Performance Analysis of Multiclass Common Spatial Patterns in Brain-Computer Interface

Soumyadip Chatterjee[1], Saugat Bhattacharyya[1], Amit Konar[1], D.N. Tibarewala[2], Anwesha Khasnobish[2], and R. Janarthanan[3]

[1] Department of Electronics and Telecommunication Engineering
Jadavpur University, Kolkata, India
[2] School of Bioscience and Engineering
Jadavpur University, Kolkata, India
[3] Department of Computer Science
TJS Engineering College, Chennai, India
{soumyadipc7,biomed.ju,anweshakhasno}@gmail.com,
saugatbhattacharyya@live.com,konaramit@yahoo.co.in,srmjana_73@yahoo.com

Abstract. Brain-Computer Interfacing (BCI) aims to assist, enhance, or repair human cognitive or sensory-motor functions. The classification of EEG signals plays a crucial role in BCI implementation. In this paper we have implemented a multi-class CSP Mutual Information Feature Selection (MIFS) algorithm to classify our EEG data for three class Motor Imagery BCI and have presented a comparative study of different classification algorithms including k-nearest neighbor (kNN) and Fuzzy kNN algorithm, linear discriminant analysis (LDA), Quadratic discriminant analysis (QDA), support vector machine (SVM), radial basis function (RBF) SVM and Naive Bayesian (NB) classifiers algorithms. It is observed that Fuzzy kNN and kNN algorithm provides the highest classification accuracy of 92.65% and 92.29% which surpasses the classification accuracy of the other algorithms.

Keywords: Brain-Computer Interfacing, Electroencephalography, Common Spatial Pattern, Mutual Information Features Selection, k-Nearest Neighbor, Fuzzy k-Nearest Neighbor, Linear Discriminant Analysis, Quadratic Discriminant Analysis, Support Vector Machine, Nave-Bayesian.

1 Introduction

The main function of Brain-computer Interfacing (BCI) is to process and decode the brain signals and send the resulting commands to an external assistive device, thus implementing a real-time interface between the user and his environment. This interface may be a word processor, wheel chair or a prosthetic limb [1, 2]. In this technique, the subjects use their brain signals for communication and control of objects in their environment, thereby bypassing their impaired neuromuscular system [3, 4].

P. Maji et al. (Eds.): PReMI 2013, LNCS 8251, pp. 115–120, 2013.

Common Spatial Pattern (CSP) is instrumental in implementing extraction of intended activity from the neural recordings. CSP was first applied in BCI implementation in [5, 6]. BCI is generally limited to binary classification of data due to low information transfer rates. To enhance the information transfer rate one can move from binary to multiple classes. For this purpose, we have proposed a feature selection technique based on a simple multiclass CSP OVR [7] and Mutual Information Feature Selection (MIFS) [8] and have compared the performance of the proposed technique using seven different classification methods including k-Nearest Neighbor (kNN), Fuzzy kNN, Linear and Radial Basis Function (RBF-) Support Vector Machine (SVM), Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA) and Nave Bayesian (NB) classifiers [9 - 11] in differentiating the raw EEG data obtained into left/right hand and up movement. Our proposed framework also introduces a novel voting scheme to increase the classification accuracy.

The rest of the paper is structured as follows: Section 2 elucidates the proposed CSP-MIFS framework. The organization of the experimental data and data preprocessing is explained in section 3. Performance analysis of the classifiers is given in section 4. Section 5 concludes the paper.

2 Proposed Framework

The CSP algorithm was initially developed for binary classification of motor imagery. In this section we describe the binary CSP algorithm and extend its application for the multiclass case. The MIFS algorithm requires a user defined parameter k which denotes the number of features to be selected. It is based on the filter approach.

2.1 Proposed Approach of Multiclass CSP-MIFS

In this paper we consider the One-Versus-Rest (OVR) [7] approach to extend the CSP algorithm to multiple classes. As we are considering three classes, three CSP blocks have been employed. The input to the first CSP blocks will be the signals from class1 and a combination of class2 and class3 EEG data. Similarly, the input to the second classifier will be class2 and a combination of class3 and class1 EEG data and so on. Next, the CSP projection matrix for each of the 3 combinations will be computed and the spatially filtered matrix Z is created. The first 3 and last 3 rows are selected from Z and then they are subject to feature selection by MIFS algorithm [8].The spatially filtered and feature selected signals creates the feature vector to be fed to the classifiers. A comparative study of the classification accuracies of these algorithms are carried out. Finally, the classified data from each stage is processed by a voting mechanism which gives the final classes of the EEG data. This process is graphically shown in Fig. 1.

