Available online at www.elixirpublishers.com (Elixir International Journal)

Computer Engineering

Elixir Comp. Engg. 82 (2015) 32134-32138

ISSN: 2229-712X

Static Partitioning of EEG Signals by GA Using Multi_CSP

Masoumeh Esmaeili and Morteza Zahedi

Department of Computer Engineering & IT, Shahrood University of Technology, Shahrood, Iran.

ARTICLE INFO

Article history:

Received: 24 November 2014;
Received in revised form:
19 April 2015;
Accepted: 30 April 2015;

Keywords

Genetic Algorithm (GA),
Window,
EEG Signal,
Multi CSP,
Feature Selection,
Spatial Patterns.

ABSTRACT

In this paper a method has been proposed that uses static partitioning for improving classification of time components of EEG signals. The main idea is that different windows of signals have different power in classification. So with removing some ineffective windows from signals, the power of classification might be increased. For finding best combination of windows, Genetic Algorithm (GA) was applied. For extracting appropriate features, Common Spatial Pattern (CSP) was derived for five class problem. It applied onto each window distinguishably, and the final feature vector was obtained from placing these feature vectors altogether. LDA was used for classifying tasks. The proposed method was applied on a dataset of five mental tasks in which 30% of dataset were used for testing system. The experimental results show that window selection by GA will increase the accuracy of algorithm. This technique increased the accuracy from 69% into 95.3% for 25 windows and into 100% for 50 windows. So with changing number of windows, the accuracy of algorithm will be changed. Another important parameter is 'm' that is the number of spatial patterns selected by CSP.

Introduction

The electroencephalogram (EEG)[1] is the recording of the neural activity of brain by placing electrodes in several positions on the scalp [2]. EEG is aperiodic and complex time series that has a key role in recognition of neurological diseases and helps disabled people in brain computer interface (BCI)[3]. A BCI provides a direct pathway between brain and an external device and completely opens new communication channel without the use of any peripheral nervous system [4]. The advantages of EEG that cause high usage in many applications are mainly due to ease of use, portability and low cost of its instruments, and its fine temporal resolution [5].

Many methods have been developed for extracting appropriate features from EEG signals. In many applications of signal processing such as Automated analysis system EEG, partitioning signals into continual windows is necessary and the accuracy of signal processing can be improved [6,7,8,9]. But size of window is an important element in classification. With high variation of signals, the size of window should be lower.

In this paper GA was applied for window selection from signals. It probes to find best combination of windows. For selecting such windows, first split signals into windows and extract appropriate features from each window by multi CSP. The final feature vector is made from placing these feature vectors altogether[10]. Then LDA applied as classifier for estimating the power of each solution in GA. This proposed method was applied on a dataset with five mental tasks: Baseline, letter task, Math Task, Geometric Figure Rotation, Visual Counting task. The results show that the proposed method is successful in EEG classification with error rate near 0%.

Proposed Method

Careful analysis of the EEG signals can reinforce the learning process and increase the accuracy of classifier. As EEG signals are non-stationary, those methods that analyze the signals in frequency domain are not highly successful. But techniques that use time component domain or time-frequency domain can provide promising results [11,12]. In this paper a method has been proposed that uses static partitioning for increasing the accuracy of EEG signal classification. The base of this algorithm is that, different windows of signals have different power in classification. So with removing some windows from signals the accuracy of classifier may be increased. For selecting best windows from signals, GA can be used for finding best solution. By using GA select those windows that cause high accuracy in classification and form new signals based on them. For extracting appropriate features, multi CSP applied on each window distinguishably. After making feature vectors from each window, the final feature vector can be made by placing these feature vectors altogether. The flowchart of proposed method has been shown in Fig 1.

