Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

Combination of Frequency Bands in EEG for Feature Reduction in Mental Task Classification

Farnaz Abdollahi[1], Ali Motie-Nasrabadi

***Abstract*— Brain-Computer interfaces require online processing of electroencephalogram (EEG) measurements. Therefore, speed of signal processing is of great importance in BCI systems. We present a method of feature reduction by combining frequency band powers of EEG, in order to speed up processing and meanwhile avoid classifier overfitting. As a result a linear combination of power spectrum of EEG frequency bands (alpha, beta, gamma, delta & theta) was found that reduces the dimension of feature vector by a factor of 5. This method gives a total correct classification rate of 91.71% comparing to 87.96% achieved from direct use of frequency band powers and 85.54% achieved from PCA feature reduction method applied to the same feature vector with 14 components.**

Key words- Brain-Computer interface (BCI), Electroencephalogram (EEG), Feature reduction, Mental tasks, Linear discriminant analysis

1. INTRODUCTION

BRAIN-Computer interface systems offer humankind a new communication channel, which enables him to control his surrounding by means of his thoughts [2]. Different technologies such as EEG, MEG and fMRI are used to observe brain neurophysiological activity among which EEG is the most popular one in BCI systems because of its cheapness and easy measurement.

The EEG measured and sampled during performance of different mental tasks shows specific characterizations which make the task-specific pattern detections possible. Many different methods are used for pattern recognition in EEG signals among which features extracted in frequency domain has been proved to be one of the best ways to recognize mental tasks [9].

In frequency domain analysis, five oscillation bands have been defined for EEG signals named Delta 0-3.5Hz, Theta 4-7Hz, Alpha 8-13Hz, Beta 14-34Hz and Gamma >35Hz [8]. Previous studies showed that these frequency bands change characteristics while performing mental tasks [3], [4] and one of the best ways of detecting these changes is using their power spectral density [10].

Although proper pattern recognition in EEG signals is one of the principals of a BCI system, other factors such as speed of signal processing and classification parameters must also be taken into consideration.

As mentioned before the goal of BCI systems is arming the disables by an online control device. Hence, a high speed signal processing in both feature extraction and classification steps is needed. On the other hand, speeding up the processing procedure must not result in a classification accuracy reduction because the correct detection of patients' orders is due the classifier accuracy.

Considering the two aspects mentioned above, reducing the feature vector dimension - referred to as feature reduction - can set a balance between speed of processing and results accuracy. In recent studies algorithms such as principal component analysis (PCA) and independent component analysis (ICA) has been used as a method of feature reduction [5].

The cost of feature reduction in these methods is loosing part of information. Although the lost data has a small proportion of feeding information, it may affect the classification accuracy. Getting the idea of combining frequency band powers from previous studies [1], [11-13], we present a new method of feature reduction by combining power spectrums of EEG frequency bands.

So, we achieved less dimensionality in feature vector while having the total feeding information. Feature reduction will also help us having a less complicated classifier and avoid overfitting.

2. METHODS

This work was supported by the Research Center for Intelligent Signal Processing.
[1]Biomedical Engineering Faculty, Amirkabir University of Technology, Tehran, Iran (E-mail: farnaz6112@yahoo.com)
[2]Department of Biomedical Engineering, Engineering Faculty, Shahed University, Tehran, Iran

A. Data Set

In this study, Anderson's EEG data set has been used for analysis. This data set contains EEG signals measured during three different mental tasks:

a) Baseline (relaxation)
b) Multiplication
c) Rotation

The sampling frequency was 250Hz and signals have been recorded in 10-second trials.

Signals have been recorded from six channels C3, C4, P3, P4, O1, O2 using 10-20 standard and a single EOG channel. The data recorded from four male subjects were used in this study after artifact removal.

