

Performance Analysis of LDA, QDA and KNN Algorithms in Left-Right limb movement classification from EEG data

Saugat Bhattacharyya^{1,a}, Anwesh Khasnobish^{1,b}, Somsirsa Chatterjee¹, Amit Konar², D.N Tibarewala¹

¹School of Bioscience and Engineering, ²Department of Electronics and Telecommunication

Jadavpur University

Kolkata-700032, India

Email address: ^a saugatbhattacharyya@gmail.com, ^b anweshakhasno@gmail.com

Abstract— Brain Computer Interface (BCI) improve the lifestyle of the normal people by enhancing their performance levels. It also provides a way of communication for the disabled people with their surrounding who are otherwise unable to physically communicate. BCI can be used to control computers, robots, prosthetic devices and other assistive technologies for rehabilitation. The dataset used for this study has been obtained from the BCI competition II 2003 databank provided by the University of Technology, Graz. After pre-processing of the signals from their electrodes (C3 & C4), the wavelet coefficients, Power Spectral Density of the alpha and the central beta band and the average power of the respective bands have been employed as features for classification. In one of the approaches we fed all the extracted features individually and in the other approach we considered all features together and submitted them to LDA, QDA and KNN algorithms distinctly to classify left and right limb movement. The aim of this study is to analyze the performance of linear discriminant analysis (LDA), quadratic discriminant analysis (QDA) and K-nearest neighbor (KNN) algorithms in differentiating the raw EEG data obtained, into their associative movement, namely, left-right movement. Also the importance of the feature vectors selected is highlighted in this study. The total set to feature vector comprising all the features (i.e., wavelet coefficients, PSD and average band power estimate) performed better with the classifiers without much deviation in the classification accuracy, i.e., 80%, 80% and 75.71% with LDA, QDA and KNN respectively. Wavelet coefficients performed best with QDA classifier with an accuracy of 80%. PSD vector resulted in superior performance of 81.43% with both QDA and KNN. Average band power estimate vector showed highest accuracy of 84.29% with KNN algorithm. Our approach presented in this paper is quite simple, easy to execute and is validated robustly with a large dataset.

Keywords—BCI, ERS, ERD, wavelet coefficients, PSD, average band power estimates, LDA, QDA, KNN

I. INTRODUCTION

Controlling a computer or robotic device with thought only without any physical intervention is the principal idea behind brain computer interfaces (BCI). BCIs employ communication of the brain with outer environment that does not follow brain's conventional output pathways (i.e., through peripheral nerves and muscles). The brain activities for BCI

can be measured using EEG (electroencephalography), ECoG (electrocorticography), fNIRS (functional Near Infrared spectrosopes), fMRI (functional Magnetic Resonance Imaging), MEG (magneto encephalography), LPF (Local Potential Field) [3]. EEG based BCI is preferred as it is non-invasive, cost efficient, portable, easy-to-use and provides superior temporal resolution. BCIs not only improve the lifestyle of the normal people by enhancing their performance levels, it also provides a way of communication for the disabled people with their surrounding who are otherwise unable to physically communicate. BCI can be used to control computers, robots, prosthetic devices and other assistive technologies for rehabilitation. Capturing motor intention and executing the desired movement are the primary basis of brain-computer interfaces for neural prosthetics. They restore the motor ability or communication to impaired individuals by decoding the intentions of the individual. One of the main research areas of EEG based BCI for motor control is to decode the brain signals corresponding to particular limb movements [1], [2]. It is now fast becoming a new tool for communication, and can be used in sectors like robotics, mass communication, automobiles, games, entertainment, and the like.

The aim of this study is to analyze the performance of linear discriminant analysis (LDA), quadratic discriminant analysis (QDA) and K-nearest neighbor (KNN) algorithms in differentiating the raw EEG data obtained, into their associative movement, namely, left/right hand movement. Also the importance of the feature vectors selected is highlighted in this study. Filtering is performed on the EEG signals to make them free from noise, and subsequently feature extraction and classification are performed. The features considered in this paper include wavelet coefficients, average band power and power spectral density. The raw dataset has been de-noised by filtering, followed by feature extraction by wavelet transformation, band power estimation and power spectral density methods. In one of the approaches we fed all the extracted features individually and in the other approach we fed all the extracted features to LDA, QDA and KNN classifiers distinctly to classify left and right limb movement (Fig.1).