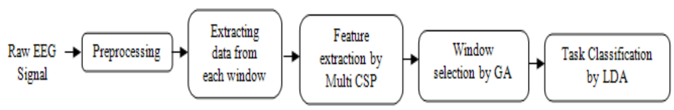


Fig 1. The flowchart of proposed method

A Static Partitioning

The proposed method is based on different resolution of signals' windows. Because if partition signals into several equal windows, and classify signals based on each window, we can see that the accuracy of each is different. For example if partition signals into fifty equal windows, and classify tasks based on each, the accuracy of them, have been illustrated in Fig 2. As you can see in this figure, the accuracy of classifier based on window 13 with 75% is highest, and windows 11 and 19 with 40%, are worst windows. So some windows can be removed from all signals, and new form of signals will be used for

Tele:
E-mail addresses: esmaeili.m2011@gmail.com

classifying tasks, that will be enhanced the accuracy of classifier.

Moreover signals might be have some noises that have been added from environment during recording them. These noises are named artifacts [13] and can be added from different resources that are out of brain [2]. Partitioning signals and removing some Windows can reduce these noises and increases the power of classification [14, 15].

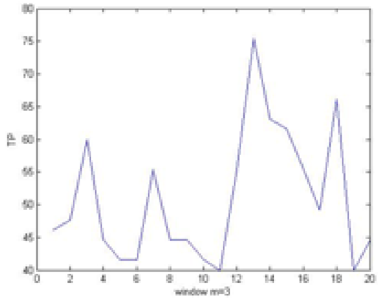


Fig 2. The power of each window in classification using multi_CSP

The method that proposed in this paper, first partition signals into equal windows and then probe best combination of windows, by GA. After finding best combination of windows, forms new signals based on best solution. Fig.3 shows a signal related to Baseline task and its first five windows, when the length of main signal is 10 seconds and the length of each window is 0.5 second.

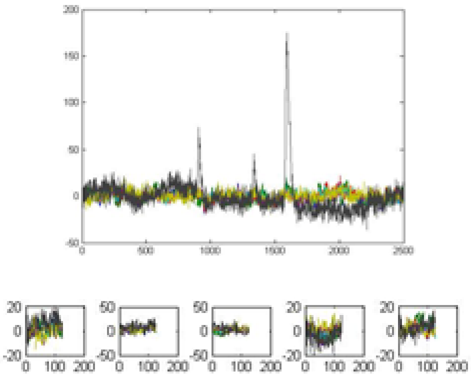


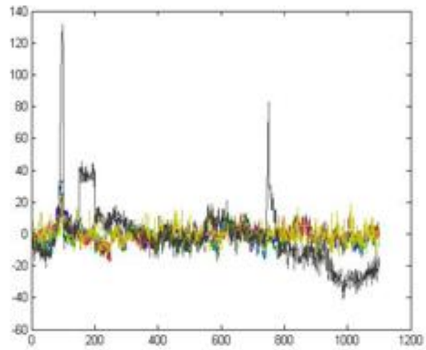
Fig 3. The EEG signal of multiplication task related to subject1 and its five first half-second windows

B Window Selection using GA

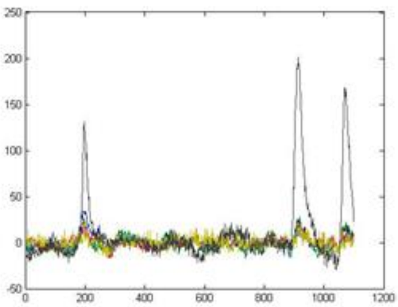
As above mentioned, removing some windows of signals can improve classification accuracy. But size of windows is important in this process. Because removing big windows from signals, causes missing a large useful data from signals. And if the size of window was small, the appropriate solution may not be found. So, the size of windows will be obtained with testing different sizes of window. For finding best formation of signals and removing fitting windows from signals, GA can be used. GA is a part of evolutionary computing, that is applicable in those problems that there are many solutions and the finding optimum solution is hard. GA represents different solutions in form of chromosomes that with combining best chromosomes, attempts to find best solutions. This iterative algorithm consists of four steps. In first step, initial population of chromosomes that represent different random solutions should be created. For this, first split signals into several equal windows and create chromosomes in length of number of windows. Each chromosome selects some of the windows randomly. In next step, the chromosomes should be evaluated. For evaluating each chromosome, select only those windows that have been chosen by it and form new signals. Then extract features from each window of new shrunk signals by multi CSP. Final feature vector will be created by placing these feature vectors altogether. For evaluating signals, classify test signals by linear discriminate analysis (LDA). True positive of classifier will be saved as fitness value of that chromosome. In the next step, best chromosomes should be selected from population as parents based on their fitness value. With combining each two parents, two children will be created. The new created children are replaced with previous population and the algorithm is resumed iteratively. The new population might be better than previous, because it has been created from best chromosomes of previous population.

