 85-86% [7]. Table 1 compares the accuracy of this work with other existing techniques in the literature.

Technique	Feature Dimension	Accuracy(%)
BCI Competition 2003-II [1]	192	84
FuRia with Fuzzufication [3]	5	85-86
Using Inverse Model [6]	27	83
CSSD+PCA+SVM (This Work)	2	90

Table1.Classification Accuracy for the Test Dataset

A significant reduction in complexity is achieved using lesser number of features, in contrast to the approach given by Wang et al. [4] which uses 192 features. PCA further reduces the number of features extracted to only two features per trial. This produces a simple two dimensional feature space allowing easier representation and visualization. Then 2-dimenssional features were classified for left and right class using a Support Vector Machine classifier with a Gaussian kernel whose classification boundary has been shown in figure 2. Improvement in the accuracy and reduction in the complexity of the feature space makes the proposed approach a promising candidate for online classification of self-paced finger tapping.

IV. CONCLUSION

This work proposes an enhanced CSSD based EEG feature extraction and classification scheme. The proposed technique classifies single-trial EEG signal during the preparation of self-paced key tapping. Using CSSD, features are extracted and their dimensionality is reduced by applying PCA. Three separate classifiers namely LDA, QDA, and MDA are used for classification purpose. Their cumulative probability of belonging to a specific class is then used to assign the final class labels. Improvement in the proposed work includes reduction in the number of features which results in significant decrease in computational complexity, while improving the accuracy of classification from earlier reported 86% to 95% using Data Set IV of BCI Competition II.

V. ACKNOWLEDGMENTS

The authors highly acknowledge the support provided by COMSATS Institute of Information Technology, Islamabad, Pakistan.

REFERENCES

[1] A. Kubler, B. Kotchoubey, J. Kaiser, J. R. Wolpaw, N. Birbaumer, “Brain-Computer Communication: Unlocking the Locked”, Psychological Bulletin, vol. 127, no. 3, pp 358-375, 2001.

[2] B. Blankertz, G. Curio, K. R. Müller, “Classifying single trial EEG: Towards brain computer interfacing” Advances in Neural Inf. Proc. Systems (NIPS 01), ed. T. G. Diettrich, S. Becker, Z. Ghahramani, pp. 157--164, vol. 14, 2002.

[3] Y. Wang, Z. Zhang, Y. Li, X. Gao, S. Gao, F. Yang, “BCI competition 2003-data set IV: An algorithm based on CSSD and FDA for classifying single-trial EEG”, Biomedical Engineering, vol. 51, no. 6, pp 1081--1086, 2004.

[4] D. P. Burke, S. P. Kelly, P. de Chazal, R. B. Reilly, C. Finucane, “A parametric feature extraction and classification strategy for brain-computer interfacing”, Neural Systems and Rehabilitation Engineering, vol. 13 1, pp 12--17, 2005.

[5] A. Bashashati, R. K. Ward, G. E. Birch, “Comparison of Using Mono-Polar and Bipolar Electroencephalogram (EEG) Electrodes for Detection of Right and Left Hand Movements in a Self-Paced Brain Computer Interface (BCI)”, Proc. Canadian Conference on Electrical and Computer Engineering, CCECE 2007, pp 725--728, 2007.

[6] F. Lotte, A. Lécuyer, B. Arnaldi, “FuRIA: A Novel Feature Extraction Algorithm for Brain-Computer Interfaces using Inverse Models and Fuzzy Regions of Interest”, Proc. 3rd International IEEE/EMBS Conference on Neural Engineering, CNE '07, pp 175--178, 2007.

[7] E. Niedermeyer, F. L. Da Silva, “Event-Related Potentials: Methodology and Quantification. In Electroencephalography: Basic Principals”, Clinical Applications and Related Fields, Williams and Wilkins, pp. 947--95, 2005.

[8] Y. Wang, P. Berg, M. Scherg, “Common spatial subspace decomposition applied to analysis of brain responses under multiple task conditions: A simulation study”, Clinical Neurophysiology, vol. 110, pp. 604--614, 1999.

[9] D. Davies, D. Bouldin, “A Cluster Separation Measure”, Pattern Analysis and Machine Intelligence”, vol. 1, no. 2, pp. 224-227, 1979.

[10] F. Wilcoxon, “Individual Comparisons by Ranking Method”, Biometrics, pp. 80-83, 1945.

[11] M. Congedo, F. Lotte, A. Lécuyer, “Classification of movement intention by spatially filtered electromagnetic inverse sloutions”, Physics in Medicine and Biology, vol 51 no. 8, pp. 1971-1989, 2006.