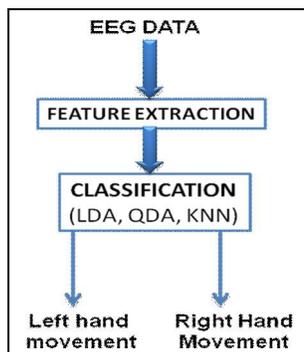


Figure 1. Block Diagram of the approach employed in the present study

During imagination or execution of body part movements, an event related synchronization (ERS) in the gamma band and an event related desynchronization (ERD) in the mu and beta band of the EEG originates in our brain. The gamma ERS and the mu-beta ERD occurs at the contralateral side of the brain near somatosensory and motor cortex area during particular limb movement. In case of ERS the power of the gamma component increases, while in case of ERD the power of the mu-beta component decreases. [4], [5]

The paper is divided into seven sections. The organization of the experimental data is explained in section II. The preprocessing of the data is introduced in section III. The feature extraction principle is briefly outlined in section IV. Motor imagery classification by LDA, QDA and KNN is explained in section V. Experimental results and their interpretations are given in section VI. Conclusions are listed in section VII.

II. EXPERIMENTAL DATA DESCRIPTION

The experimental data was obtained from BCI Competition 2003 provided by Department of Medical Informatics, Institute for Biomedical Engineering, University of Technology Graz. This dataset was recorded from a normal subject (female, 25 yr) during a feedback session where the subject was made to sit in a relaxing chair with armrests. The task was to control a feedback bar by means of imagery left-right movement in which the order of left and right cues were random. The recording was made using a G.tec amplifier and a Ag/AgCl electrode and three bipolar EEG channels were measured over C3, CZ and C4 electrode (Fig.2). The experiment consists of 7 sessions with 40 trials each conducted on the same day with several minutes break in between. In each trial, the first 2 seconds was quite. In the 2nd second an acoustic stimulus indicates the beginning of the trail with a fixation cross '+' displayed on the screen and at the 3rd second the visual cue (left-right arrow) is displayed. At the same time the subject was asked to move the bar in the direction of the cue as feedback. The feedback was based on AAR parameters of channel C3 and C4 and the AAR parameters were combined with a discriminant analysis into one output parameter (Fig.3). The EEG data was sampled at 128Hz.

I. PRE-PROCESSING OF THE EXPERIMENTAL DATA

A total of 280 trials were given of 9 second each. Out of the three electrodes used, C3 and C4 are selected for this study. CZ is left out because it is of little relevance for extracting information on left-right movement [6]. Thus, the total dataset comprised of $1152 \times 2 \times 280$ data. The trials for training and testing were selected randomly to prevent any systematic effect due to feedback. So, a total of 140 trials were selected for training and the rest 140 trials for test. As the visual cue started from $t=3$ sec to $t=9$ sec, thus, only the data for this time interval was selected. Now it is known that the brain electrical activities mainly occur in the 0.3-40Hz bands, and the higher frequencies can be considered as noise based on their environments and recording techniques. Thus a bandpass filter is used to filter in the frequency band: 0.5-30 Hz.

