Classification of Multichannel EEG Signal by Linear Discriminant Analysis

Mohammad Rubaiyat Hasan, Muhammad Ibn Ibrahimy, S. M. A. Motakabber, and Shahjahan Shahid

1 Introduction

By using the classification algorithm for EEG signal it becomes easy to find out performance of Brain computer interface(BCI). BCI causes direct operation between brain and computer. Studies showed that a person with severe neuromuscular disabilities can learn to use a BCI system by modulating the various features in EEG signal [1]. The efficiency of a BCI depends on 3 operations. They are: signal recording; feature extraction from the recorded signal and classification of the extracted information [2].

The output of the feature extraction unit highly impacts on the performance of the feature classification unit. The probability of correctness identification can be increased if the feature extraction unit transforms the EEG signal in such a way that the signal to noise ratio (SNR) can be maximized as much as possible [3]. This paper presents Linear Discriminate Analysis (LDA), a signal classification algorithm for a MI BCI. For classify EEG signal, it has been used the signal recorded from the motor cortex area while a subject performs the imagination of a motor movement.

The rest of the paper is organized as follows. In Section 2, the materials and methods applied here are mentioned. Performance measurements are discussed in Section 3 and results have been showed in section 4. Finally, section 5 concluded the paper.

M.R. Hasan (✉) • M.I. Ibrahimy • S.M.A. Motakabber
Dept. of Electrical and Computer Engineering, International Islamic University, Gombak, Malaysia
e-mail: mdrubaiyat@yahoo.com; Ibrahimy@iium.edu.my; amotakabber@iium.edu.my

S. Shahid
Dept. of Computing Science, University of Glasgow, Scotland, UK
e-mail: shahjahan.shahid@gmail.com

2 Materials and Methods

In first part the feature extraction has been done by Power spectral density (PSD) for 22 different channel data. Two different frequency bands are calculated for those 22 EEG channels. Then the LDA classifier is used to get more accurate results. The proposed technique has been used to devise to an MI related BCI which has been evaluated with the data provided by the Graz BCI lab as part of the BCI competition IV data-2a. To validate accuracy measurement of accuracy and Cohen's kappa are used in this paper. The accuracy results of the classification are shown in Sections 4.

2.1 Data Selection

In research work, the data set consists of EEG data from 9 subjects. The BCI paradigm consisted of 2 different motor imagery tasks. They are: the imagination of movement of the left hand (class 1) and right hand (class 2). For each subject the two sessions were recorded on different times. Each session is comprised of 6 runs separated by short breaks. One run consists of 24 trials (12 for each of the two possible classes), yielding a total of 144 trials per session. For each session, at the beginning a recording of approximately 5 minutes was performed to estimate the EOG influence. The recording was divided into 3 blocks: two minutes with eyes open (looking at a fixation cross on the screen), one minute with eyes closed and one minute with eye movements.

Fig. 1 shows the timing scheme of paradigm. In this paradigm 0-6 seconds time for one session. Around 2 seconds break makes a total of 8 seconds time for each session. First 3 seconds for for fixation and maintaining the cue then 3-6 seconds the potential time for recording the MI based EEG signal.

H. Selvaraj et al. (eds.), *Progress in Systems Engineering: Proceedings of the Twenty-Third International Conference on Systems Engineering*, Advances in Intelligent Systems and Computing 1089,
DOI 10.1007/978-3-319-08422-0_42,

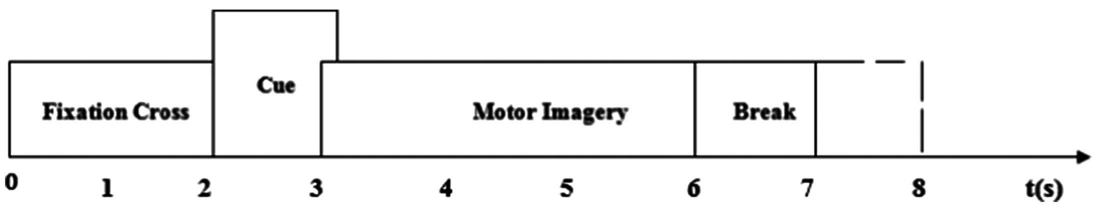

Fig. 1 Timing scheme of the paradigm [4].