B. Feature Extraction

For an effective analysis of reduced feature vector, we investigated the following processing methods:

I) Calculation of band power in predefined frequency bands in 2 second intervals;

II) Linear combination of frequency band powers calculated in section I;

III) Implementation of PCA algorithm on the feature vector of section I;

In the first method, after segmenting each 10-second record by 2-second rectangular windows with 50% overlap, five frequency band powers; delta, theta, alpha, beta & gamma were extracted by means of FFT and Parceval's law. As a result for each 10-second record 9 sets of frequency bands were achieved and we consider each set as an independent data. Putting band powers calculated in each segment for data recorded from six channels in a row leads to a dimension of 30 for feature vector. The overall view of the feature vector is shown in figure1.

$$\left[\overbrace{\delta_{11}^{C_3}\ \ \theta_{11}^{C_3}\ \ \alpha_{11}^{C_3}\ \ \beta_{11}^{C_3}\ \ \gamma_{11}^{C_3}}^{\textit{Frequency band powers of the first channel}}\ \ .\ .\ .\ \ \delta_{11}^{O_2}\ \ .\ .\ \ \gamma_{11}^{O_2}\right]$$

First window of the first trial

Figure1- Feature vector in the first method

In the next step, in order to reduce the dimensionality a combination of frequency band powers in each channel was considered. According to previous studies, using the terms in form of $(A-B)/(A+B)$ manifests the different behavior of A & B. Also, it was shown that delta, theta and alpha power spectrums show different characteristics during mental tasks in comparison with those of beta and gamma bands. Therefore, we put each of the three frequency bands in the first group in the place of A and those of the second group in the place of B consequently.

Considering these points we came to 6 different combinations shown in figure 2.

$$(\alpha-\beta)/(\alpha+\beta)$$
$$(\alpha-\gamma)/(\alpha+\gamma)$$
$$(\delta-\beta)/(\delta+\beta)$$
$$(\delta-\gamma)/(\delta+\gamma)$$
$$(\theta-\beta)/(\theta+\beta)$$
$$(\theta-\gamma)/(\theta+\gamma)$$

Figure2- 2 frequency band powers combination

These combinations suffer from lack of information so we decide to form a linear combination of the form shown in figure 3 in order to gain both less dimensionality and more feeding information simultaneously. As a result the dimension was reduced to 6 -same as previous 2 frequency band powers combination- and at the same time we keep the total frequency information of each segment.

$$((\delta+\theta+\alpha)-(\beta+\gamma))/(\delta+\theta+\alpha+\beta+\gamma)$$

Figure3- Linear combination used for feature reduction

Finally, for a better comparison between the results we implement the PCA algorithm on the feature vector of section I and we change the components from 1 up to 14.

C. Classification Procedure

As a classifier we choose a Linear Discriminant Analysis based on perceptron algorithm [6-7], [14-15]. The equation describing a single layer perceptron is:

$$g(x) = w^T x + w_0$$

In which ***w*** is the weight vector and ***x*** is the feature -

vector. The flow graph of this equation is shown in figure 4. In this procedure one third of the data was used as test data and the remaining as training data.

3. RESULTS

A. Frequency band powers

In this section three classes have been separated with a feature vector of dimension 30 and a total classification accuracy of 87.96% was resulted.

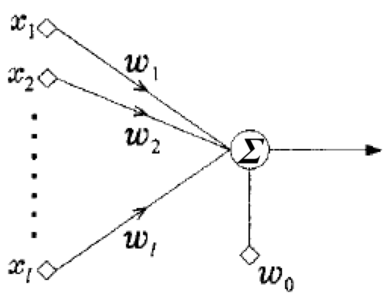

Figure4- The basic perceptron model

The classification accuracy for each of the four subjects is shown in table 1.

Table1- classification results for the first method

Sub1	Sub2	Sub3	Sub4	mean
94.69±4.6	86.39±1.22	91.67±5.24	79.09±6.4	87.96

In the next step we replace five frequency band powers of each channel with the two frequency band powers combination shown in figure 2 resulting in maximum correct classification rate of 68.8 %. The classification accuracy for each of the six features is shown in table 2.

Next, the features were computed with the linear combination shown in figure 3. In this section a total classification accuracy of 91.71% was resulted. Results are shown in table 3.