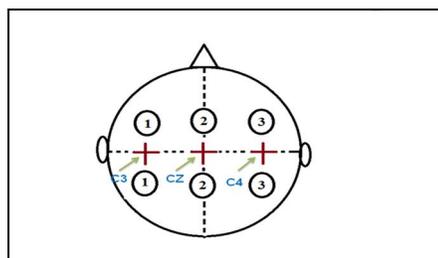


Figure 2. Electrode placement based on the experiment

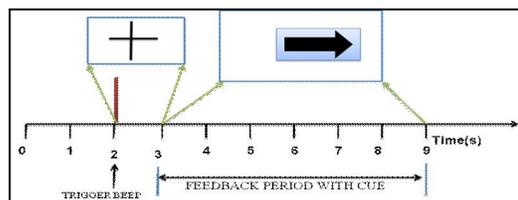


Figure 3. Timing scheme of the experiment

III. DESCRIPTION OF THE FEATURES EXTRACTED

A. Wavelet Features Feature Extraction

Wavelet transforms is a very effective way to extract features from an EEG signal [7, 8]. Their ability to discriminate both the temporal and spectral domain features of the signals makes them an important asset for EEG analysis. Also the wavelet transform do not suffer from the time-frequency trade off inherent in Short Time Fourier Transform (STFT) and Fourier Transform (FT) as their multi-scale approximation allows for effective localization of the signal with various spectral-temporal characteristics. Thus for a non-stationary signal like EEG, it is an effective analysis tool. The discrete wavelet transforms analyzes the signals at different resolutions by decomposing the signal into coarse approximation and detail information. Each level comprises of two digital filters and two down-samplers by 2. The down-sampled outputs of the first high-pass and low-pass filters provides the detail D1 and approximation A1, respectively. The first approximation is

further decomposed and the process continued, until the desired result is obtained. [9, 10].

In the present study, Daubechies (db) mother wavelet of order 4 is used. After trials with the EEG data, the D3 features i.e., the third level coefficient for the respective electrodes were selected as one of the feature components for the final feature vector. Figure 4 and 5 shows the wavelet decomposition for left-right imagery for C3 and C4 electrode.

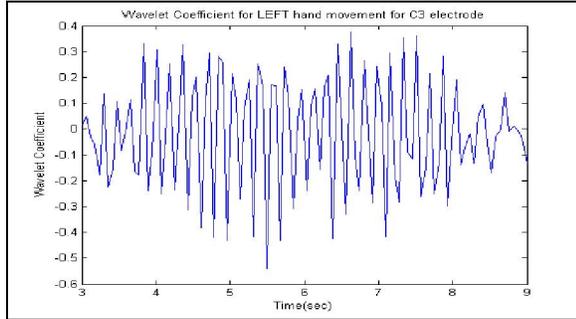


Figure 4a. Wavelet Coefficient for Left movement for C3 electrode

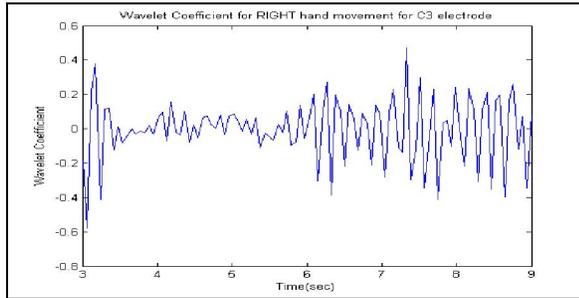


Figure 4b. Wavelet Coefficient for Left movement for C3 electrode

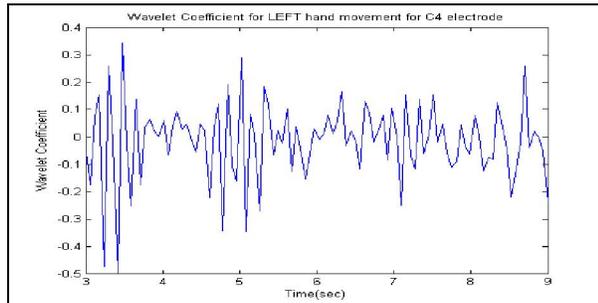


Figure 5a. Wavelet Coefficient for Left movement for C4 electrode

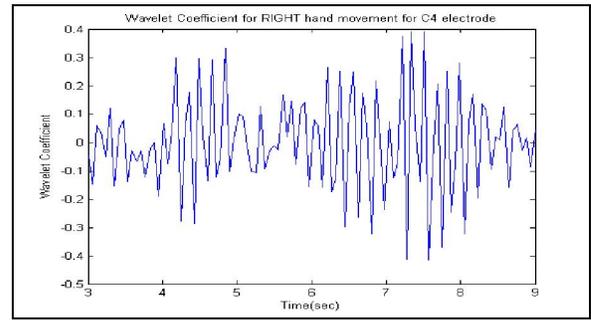


Figure 5b. Wavelet Coefficient for Left movement for C4 electrode

B. Spectral Estimation Method

Spectral density methods extract information from a signal to describe the distribution of its power in the frequency domain. The power spectral density (PSD) is defined as the Fourier transform (FT) of the signal's autocorrelation function provided that the signal is stationary in a wide sense [10]. Thus for an EEG signal segmenting the complete time series data would be an ideal approach.