2.2 The Voting Mechanism

The input to the voting stage is the predicted classes of the data from the three classifiers. We denote the classes used in the first classifier as 1 and 23 where 1

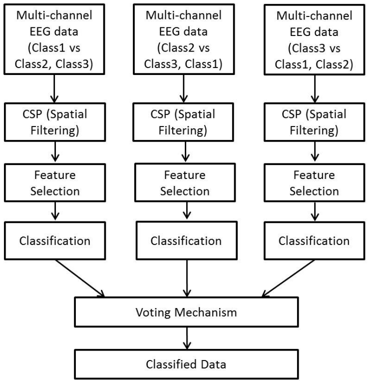

Fig. 1. Different stages of the proposed scheme for 3 classes

denotes the data is of class1 and 23 denotes the data is either of class2 or class3. Similarly the classes in the second classifier are denoted by 2 and 31 and for the third classifier as 3 and 12.c(x, y) is a function that computes the binary CSP between classes x and y and gives the predicted class. When there is an equal probability of a test data to belong to any one of two classes, then this function is called to resolve the matter. This happens in cases 2, 3 and 5. When the test data cannot be classified correctly to any of the classes, a random class is assigned to them. This is the case in 1 and 8. This is illustrated in Table 1.

Table 1. The Voting Mechanism

Sl. No.	**Classifier 1**	**Classifier 2**	**Classifier 3**	**Class**
1	1	2	3	rand(1,2,3)
2	1	2	12	c(1,2)
3	1	31	3	c(1,3)
4	1	31	12	1
5	23	2	3	c(2,3)
6	23	2	12	2
7	23	31	3	3
8	23	31	12	rand(1,2,3)

3 Data Analysis

All the experiments were conducted in our lab at Jadavpur University. 10 subjects (6 female and 4 male) performed the experiments in which they were instructed to imagine moving left (Class 1), right (Class 2) or forward (Class 3), according to the instructions displayed through a visual cue. The subjects performed the experiment in a single session, containing 120 trials each, i.e., 40 trials for each class.

Visual Cue. The visual cue is designed as follows: In the first 30 seconds of the session, a blank screen is displayed during which the baseline of the subject is measured followed by 60 trials of 6 seconds each. Each trial began with a fixation + for 1 second, which is an instruction to the subject to focus on the screen. Then a left/right/up arrow is displayed on the screen for 3 seconds as instruction to the subject. After 3 seconds, a blank screen would be displayed for 1.75-2.25 seconds to eliminate the cognitive effect of the current trial in the next one. The timing scheme of the visual cue is shown in Fig.2.

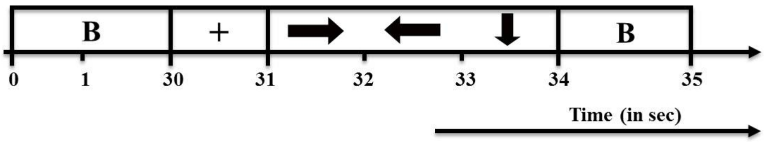

Fig. 2. Timing scheme of the visual cue displayed to the subjects

Experimental Setup. The EEG was recorded using an Emotiv Epoc system, which is a high resolution, multi-channel, wireless neuroheadset obtaining the EEG signals from the 14 electrode locations, based on the 10-20 electrode system. The electrode channels are AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8 and AF4. The sampling rate of the EEG system is 128 Hz.

Data Extraction and Pre-processing. Following the acquisition of the EEG signals, the raw EEG signal is band pass filtered using an IIR elliptical filter of order 6 between 8-24 Hz, as movement related signals are obtained from the 8-12 Hz mu- and 16-24 central beta band and to filter out any artifacts obtained during the recording. The filtered signal is further epoched into 1/16th of a second (i.e., 0.0625 seconds) and fed to the feature extraction and selection algorithm to form the feature vector.

Feature Extraction. Our proposed approach is applied to the epoched signal. First, the spatially filtered epoched signals are obtained and then fed to the MIFS feature selector to select the best features among the six rows. The final size of the features selected from each row is 4.

Classifiers. In this paper we have used both linear and non-linear classifiers which include k-Nearest Neighbor (kNN), Fuzzy kNN, Linear and Radial Basis Function (RBF-) Support Vector Machine (SVM), Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA) and Nave Bayesian (NB) classifiers [7-9].

4 Results and Discussion

The whole experiment is conducted in MATLAB version 7.9 environment. The specification of the system in which the experiment was conducted is as follows: Processor- Intel Core2Duo, 1.19 GHz, 3.2 GB RAM.

The total feature vector is partitioned into two different datasets, the training dataset and the testing dataset using k-fold cross-validation technique [10]. In our study, k is taken as 10. The feature vector is fed to the above classifiers and the classification accuracies are used as a parameter for performance analysis of the classifiers (as shown in Table 2). From Table 2, we observe that Fuzzy kNN gives the mean classification accuracy of 92.65% whereas kNN gives an accuracy of 92.29%.