For preventing from finding local optimum, chromosomes should be changed randomly, to be variate from their parents. Mutation is used for maintaining diversity of chromosomes in new population [16]. In this step that is final step, choose some chromosomes from last population randomly and change one gene of them randomly.

This optimization algorithm will be iterated until a satisfactory fitness value is obtained by one chromosome or the maximum number of iterations is achieved. Afterwards best chromosome from last population will be chosen and final new signals are shrunk based on it. The volume of new signals is lower than previous, but the power of the classifier has been raised based on it [17].



(a)



(b)

Fig 4. Forming new shrunk EEG signal based on best chromosomes:

(a) with selecting the windows 3, 4, 6, 8, 10, 12, 16, 18, 21, 25, 26, 29, 31, 34, 35, 37, 38, 39, 41, 47, 49, 50

(b) with selecting the windows 1,2,3,4,5, 14,15,16,18,19, 21, 23,25,27,29,32,34, 35,36,37,41, 42,44,45,49,50

C Feature Extraction Using Multi CSP

One of the most successful techniques for classifying tasks of brain patterns is common spatial patterns (CSPs) [18]. The primary version of CSP was designed with only two classes of

EEG signals. But Multi classes BCI was a difficult problem in BCI and feature extraction by multi CSP can largely improve the performance of classifiers. So this algorithm was extended into multi classes BCI. J.Muller-Gerking first proposed the multi-class extension of CSP, which is based on pairwise classification and voting [16]. G.Dornhege also proposed two new algorithms for multi class CSP, which improved the accuracy of classifier [5]. In [19] multi CSP for three classes is proposed that we extend it for five classes. With applying this method for extracting features from five mental tasks, the power of the classifier was raised completely. As follow, we will present the mathematical express of the algorithm.

With the theory of classifying three-class problems being the same, we will derive the algorithm for five class case, without loss of generality. Consider five mental tasks matrixes as Xa, Xb, Xc, Xd and Xe, with dimension of N × K, where N is the number of channels and K is the number of samples in time.

First the auto-covariance matrixes should be computed for each task:

$$R^{(i)}=\sum_{k=1}^{K}(x_k^{(i)}-\frac{1}{k}\sum_{k=1}^{K}x_k^{(i)}) \qquad (1)$$

Here $x_k^{(i)}$ is a N-dimensional vector at time k. *t* denotes transpose operator, i denotes the index of five classes (i.e. a, b, c, d and e classes respectively). The normalized covariance matrixes are

$$R_i=\frac{1}{l}\sum_{i=1}^{l}\frac{R^{(i)}R'^{(i)}}{trace(R^{(i)}R'^{(i)})} \qquad (2)$$

Where R is the normalized covariance matrixes and l denote numbers of trials. Trace(x) is the sum of the diagonal elements of x. As in binary CSP, We can build the composite covariance matrix as:

$$R=R_a+R_b+R_c+R_d+R_e \qquad (3)$$

Now extract the eigenvectors and eigenvalues from the matrix *R*:

$$R=U_0\Lambda U_0^T \qquad (4)$$

Where *U0 is* the eigenvector matrix and Λ is the eigenvalue matrix of *R*. then calculate the whitening matrix:

$$W=\Lambda^{-\frac{1}{2}}U_0^t \qquad (5)$$

To extract common spatial patterns of condition a, let:

