International Journal of Smart Electrical Engineering, Vol.5, No.1,Winter 2016
pp. 1:4

ISSN: 2251-9246
EISSN: 2345-6221

Applying Genetic Algorithm to EEG Signals for Feature Reduction in Mental Task Classification

Alireza Rezaee[1]

[1]Assistant Professor of Department of system and Mechatronics Engineering, Faculty of New Sciences and Technologies, University of Tehran, Tehran, Iran, arrezaee@ut.ac.ir

---

Abstract

Brain-Computer interface systems are a new mode of communication which provides a new path between brain and its surrounding by processing EEG signals measured in different mental states. Therefore, choosing suitable features is demanded for a good BCI communication. In this regard, one of the points to be considered is feature vector dimensionality. We present a method of feature reduction using genetic algorithm as a wide search method and we choose 6 best frequency band powers of EEG, in order to speed up processing and meanwhile avoid classifier over fitting. As a result a vector of power spectrum of EEG frequency bands (alpha, beta, gamma, delta & theta) was found that reduces the dimension while giving almost the same correct classification rate

Keywords: Brain-Computer interface (BCI), Electroencephalogram (EEG), Feature reduction, Genetic algorithm (GA), Mental task, Linear discriminant analysis (LDA)

---

1. Introduction

Brain-computer interface systems based on EEG signals recorded from scull surface offer humankind a new path beside nerves and muscles between man and his surrounding, which enables him to control machines by means of his thoughts [2].

The EEG measured and sampled during performing of different mental tasks shows specific characterizations which makes 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 show that these frequency bands' characteristics change 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. The question is that there are many possible choices for reducing a feature vector dimension using its components. This wide range of possibilities can only be covered with wide search algorithms like Genetic Algorithm.

Getting the idea of choosing specific frequency band powers, we present a new method of feature reduction applying genetic algorithm to reduced feature vector as populations and try to minimize the classification error.

So, we achieved less dimensionality in feature vector while having less dimensionality. Feature reduction will also help us having a less complicated classifier and avoid over fitting.

2. Method

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:

- Baseline (relaxation)
- Multiplication
- 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 artefact removal.

B. Feature Extraction

In this section, 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.

Frequency band powers of the first channel

$$\left[\overbrace{\delta_{11}^{C_3} \quad \theta_{11}^{C_3} \quad \alpha_{11}^{C_3} \quad \beta_{11}^{C_3} \quad \gamma_{11}^{C_3}} \quad . \; . \; . \quad \underset{\nearrow}{\delta_{11}^{O_2}} \quad . \; . \quad \gamma_{11}^{O_2}\right]$$

First window of the first trial

Fig. 1. Feature vector in the first section

C. Genetic Algorithm

The typical flowchart of genetic algorithm used in this study is shown in figure 2.

- *Initial Population*

As for chromosomes, we choose 1x6 vectors among 593,775 possible choices of choosing 6 features among 30.

So we generate 50 random groups of numbers, each containing 6 numbers between 1 & 30. And then we extract frequency band powers from the vector generated in figure 1 using the group of numbers as their places.

As a result, we have 50 different feature vectors each of which containing 6 frequency band powers in the initial population.

- *Fitness Measure*

The fitness for each chromosome is evaluated by training a LDA classifier with 60% data and test the correct classification rate with the remaining 40%. Then we use classification error as a fitness parameter to be minimized.

- *Genetic Operators*

Selection: We used a statistical selection scheme for the selection process. In this scheme, the chromosomes were put into order from the highest fitness (the minimum classification error) to the lowest fitness. Then the chromosomes were selected with the probability of .8 if the chromosome was placed in the upper half of the vector and the probability of .2 if the chromosome was placed in the lower half of the vector of classification errors.

Crossover: Crossover combines the features of two parent individuals to form two similar offspring by swapping corresponding segments of the parents. The intention of the crossover operator is information exchange between different potential solutions. In this part, we choose two chromosomes from the previous generation as parents. Then we generate a random number between 1 and 5 and then put the first up to that number of the mother chromosome in the offspring and the rest of that is chosen from the remaining part of the father chromosome.

Mutation: The mutation operator used for generation of offspring in this study is also a statistical scheme. The intention of the mutation operator is the introduction of some extra variability into the population. In this part, we use random generated numbers again between 1 to 30 to produce offspring that were among 593,775 possible choices but not among those used in previous generations. In this case a number of chromosomes were added to each generation that did not exist in the previous one.