For this paper, the Welch approach was applied along with a Hamming window of length 64. The Welch method divides the times series data into overlapping segments, computing a modified periodogram of each segment and then the PSD estimates is averaged. The PSD estimates were obtained for two frequency bands, namely the alpha or mu band (8-12Hz) and the central beta band (18-25Hz) for each respective electrode. Also the average power was obtained for each band. Then the difference of the PSD estimates (formula 1) and average power (formula 2) is selected as another feature for this study.

$$F_{PSD} = \sum_{f=a}^b Psd_{C4}(f) - \sum_{f=a}^b Psd_{C3}(f) \quad (1)$$

$$F_{POW} = Pow_{C4} - Pow_{C3} \quad (2)$$

Where $Psd_{C3/C4}$ is the PSD estimates of the respective electrodes in $[a, b]$, where a & b is the frequency range (8-12Hz for alpha band and 18-25 for beta band), $Pow_{C4/C3}$ is the average power of the respective electrodes in $[a, b]$.

Figure 6 and 7 shows the PSD estimates for left-right imagery for C3 and C4 electrodes, for both alpha and beta band. Figure 8 shows the PSD estimates for left-right imagery for the difference of the two electrodes.

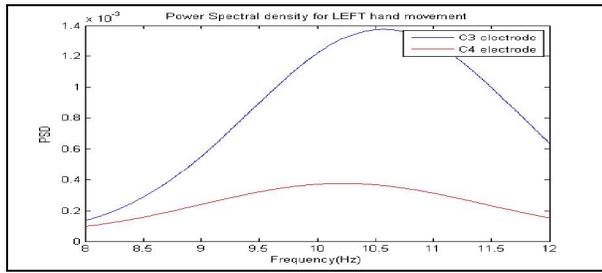


Figure 6a.

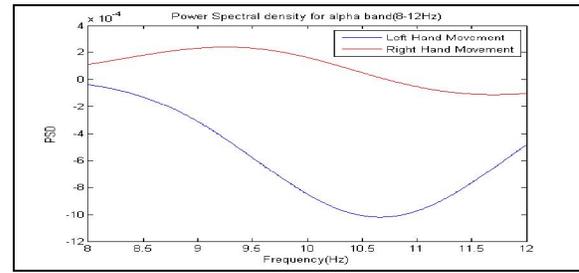


Figure 8a

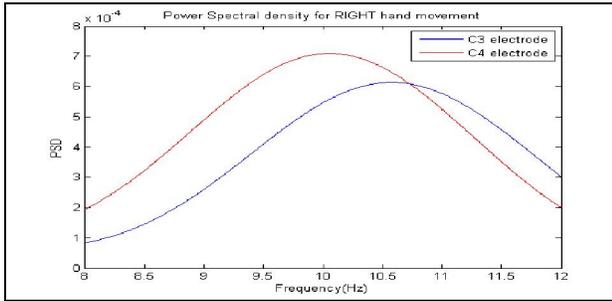


Figure 6b

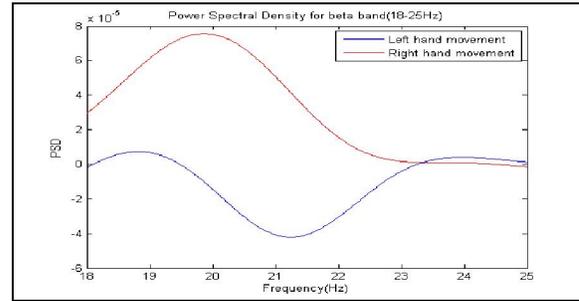


Figure 8b

Figure 6. Alpha Band PSD estimate for a) left movement b) right movement

Figure 8. Power Spectral Density of a) alpha band b) beta band for the difference of the two electrodes C3 and C4 for left/right hand movement