Table 2. Average Classification Accuracy

Subjects	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**Mean**	**Std**
kNN	96.15	95.31	92.43	90.17	93.59	90.72	89.56	94.57	85.11	95.29	**92.29**	**3.43**
Fuzzy kNN	96.88	95.35	92.40	90.32	93.65	90.71	89.66	94.70	86.77	96.01	**92.65**	**3.24**
SVM	72.10	73.52	75.21	76.22	73.89	74.20	75.37	72.66	75.12	72.98	**74.13**	**1.34**
RBF-SVM	85.12	84.56	82.35	82.01	83.23	81.36	81.02	83.14	78.85	84.93	**82.66**	**1.97**
LDA	65.87	67.23	70.52	71.41	70.25	71.17	72.07	68.03	74.21	66.10	**69.69**	**2.76**
QDA	81.27	80.78	79.54	77.16	80.16	76.98	76.85	80.51	72.29	80.92	**78.65**	**2.82**
NB	85.07	83.54	81.29	78.96	80.23	77.23	75.20	81.29	70.23	84.19	**79.72**	**4.54**

We have also employed Friedman Test [12] to statistically validate our results. The significance level is set at α =0.05. The null hypothesis here, states that all the algorithms are equivalent, so their ranks should be equal. We consider the mean classification accuracy, obtained from Table 2 as the basis of rank. Table 3 provides the ranking of each classifier algorithm.

Table 3. Ranking of the classifiers based on their average classification accuracy

Classifier j	**kNN**	**Fuzzy kNN**	**SVM**	**RBF-SVM**	**LDA**	**QDA**	**NB**
Rank	2	1	6	3	7	5	4

Now, from Table 3, we obtain r_j, χ_F^2 = 79.291. Now, the χ_F^2 for our given dataset > $\chi^2_{7,0.05}$= 14.067. So, the null hypothesis, claiming that all the algorithms are equivalent, is wrong and, therefore, the performances of the algorithms are determined by their ranks only. It is clear from the table that the rank of Fuzzy kNN is 1, claiming Fuzzy kNN outperforms all the algorithms by Friedman Test.

5 Conclusion

The paper proposes a novel feature extraction and selection technique based on Multi-class Common Spatial Pattern and Mutual Information Feature Selection classification of multi-class problems. The resultant feature vector is fed to seven classifiers for a comparison on their performances. It is noted that Fuzzy kNN and kNN give the best results among all the classifiers and most of the classifiers give a result of more than 75%. Thus, our algorithm can be employed for further real time processing of multi-class problems. Further study in this direction will aim to optimize the feature selection, extraction and classification techniques to be implemented in real time application of Brain-Computer Interfacing.

Acknowledgments. I would like to thank University Grants Commission, India, University of Potential Excellence Programme (Phase II) in Cognitive Science, Jadavpur University and Council of Scientific and Industrial Research, India.

References

1. Lebedev, M.A., Nicolelis, M.A.: Brain-machine Interface: Past, Present and future. J. Trends Neurosci. 29(9), 536–546 (2006)
2. Vaughan, T.M., et al.: Brain-computer Interface Technology: A review of the second international meeting. IEEE Trans. Neural Sys. Rehab. Engg. 11(2), 94–109 (2003)
3. Daly, J.J., Wolpaw, J.R.: Brain-computer interfaces in neurological rehabilitation. Lancet Neurol. 7, 1032–1043 (2008)
4. Wolpaw, J.R., et al.: Brain-computer interfaces for Communication and Control. Clin. Neurophy. 113(6), 767–791 (2002)
5. Mller-Gerking, J., Pfurtscheller, G., Flyvbjerg, H.: Designing optimal spatial filters for single-trial EEG classification in a movement task. Clin. Neurophy. 110(5), 787–798 (1999)
6. Ramoser, H., Muller-Gerking, J., Pfurtscheller, G.: Optimal spatial filtering of single trial EEG during imagined hand movement. IEEE Trans. Rehab. Engg. 8(4), 441–446 (2000)
7. Wu, W., Gao, X., Gao, S.: One-versus-the-rest (OVR) algorithm: An extension of common spatial patterns (CSP) algorithm to multi-class case. In: 27th Annual International Conference of the Engineering in Medicine and Biology Society (IEEE-EMBS), pp. 2387–2390. IEEE Press, Shanghai (2006)
8. Battiti, R.: Using mutual information for selecting features in supervised neural net learning. IEEE Trans. Neural Net. 5(4), 537–550 (1994)
9. Lotte, F., Congedo, M., Lecuyer, A., Lamarche, F., Arnaldi, B.: A Review of Classification Algorithms for EEG-based Brain-Computer Interfaces. J. Neural Eng. 4, R1–R13 (2007)
10. Theodoridis, S., Koutroumbas, K.: Pattern Recognition. Academic Press (2006)
11. Keller, J.M., Gray, M.R., Givens, J.A.: A fuzzy k-nearest neighbor algorithm. IEEE Trans. Systems, Man and Cybernetics SMC 15(4), 580–585 (1985)
12. Conover, W.J., Iman, R.L.: Rank transformations as a bridge between parametric and nonparametric statistics. The American Statistician 35(3), 124–129 (1981)