2.2 Feature Extraction

PSD shows the strength of the variations as a function of frequency. In other words, it shows at which frequencies variations are strong and at which frequencies variations are weak. The unit of PSD is energy per frequency (width) and you can obtain energy within a specific frequency range by integrating PSD within that frequency range. Computation of PSD is done directly by the method called FFT or computing autocorrelation function and then transforming it.

2.3 Classification

Classification is done by LDA. The aim of is to use hyperplanes to separate the data representing the different classes. LDA assumes normal distribution of the data, with equal covariance matrix for both classes. The separating hyperplane is obtained by seeking the projection that maximize the distance between the two classes means and minimize the interclass variance. To solve an N-class problem (N > 2) several hyperplanes are used in LDA. It can be defined by the equation, which is maximized over all linear projections, w:

$$J(w) = |\mathrm{m}_1 - \mathrm{m}_2|^2 / (S_1^2 + \mathrm{S}_2^2) \tag{1}$$

Here, m represents the mean, S represents a variance, and the subscripts denote the two classes [5]. Limitation of LDA is that, for nonlinear classification it does not provide good performance always. But the high-dimensional and noisy nature of EEG often limits the advantage of nonlinear classification methods [6].

3 Performance Measures

Performance of a BCI system is measured by percentage of accuracy[7].The performance have been measured to see evaluate the classification results of LDA for those data. This is done by percentage of left and right accuracy measurement following the detection of accuracy and Cohen's kappa

3.1 Accuracy

The accuracies have been computed for each instant of data by the classifier's output (estimated label = the sign of d(m), where right means positive and left means negative) is compared with actual left or right of MI to prepare a confusion matrix (CM). Using the CM, the left and right accuracies for each instant of the paradigm are computed by the following formulae:

$$\text{Left accuracy} = \frac{(\text{finally obtained negative in CM} \times 100)}{(\text{total number of input as left})} \tag{2}$$

$$\text{Rightaccuracy} = \frac{(\text{finally obtained positive in CM} \times 100)}{(\text{total number of input as right})} \tag{3}$$

Within the paradigm a total of 50 equi-spaced points were considered from the 3-8 (5 seconds). Considering all trials and their actual and estimated labels (left and right), 50 CM have been made; where positive means both actual and estimated labels are right and they match; negative means both actual and estimated label are left and they match. The left and right accuracies are then computed using equation (1) for each time point of the paradigm. The mean of left and right hand MI accuracy was called here as the overall accuracy.

3.2 Cohen's Kappa

Cohen's kappa is a statistical measurement. It provides an index of interrater reliability. It is an improvement over using the percent of accuracy, as the procedure of computing accuracy does not involve the false positive or false negative effects. The computation of kappa at each instance starts from the CM prepared by comparing the appearance of two raters: the actual events and the estimated events (observed at classifier's output). From the definition, Cohen's kappa can be written as,

$$k = \frac{(P_0 - \mathrm{P_c})}{(1 - P_c)} \tag{4}$$

where Po means the relative observed agreement between raters, and Pc for the hypothetical probability of chance agreement. The maximum possible value of Cohen's kappa is limited to 1 and then the raters are in complete agreement. If there is no agreement among the raters, k = 0 [8].