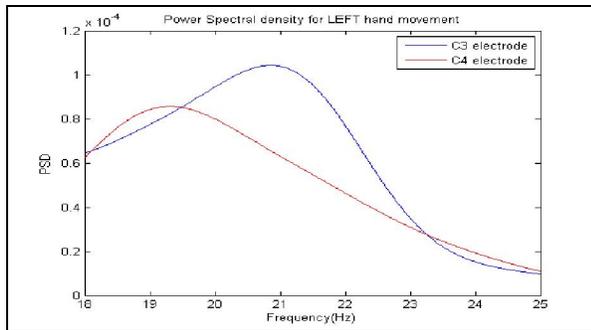


Figure 7a

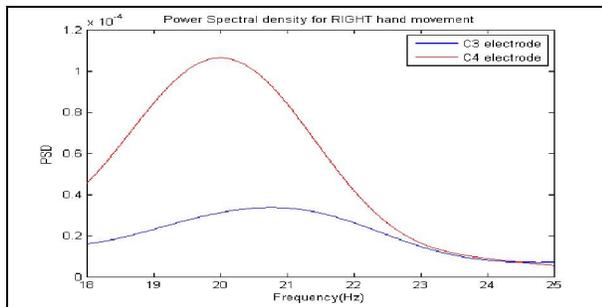


Figure 7b.

Figure 7. Beta Band PSD estimate for a) left movement b) right movement

Table 1 gives an overview of the feature vectors taken.

TABLE I. FEATURE VECTORS WITH THEIR RESPECTIVE SIZE

| FEATURE VECTORS | SIZE (No. of Features per Trial \times No. of Trials) |
|--------------------------|---|
| Wavelet Coefficient (D3) | 204×140 |
| Alpha band PSD estimates | 768×140 |
| Alpha Band Average Power | 1×140 |
| Beta Band PSD estimates | 768×140 |
| Beta Band Average Power | 1×140 |
| Total | 1742×140 |

IV. LIMB MOVEMENT CLASSIFICATION FROM MOTOR IMAGERY

One of the main prerequisites for EEG based BCI to control prosthetic or assistive devices are to classify the EEG data into its corresponding particular limb movement. Linear Discriminant Analysis (LDA), Support Vector Machines (SVM), Bayesian methods are few of the common classifiers used for BCI research. In this paper, LDA, Quadratic Analysis (QDA) and K- Nearest Neighbor (KNN) are performed on each of the single vectors and the complete feature set.

LDA is a Bayes optimal classifier provided the distribution of features in each of two classes is normal with the same covariance matrix [10]. The LDA finds a one-dimensional subspace in which the classes are usually well separated by a linear separating hyper plane. The discriminant function is given by

$$d_k(x) = 2\mu_k^T \sum_k^{-1} X - \mu_k^T \sum_k^{-1} \mu_k - 2\log\pi(k). \quad (3)$$

Where, k =class, X is the set of measurements, μ_k is the mean vector, π_k is the prior probability and \sum_k is the covariance matrix.

QDA is a generalized version of LDA, provided there are only two classes of points and the measurements are normally distributed. However unlike LDA, the assumption that the covariance of each class is identical is not taken into consideration in QDA. Further, the surface that separates the subspaces will be a conic section (like parabola, hyperbola, etc.). When the equation III is multiplied by -2, the discriminant function is given by

$$d_k(X) = (X - \mu_k)^T \sum_k^{-1} (X - \mu_k) + \log|\sum_k| - 2\log\pi_k \quad (4)$$

And the discriminant rule is given by

$$d_k(X) = \min_{1 \leq k \leq K} d_k(X) \Leftrightarrow \max_{1 \leq k \leq K} p(k/x) \quad (5)$$

where, $p(k/x)$ =posterior distribution. Using this rule is called the QDA [11].