D. Classification Procedure

As a classifier we choose a Linear Discriminant Analysis based on Bays Optimal Classifier theory [6], [7]. The equation describing a linear function is:

$$g(x) = w^T x + w_0 \tag{1}$$

In which w is the weight vector and x is the feature vector.

Using a probability model describing the discriminant function and supposing class distribution of fi(x) and prior probability of πi, where x is the observation of dimension q:

$$P(class = i \mid x) = \frac{f_i(x)\pi_i}{\sum_j f_j(x)\pi_j} \tag{2}$$

It can be shown that the rule that maximizes the conditional probability shown above will give the smallest number of misclassifications which is known as Bays' rule. If we further assume that the classes have a Gaussian distribution with mean μi and covariance Σ then maximizing the conditional probability is equivalent to finding the ith class which maximizes Li.

$$L_i = x^T \Sigma^{-1} \mu_i - \mu_i^T \Sigma^{-1} \mu_i / 2 + \log \pi_i \tag{3}$$

Using maximum likelihood estimation for μi and Σ, we arrive at a linear discriminant analysis.

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 88.96% was resulted.

B. Genetic Algorithm Results

In the next step we replace five frequency band powers of each channel with a single frequency band power chosen via genetic algorithm.In order to make the best replacement, we set the genetic algorithm steps as follows;

I. Generating the first population and calculating the fitness function for each chromosome.
II. Selecting 10 chromosomes from the previous population to go directly to the next generation. This procedure is the same as one explained in Genetic Operators part (a) and has a 20% proportion in the new generation.
III. Generating 35 offspring chromosomes using 35 chromosomes of the previous generation as mothers and fathers. In this section we use the procedure explained in Genetic Operators part (b) and has a 70% proportion in the new generation.
IV. Generating 5 offspring chromosomes that did not exist in the previous generations. In this section we use the procedure explained in Genetic Operators part (c) and has a 10% proportion in the new generation.
V. If the final conditions were not met go back to step II.
VI. Performing the steps above two vectors were found having the same correct classification rate of 82.22%. These two vectors are:

[10 16 2 27 13 3]

Which means:

$$\left[\gamma_{C4} \quad \delta_{P4} \quad \theta_{C3} \quad \theta_{O2} \quad \alpha_{P3} \quad \alpha_{C3}\right]$$

And

[30 24 4 14 26 1]

Which means:

$$\left[\gamma_{O2} \quad \beta_{O1} \quad \beta_{C3} \quad \beta_{P3} \quad \delta_{O2} \quad \delta_{C3}\right]$$

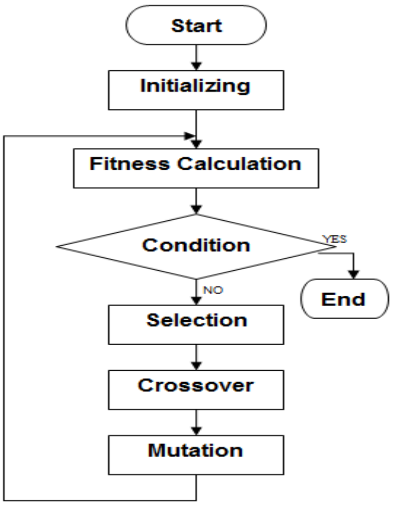

Fig. 2. Genetic algorithm flowchart.

4. Discussion and Conclusion

In this study, we have developed a new method of feature reduction, using genetic algorithm as a wide search tool.

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. In this section using a linear discriminant function we get a total classification comparing the results of these two parts we can see that the dimension is reduced by a factor of five while having only 6% reduction in the correct classification rate.

On the other hand, considering 30seconds for each possible combination if we wanted to examine every possible choices it took 17.2 month long. So genetic algorithm is a very useful procedure in this case.