Table3- classification results for linear combination method

Sub1	Sub2	Sub3	Sub4	mean
92.38±6.78	94.42±3.65	89.31±4.1	90.71±4.98	91.71

Finally, we impose PCA algorithm based on eigen values to the feature vector of first method. In this case a feature vector of a total dimension of 30 was resulted. The features were put in place in a descending order of importance. So we start classifying the data choosing 1 to 14 components from the original feature vector. Results are shown in table 4.

Figure4 shows a better representation of the results.

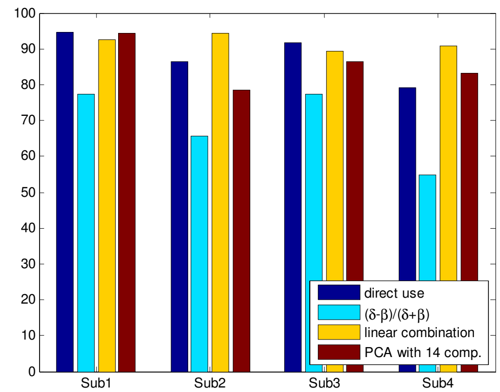

Figure4- comparison of methods

4. DISCUSSION AND CONCLUSION

In this study, we have developed a new method of feature reduction, using linear combination of frequency band powers.

First, we form a feature vector using five frequency band powers; delta, theta, alpha, beta and gamma for each channel. The resulted vector is of dimension 30. Classification with a linear discriminant function results in an average correct classification rate of 87.96% shown in table 1.

Then, as a first step to feature reduction we form combinations as shown in figure 2 with frequency band powers. In this case, the maximum correct classification rate resulted using $(\delta-\beta)/(\delta+\beta)$ as features was 68.8%.

Comparing this result with that of table 1, we can see that a severe reduction occurs in classification rate. So we decide to find a combination that keeps the feeding data and reduces the feature vector dimension both at the same time.

Therefore, we make combinations in form of figure3

Table2- classification results using combination of two frequency band powers for feature reduction

	Sub1	Sub2	Sub3	Sub4	mean
$(\alpha-\beta)/(\alpha+\beta)$	52.9±14.7	47.4±8.9	72.6±6.7	76.1±4	62.25
$(\alpha-\gamma)/(\alpha+\gamma)$	65.3±11.5	58.7±13.4	64.6±13.9	84.5±10.1	68.28
$(\delta-\beta)/(\delta+\beta)$	77.3±5.2	65.7±11.2	77.4±9.6	54.7±22.7	68.8
$(\delta-\gamma)/(\delta+\gamma)$	77.8±3.6	61.5±11.4	66.3±14.6	66.3±21.5	67.99
$(\theta-\beta)/(\theta+\beta)$	71.7±9	62.3±13.8	65.7±8.3	59.5±16.3	64.79
$(\theta-\gamma)/(\theta+\gamma)$	77.9±9.2	60.1±15.4	49.4±12.3	74.9±18.1	65.6

Table4- classification results for third section

	Sub1	Sub2	Sub3	Sub4	mean
1	44.93±23	43.65±22.7	50.63±23.4	39.16±28.2	44.59
2	45.93±21.1	44.1±17.2	53.76±28.1	53.82±9.1	49.40
3	45.1±21	47.1±23.6	58.23±22.4	54.68±7.2	51.27
4	46.79±20	56.85±9.2	69.78±6.2	53.8±8.23	58.05
5	58.94±13.1	74.02±2.8	81.57±5.2	64.82±10.5	69.83
6	76.51±9	76.14±2	81.62±5.8	63.95±7.12	74.55
7	85.04±4	77.56±3.7	82.73±6.8	71.81±7.3	79.28
8	90.44±2	78.77±8.1	84.41±6.8	75.58±14.3	82.22
9	90.32±0.9	77.7±8.5	85.01±6.8	75.28±12.9	82.08
10	94.11±4.9	78.64±7.4	83.65±6.2	79.75±9.9	84.04
11	94.47±4.7	77.74±9.5	85.4±9.35	79.15±10.3	84.25
12	94.72±4.8	78.79±8.3	86.06±9.7	81.26±9.4	85.21
13	94.97±4.2	77.98±9.7	86.26±9.3	76.58±23	83.95
14	94.33±4.6	78.41±8.6	86.37±9.9	83.07±9.2	85.54

for each channel. In this case, although the feature vector size was reduced by a factor of 5, the average correct classification rate as shown in table 3 is about 3.75% better than the direct use of same frequency band powers.