The k-nearest neighbors (KNN) algorithm finds the K-Nearest Neighbors among the training set, and the categories of the k-nearest neighbors are used to weigh the category candidates. The performance of this algorithm greatly depends on two factors: a suitable similarity function and an appropriate value for k. if k is too large, big classes will overwhelm the small classes whereas if k is too small, the advantage of KNN algorithm is not exhibited. Equation (6) and (7) is the widely used strategies for this algorithm [12].

$$y(d_i) = \arg \max_k \sum_{x_j \in kNN} y(x_j, c_k) \quad (6)$$

$$y(d_i) = \arg \max_k \sum_{x_j \in kNN} Sim(d_i, x_j) y(x_j, c_k) \quad (7)$$

where, d_i is a test document, x_j belongs to class c_k , $Sim(d_i, x_j)$ is the similarity function for d_i and x_j . Equation (1) shows that the prediction will be the class that has the largest number of members in the k nearest neighbor. Equation (2) means the class with maximal sum of similarity will be the best result.

V. RESULTS AND DISCUSSIONS

The test data was used for validation of the classifiers. The true labels of the test data were obtained from the website of the BCI competition 2003. All of the data were bandpass filtered between the frequency ranges of 0.5-30Hz. From the two electrodes of interest, namely, C3 and C4, wavelet coefficients, PSD estimates for the alpha and beta bands and their corresponding powers were selected as the features for this study. The feature vectors have been validated using paired t-test and their respective probability of the occurrence of Type I and Type II error are shown in Table II. Each single feature vector and the complete feature set were fed into LDA, QDA and KNN classifiers separately in a MATLAB environment. The results of the classification are shown in Table II. The error in Table II gives the misclassification error while training the dataset and the accuracy is obtained when the test data is fed to the trained classifier.

Fig 7 illustrates the accuracy with different number of features for the three classifiers. It is seen that the accuracy is almost the same for the different number of features taken. It is also observed from Table II that when only the wavelet coefficient feature vector is used, it gave poor classification accuracy with the classifiers (i.e., LDA, QDA and KNN) due to its complete non linearity. Wavelet coefficients classification with QDA showed highest accuracy of 80%. The power spectral density estimate feature vector showed higher classification accuracy with respect to wavelet coefficients and average band power estimates. The full feature vector set comprising all the extracted features with greater dimensionality indicated higher performance accuracy of 80%, 80% and 75.71% with LDA, QDA and KNN respectively. LDA showed better classification with PSD vector and complete feature set with an accuracy of 80%. QDA performed better with PSD vector with an accuracy of 81.43%. KNN showed highest performance with average band power estimate vector with an accuracy of 84.29%.

TABLE II. RESULT OF CLASSIFICATION WITH THE SELECTED FEATURES

| FEATURES | TYPE I ERROR | TYPE II ERROR | CLASSIFIER | ERROR (IN %) | ACCURACY (IN %) |
|------------------------|--------------|---------------|------------|--------------|-----------------|
| Wavelet Coefficient | 0.05 | 0.1943 | LDA | 35.71 | 48.57 |
| | | | QDA | 8.57 | 80 |
| | | | KNN | 34.29 | 65.71 |
| Power Spectral Density | 0.05 | 0.0223 | LDA | 20 | 80 |
| | | | QDA | 20.71 | 81.43 |
| | | | KNN | 13.57 | 81.43 |
| Average Band Power | 0.05 | 0.0245 | LDA | 19.29 | 78.57 |
| | | | QDA | 19.29 | 77.86 |
| | | | KNN | 12.86 | 84.29 |
| All | 0.05 | 0.0008 | LDA | 19.29 | 80 |
| | | | QDA | 20.71 | 80 |
| | | | KNN | 12.14 | 75.71 |