4 Results and Discussion

The study presented the aim of better performance after using LDA. Table 1 shows the performances of PSD based BCI after applying LDA. These results are obtained upon applying the training and evaluation signals (from 9 subjects) to the respective BCI. Here, column in the table provides the maximum value of a performance by measuring (accuracy or kappa). The maximum measurement has been picked out from its average distribution obtained after averaging across all trials of a session of EEG signal. From results of training stages in Table 1, the average of max accuracy from 9 subjects is above 85%; where, there is around 80% of chances to separate left and right motor imagery signal if we use the ERD/ERS phenomenon in the motor cortex EEG signal. It is only a predictive value of accuracy as the classifier was applied to the same signal on which it was trained. The training stage's results for each subject indicate that the left and right imagination. Hence, the LDA is an acceptable discrimination technique to identify the PSD based features into left and right motor imagery.

Table 1 represents over all accuracy and max kappa for training and evaluation data of 9 different subjects. Comparing with BCI – IVit is found better accuracy applying LDA here. In BCI – IV for 2a data the maximum average kappa was 57%. Here, we have got 61% maximum average kappa using training and valuation data. Individually, for training data we have obtained 73% average kappa and 49% for evaluation data.

Fig. 2 depicts the percentage of overall accuracies for training and evaluation phases. Most of the cases training accuracies are higher than the evaluation accuraceis. Average training accuracy 78% where average evaluation accuracy around 72 in percentage.

Table 1 Results by classification of EEG by LDA based technique

	Training stage		Evaluation stage		Both training and evaluation stages
Subject	Overall accuracy Max (in %)	Max. kappa	Overall accuracy	Max. kappa	Average max kappa
A01	56	0.76	56	0.17	0.47
A02	50	0.56	52	0.40	0.48
A03	99	0.97	93	0.86	0.92
A04	82	0.97	61	0.22	0.60
A05	81	0.64	78	0.57	0.61
A06	69	0.61	69	0.38	0.50
A07	88	0.38	84	0.68	0.53
A08	96	0.75	84	0.68	0.72
A09	81	0.92	72	0.43	0.68
Average	78	0.73	72	0.49	0.61

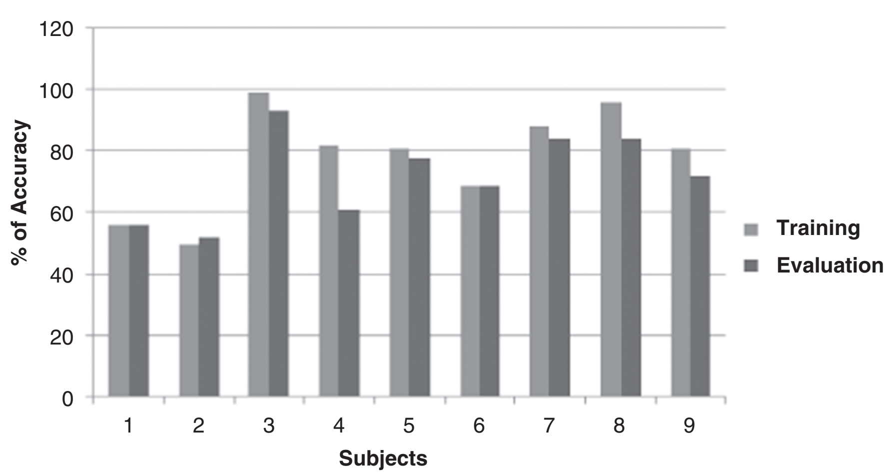

Fig. 2 Comparison of accuracy between training and evaluation data. Blue bar represents training accuracy and red bar represents evaluation accuracy.