Finally, we compare the results with that of PCA algorithm based on Eigen values which is known as one of the best means of feature reduction. Table 9 shows, using 14 principal components to make the feature vector – almost twice the dimension of feature vector in linear combination method of figure 3 – we achieve less classification accuracy. The 6.17% difference between these two reduction methods proves the effectiveness of the proposed method.

5. REFERENCES

[1] Z.A. Keirn, J.I. Aunon, " *A new mode of communication between man and his surrounding*" , IEEE transactions on biomedical engineering, vol. 37, no. 12, Dec. 1990

[2] J.R. Wolpaw, N. Birbaumer, D.J. McFarland, G. Pfurtscheller, T.M. Vaughan, "*Brain-computer interfaces for communication and control*", Clinical Neurophysiology 113, 2002

[3] E. Basar, C.B. Eroglu, S. Karakas, M. Schurmann, "*Gamma, alpha, delta and theta oscillations govern cognitive processes*", International journal of psychophysiology 39, 2001

[4] G. Pfurtscheller, F.H. Lopes da Silva, "*Event-related EEG/MEG synchronization and desynchronization: basic principles*", Clinical neurophysiology 110, 1999

[5] S. Park, H. Lee, S. Choi, "ICA+OPCA for artifact robust classification of EEG data", IEEE XII workshop 2003

[6] S. Theodoridis, K. Koutroumbas, "*pattern recognition*", Academic press, 1999

[7] R. Duda, P. Hart, D. Stork, "*pattern classification*", 2nd edition, New York, Wily, 2001

[8] E. Niedermeyer, F. Lopes da Silva, "*Electroencephalography: basic principles, clinical applications and related fields*", Urban and schwarzenberg, 1987

[9] A.M. Nasrabadi, M.H. Golpaygani, Khalilzadeh, Sharifi, "*Assessing linear and nonlinear dynamics of the EEG signal for mental task Discrimination*", Amirkabir research and scientific journal, 14th year, Vol. A-55, pp. 592-600, 2003

[10] S. Solhjoo, M.H. Moradi, "*Mental task recognition: a comparison between some of classification methods*", BIOSIGNAL 2004

[11] A.M. Nasrabadi, M.H. Golpaygani, K. Tavakolian, "*Comparison between Linear and Nonlinear EEG Signal Processing during Different Mental Activities*", SCI2003, Orlando, Florida, USA, 27-30 Jul. 2003

[12] K. Tavakolian, A.M. Nasrabadi, "*A Novel Hybrid Linear and Nonlinear Feature Vector for Discriminating Five Different Mental Tasks from Spontaneous EEG signals* ", SCI2003, Orlando, Florida, USA, 27-30 Jul. 2003

[13] K. Tavakolian, A.M. Nasrabadi, "*Effects of Different Feature Vectors and Neural Network Topology on EEG Mental Tasks Classification*", IASTED International Conference in Biomedical Engineering, pp. 39-43, Salzburg, Austria, 25-27 Jun. 2003

[14] K. Tavakolian, S. Rezaei, A.M. Nasrabadi, "*Classification of Different Mental Task using Neural Network*" World Scientific and Engineering Academic and society, Issue 2, Volume 1, pp. 343-347, Apr. 2004

[15] A. Akrami, S. Solhjoo, A.M. Nasrabadi, M.H. Golpayegani, "*EEG-Based Mental Task Classification: Linear and Nonlinear Classification of Movement Imagery*", Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference, Shanghai, China, 1-4 Sept. 2005