$$R'_a = R_b+R_c+R_d+R_e \qquad (6)$$

As the same to three-class, we can then evaluate the transformed covariance matrixes S_a and S'_a respectively for each class a:

$$S_a=WR_aW^T \qquad (7)$$

$$S'_a=WR'_aW^T \qquad (8)$$

From statistics we know that S_a and S'_a share common eigenvectors and the sum of corresponding eigenvalues of the matrixes will be one. So S_a and S'_a can be decomposition:

$$S_a=U\Lambda_aU^T \qquad (9)$$

$$S'_a=U\Lambda'_aU^T \qquad (10)$$

And following clause is true:

$$\Lambda_a+\Lambda'_a=I \qquad (11)$$

Sort the eigenvectors in U in descending order in respect of the eigenvalues Λ_a (or in ascending order in respect of Λ'_a). In consequence, by the projection matrix:

$$SF_a=UW \qquad (12)$$

SF can be seen as spatial filter of a condition. if class ‘a’ and a’ are both projected onto the first Eigenvector U1, then class ‘a’ yields the maximal variance and class a’ causes the minimal variance. Whereas when the classes are projected onto the last Eigenvector U1,then class ‘a’ yields the minimal variance and class a’ causes the maximal variance. So, we get the mapping of each EEG trial as follow:

$$Z_a=SF_aX_a \qquad (13)$$

In practice, only few Eigenvectors are chosen, $U_m=(U_1,\dots U_m,U_{N-m+1},\dots,U_N)$, where m is low(m<<N). Finally the projection matrix is defined as

$$SF_a^s=U_mW \qquad (14)$$

The final projection matrix is defined as

$$Z_a^s=SF_a^sX_a \qquad (15)$$

As above, we only consider task ‘a’ with conditions b, c, d and e. We can get four other matrixes Z^a_b, Z^a_c, Z^a_d and Z^a_e with other conditions of a, b, c, d and e. So we can get each spatial pattern.

For extracting appropriate features from Z, in last step, calculated the logarithm transformed, normalized variance of Z:

$$f_i^k=\log(\frac{\mathrm{var}(Z_{pi}^k)}{\sum_{p=1}^{2m}\mathrm{var}(Z_p)}) \qquad (16)$$

Each row of matrix f_k^i is one of the reduced data related to task ith. The number of extracted features is twice of the number of selected eigenvectors (m). In this paper CSP was examined with two different values of m=2 and 3. So in these cases the number of features for each window reduced to 4 and 6 respectively.

III Experimental Results

To experiment the performance of proposed method, a dataset with 325 EEG signals was considered. It is from CEBL laboratory of Colorado University. For recording these signals, one electrode cap elasticwas placed on the scalp, in positions O1, O2, P3, P4, C3 and C4 (Fig 5) [20,21]. Seven subjects seated in a sound controlled booth with dim lighting and noiseless fans for ventilation. The subjects were asked to perform five mental tasks, each with five or seven trials. Each trial was recorded for 10 seconds that consists of 2500 observations. For our training dataset we used 230 signals, each task with 46 data. The test dataset comprised 95 signals, each task with 19 signals.

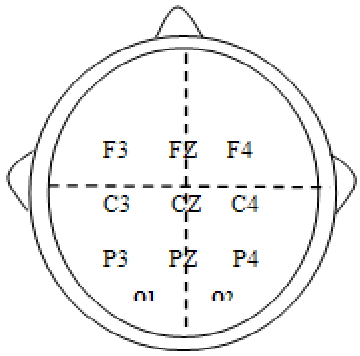


Fig 5. The position of different electrodes C3, C4, P3, P4, O1 and O2 on the scalp