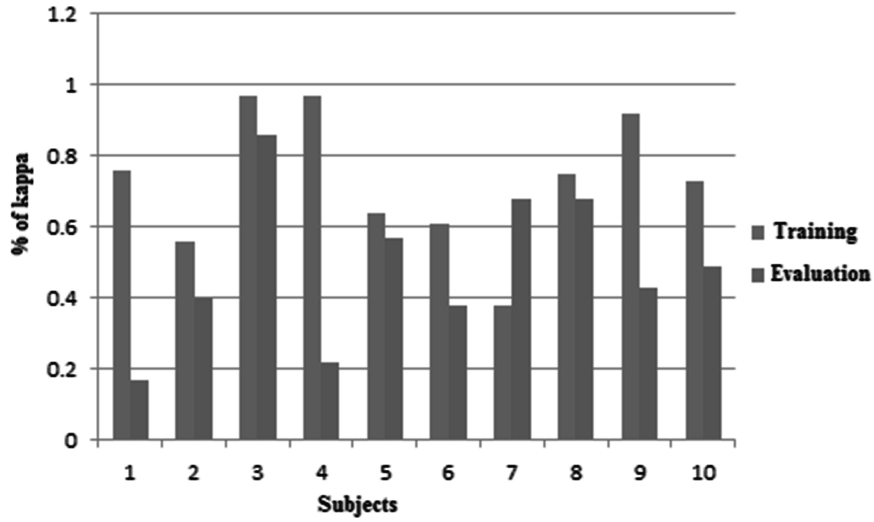

Fig. 3 Comparison of kappa between training and evaluation data. Blue bar represents training kappa and red bar represents evaluation kappa.

Fig. 3 represents the percentage of kappa for training and evaluation phases. Except for subject 7 all other cases the training kappa are higher than the evaluation kappa. Average training kappa73% where average evaluation kappa around 49 in percentage. Overall average kappa for traininig and evaluation is around 61%.

5 Conclusion

Proposed technique shows remarkably much higher and consistent MI task detection accuracy and Cohen's kappa in most of the cases. This paper shows LDA classifier with PSD based feature extraction technique with its application to an MI based BCI. In BCI competition IV average maximim kappa accuracy was 57%, we have obtained 61% by using LDA in this paper. The advantages of using this classification technique are: it uses a simple computation from a sliding windowed EEG signal; it provides performance measures are observed in training and evaluation sessions, and the accuracy or kappa distribution over the time course of paradigm is very similar; In nutshell it is observed that the LDA classification exposes a propitious technique for detecting different brain states.

Acknowledgment This research has been supported by the Ministry of Higher Education of Malaysia through the Exploratory Research Grant Scheme ERGS12-026-0026.

References

1. Ahsan, M.R., Ibrahimy, M.I., Khalifa, O.O.: EMG signal classification for human computer interaction: A review. Eur. J. Sci. Res., 3, 480–501 (2009)
2. Shahid, S., Prasad, G.: Bispectrum based feature extraction technique for devising a practical Brain-Computer Interface. J. Neural Eng., 8(2), Article Number: 025014 (2011)
3. Wolpaw, J.R., Birbaumer, N., McFarland, D.J, Pfurtscheller, G., Vaughan, T.M.: Brain-computer interfaces for communication and control. Clin Neurophysiol, 113(6):767–791, (2002)
4. Brunner, C., Leeb, R., Mü,ller-Putz, G., Schló,gl, A., Pfurtscheller, G.: BCI Competition 2008, Graz Data Set a, Laboratory of Brain-Computer Interfaces, Inst. for Knowledge Discovery, Graz Univ. of Technology (2008)
5. Rathinave, S., Arumugam, S.: Full Shoe Print Recognition based on Pass Band DCT and Partial Shoe Print Identification using Overlapped Block Method for Degraded Images. International Journal of Computer Applications (0975 – 8887), vol. 26 – No.8 (2011)
6. Garrett, D., Peterson, D.A., Anderson, C.W., Thaut, M.H.: Comparison of linear, nonlinear, and feature selection methods for EEG signal classification. IEEE Transactionson Neural System and Rehabilitation Engineering, vol. 11, pp. 141-144 (2003)
7. Hasan, M.R., Ibrahimy, M.I., Motakabber, S.M.A.: Performance Analysis of Different Techniques for Brain Computer Interfacing. International Conference on Computing, Electrical and Electronic Engineering (ICCEEE), pp. 730 – 734 (2013)
8. Sim, J., Wright, C.C.: The kappa statistic in reliability studies: Use, interpretation, and sample size requirements. Physical Therapy, vol. 85(3), pp. 257–268 (2005)