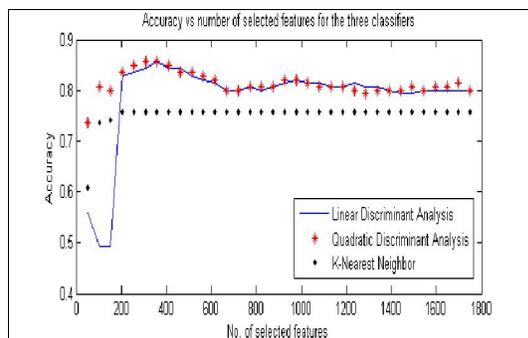


Figure 7. Beta Band PSD estimate for a) left movement b) right movement

VI. CONCLUSION

In this paper, features are extracted from the preprocessed EEG signal and fed to the motor imagery classifiers for differentiating the EEG signal to its corresponding left-right limb movement. Wavelet transform, power spectral density estimate and average band power estimates are techniques followed in this study for feature extraction. In one of the approaches we fed all the extracted features individually and in another approach we formed a feature vector and fed it to LDA, QDA and KNN algorithms distinctly to classify left and right limb movement. It is evident from the results that due to the non-linearity of the wavelet coefficients it contributed to poor classification accuracy when used individually. When each feature vector is fed for classification, PSD showed highest accuracy than the rest feature vectors. The total set to feature vector comprising all the features (i.e., wavelet coefficients, PSD and average band power estimate) performed better with the classifiers without much deviation in the classification accuracy. A lot of the classification depends on the process of the feature vectors selected and the parameters that define these vectors. The processing of the features requires further validation and study to improve the accuracy of the classifiers. Also, it is held that the combination of feature vector is vital for proper classification, thus newer features need to be tried out to further improve the classification of left-right motor imagery. Our approach of feature extraction and classification presented here is very simple and robust. To control EEG based BCI devices it is required to find out more relevant features with less computational time and with higher computational efficiency.

Future study in this direction will aim at techniques for optimizing feature selection, extraction and classification methodologies to be implemented in online classification of EEG data for BCI research.

REFERENCES

- [1] Schwartz A.B., Cui X.T., Weber D.J., Moran D.W. "Brain Controlled Interfaces: Movement Restoration using Neural Prosthetics." *Neuron* vol.52, October 2006, pp. 205-220
- [2] Lebedev M.A., Nicolelis, "Brain-machine interface: Past, present and future", *Trends Neurosci.* Vol. 29(9), September 2006, pp.536-546
- [3] Anderson R.A., Musallam S., Pesaran B., "Selecting the signals for a brain-machine interface", *Curr Opin Neurobiol* Vol.14 (6), December 2004, pp.720-726
- [4] Matsunaga T., Katayama Y., Hayami T., Iramina K. "Measurement of mu/beta ERD and gamma ERS during the imagination of body parts movement." 30th Annual International IEEE EMBS Conference Vancouver Canada, August 2008
- [5] Hema C.R., Paulraj M.P., Yaacob S., Adom A.H., Nagarajan R. "Recognition of motor imagery of hand movements for a BMI using PCA features." 2008 International Conference on Electronic Design, Penang, Malaysia, December 2008
- [6] Xu Huaiyu, Lou Jian, Su Ruidan, Zhang Erpang "Feature Extraction and Classification of EEG for imaging left-right hands movement."
- [7] Schloegl A., "Dynamic spectral analysis based on an autoregressive model with time-varying coefficients", *IEEE-EMBC and CMBEC*, 1995
- [8] Darvishi S., Al-Ani A. "Brain-computer interface analysis using continuous wavelet transform and adaptive neuro-fuzzy classifier", *Proc. 29th Int. Annu. Conf. IEEE Eng. Med. Biol. Soc.*, August 2007, pp. 3220-3223
- [9] Xu Q., Zhou H., Wang Y., Huang J. "Fuzzy support vector machine for classification of EEG signals using wavelet based features." *Medical Engineering & Physics* 31, 2009, pp. 858-865
- [10] Herman P., Prasad G., McGinnity T.M., Coyle D. "Comparative analysis of spectral approaches to feature extraction for EEG-based motor imagery classification." *IEEE Trans. Neural sys. Rehab eng.* 16(4), August 2008, pp. 317-326
- [11] Daniella Birkel, "Regularized Discriminant Analysis"
- [12] Baoli L., Shiwen Y., Qin L. "An improved K- nearest neighbor algorithm for text categorization." *Proceedings of the 20th International conference on computer processing of Oriental Languages, Sheyang, China, 2003.*