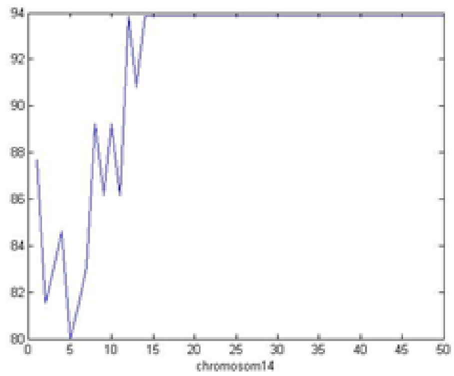
The subjects were asked to perform five following mental tasks, First Baseline, which subjects were asked to be relaxed without any thinking about anything and without any movement. Second is Letter Task, which subjects were shown images of words that each word was indicative of a friend or family member. The subjects were instructed to mentally compose a letter to a friend or relative without vocalizing or making any physical movements. Third is Math Task that subjects were given nontrivial multiplication problems, such as 63 times 84, and were asked to solve them mentally. Forth is Geometric Figure Rotation, that subjects were shown images of three-dimensional figures, and asked to visualize them being rotated about an axis. Fifth is Visual Counting task that the subjects were asked to imagine a blackboard and to visualize numbers being written on the board, one after another, sequentially in ascending order, which the previous numeral being erased before the next being written.

For training system, 70% from data were used and the other 30% of data were used for test. The proposed method was examined with two different sizes of windows and the efficiency of algorithm was examined for each. First signals partitioned into 25 windows each with length 400ms and next partitioned into 50 windows each with 200ms, respectively. GA was applied with 68 chromosomes during 50 epochs.

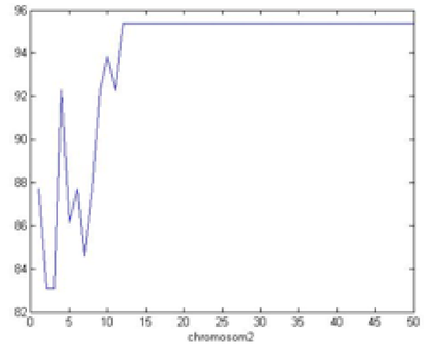
For finding best combination of windows by GA, LDA [22,23] was used for classifying tasks. Experimental results show that partitioning signals into 50 windows is more efficient than 25 windows. In other words, partitioning signals into 50 windows raised the performance of window selection by GA. Fig 6 shows the TP of best chromosome during epochs, for 25 and 50 windows respectively. In this work CSP was experienced with m=2 and m=3. As you can see, using GA with 25 window, if m=2 the average of TP is better than it with m=3. Also, with 50 window, if m=2 the average of TP is better than when m=3. In table1 this results have been compared.

Conclusion

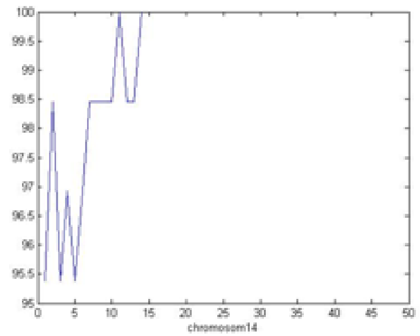
In this paper a method was proposed for classifying time components of EEG signals that uses static partitioning for improving the power of classification. Because of different windows of signals have different power in classification, so with removing some windows from signals, the power of classifier might be improved. In order to finding best combination of windows, GA was applied. For extracting appropriate features from signals multi CSP for five classes is derived. This method was applied onto all windows of a signal distinguishably. The final feature vector was obtained from placing these feature vectors altogether. The classification results by LDA illustrate that window selection by GA can improve the power of the classifier. With 50 windows, the accuracy of classifier without using window selection was 83.15% for m=2 and 84.15% for m=3. But by window selection using GA, the accuracy of the algorithm is 100%. And if the number of partitions is 25 the accuracy of algorithm is 95.3%. So with changing m (the parameter of CSP) and also with changing number of windows, the accuracy of algorithm will be improved. In this paper with m=2 and with 50 windows, best results was obtained.