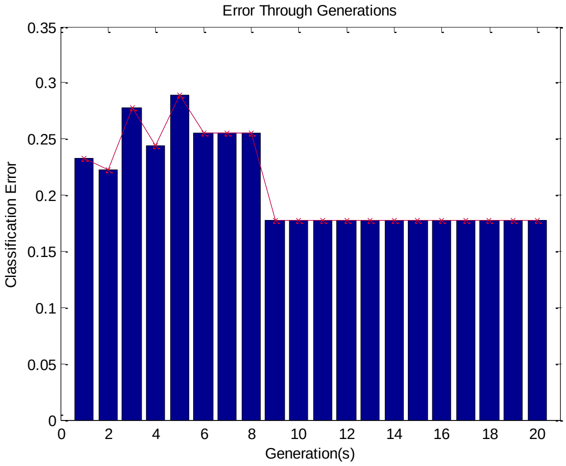

Fig. 3. Fitness function plotted every 10 generations.

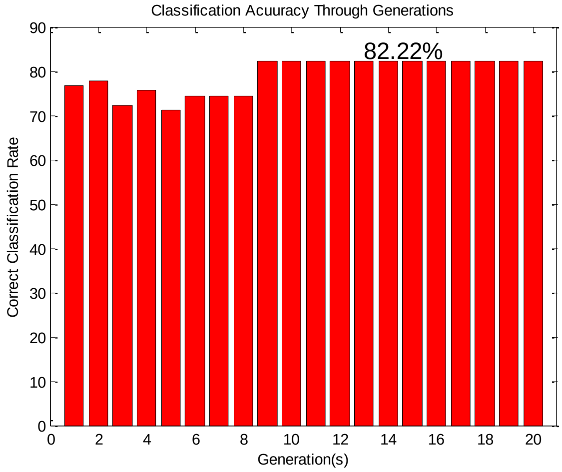

Fig. 4. Maximum correct classification rate in each 10 generations

References

[1] Zachary A. Keirn, Jorge I. Aunon, " A new mode of communication between man and his surrounding" , IEEE transactions on biomedical engineering, Vol. 37, No. 12, December 1990

[2] Jonathan R. Wolpaw, Niels Birbaumer, Dennis J. McFarland, Gert Pfurtscheller, Therese M. Vaughan, "Brain-computer interfaces for communication and control", Clinical Neurophysiology 113, 2002

[3] Erol Basar, Canan Basar-Eroglu, Sirel Karakas, Martin 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] Sugcheol Park, Hyekyoung Lee, Seungjin 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] Ernst Niedermeyer, Fernando Lopes da Silva, "Electroencephalography: basic principles, clinical applications and related fields", Urban and schwarzenberg, 1987

[9] A.M. Nasrabadi, M. Hashemi Golpaygani, Khalilzadeh, Sharifi, "Assesing 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, summer 2003

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

[11] A.M. Nasrabadi, M. Hashemi Golpaygani, et al “Genetic Algorithm as the Selection Criteria of Superior EEG Channels in Classification of Mental Tasks”, SCI2003, Orlando, Florida, USA, 27-30 July 2003

[12] A.M. Nasrabadi, M. Hashemi Golpaygani, K. Tavakolian, et al “Comparison between Linear and Nonlinear EEG Signal Processing during Different Mental Activities”, SCI2003, Orlando, Florida, USA, 27-30 July 2003

[13] K. Tavakolian, A.M. Nasrabadi, et al “A Novel Hybrid Linear and Nonlinear Feature Vector for Discriminating Five Different Mental Tasks from Spontaneous EEG signals ”, SCI2003, Orlando, Florida, USA, 27-30 July 2003

[14] K. Tavakolian, A.M. Nasrabadi, et al “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 June 2003

[15] K. Tavakolian, A.M. Nasrabadi et al " Selecting better EEG channels for classification of mental tasks", IEEE International Symposium on Circuits and Systems (ISCAS'2004), pp. 537-540, Vancouver, 23-26 May 2004

[16] 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, April 2004

[17] A. Akrami, S. Solhjoo, A. M. Nasrabadi, M. Hashemi 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 September 2005

[18] S. Solhjoo, A. M. Nasrabadi, M. Hashemi Golpayegani, "EEG-Based Mental Task Classification in Hypnotized and Normal Subjects", Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference, Shanghai, China, 1-4 September 2005

[19] K. Tavakolian, A.M. Nasrabadi, et al “Using Genetic Algorithm to Reduce Channels and Feature in EEG Mental Task Classification”, 7th IASTED, pp. 795-800, Banff, Alberta, Canada, 14-16 July 2003