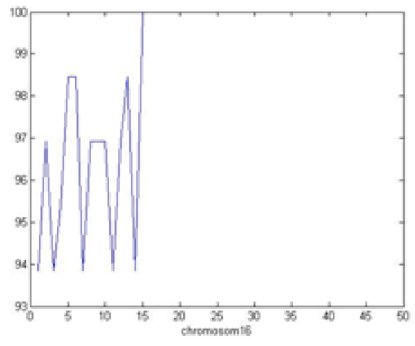
(a)



(b)



(c)



(d)

Fig 6. TP of proposed method related to best chromosomes with different size for windows and CSP argument m (a) Number of windows=25, m=2, (b) Number of windows=25, m=3 (c) Number of Windows=50, m=2, (d) Number of Windows=50, m=3

Table 1.1. Performance analysis of GA with different number of windows and different values for m

Number of windows	**Without window selection (25)**	**Without window selection (50)**	**Num. Windows = 25**	**Num. Windows = 50**
Max TP (m=2)	92.63	83.1579	95.3	100
Max TP (m=3)	81.053	84.1579	93.9	100
Average TP (m=2)	82.63	80	93.40	99.24
Average TP (m=3)	77.9	81.02	91.63	98.54

References

[1] N.J. Hill, T.N. Lal, M. Schr, T. Hinterberger, G. Widman, C.E. Elger, B. Sch, and N. Birbaumer (2006) “Classifying Event-Related Desynchronization in EEG, ECoG and MEG signals”, Interface, pp. 404-413.

[2] Knight J.N. (2003) “SIGNAL FRACTION ANALYSIS AND ARTIFACT REMOVAL IN EEG”, Master of Science THESIS, Colorado State University, Fort Collins, Colorado.

[3] F. Lotte and M. Congedo (2007) “A review of classification algorithms for EEGbased brain computer interfaces”, Journal of neural engineering, vol. 4.

[4] A.S.Muthanantha Murugavel , S.Ramakrishnan , “TREE BASED WAVELET TRANSFORM AND DAG SVM FOR SEIZURE DETECTION”, Signal & Image Processing : An International Journal (SIPIJ) Vol.3, No.1, February 2012 DOI .

[5] C.R. HEMA, M.P.PAULRAJ, R. NAGARAJAN, S. YAACOB, A.H. ADOM , “Application of Particle Swarm Optimization for EEG Signal Classification”, Biomedical Soft Computing and Human Sciences, Vol.13, No.1, pp.79-84 2008.

[6] C. Macdonald, “Long-term EEG monitoring in the Neonatal Intensive Care Unit”, *IEEE International EMBS conference of the Engineering in Medicine and Biology Society*, pp. 2483-2485, 2006.

[7] R. Agarwal, J. Gotman, D. Flanagan and B. Rosenblatt, “Automatic EEG analysis during long-term monitoring in the ICU”, Electroencephalography and ClinicalNeurophysiology, Vol. 107, pp. 44-58, 1998.

[8] K. Kosar, L. Lhotska, and V. Krajca, “Classification of Long-Term EEG Recordings,” *Book Series in Computer Science*, Vol. 3337, pp. 322-332, 2004.

[9] R. Agarwal and J. Gotman, “Adaptive Segmentation of Electroencephalographic Data Using a Nonlinear Energy Operator” Proc. IEEE International Symposium on Circuits and Systems (ISCAS'99), vol. 4, pp. 199-202, 1999.

[10] F. Goksu, N.F. Ince, A.H. Tewfik (2011) “SPARSE COMMON SPATIAL PATTERNS IN BRAIN COMPUTER INTERFACE APPLICATIONS”, IEEE ICASSP, PP 533-536.

[11] Molin G.G. Tsoneva T. Nijholt A. (2009) “Emotional Brain-Computer Interfaces” 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops، pp. 1-9.

[12] Mcfarland D.J. Mccane L.M. David S.V. and Wolpaw J.R. (1997) “Spatial filter selection for EEG-based communication”، Electroencephalography and clinical Neurophysiology، vol. 103، pp. 386-394.

[13] Dhiman R. Saini J.S. Priyanka, Mittal A.P (2010) “ARTIFACT REMOVAL FROM EEG RECORDINGS -AN OVERVIEW”, NCCI National Conference on Computational Instrumentation, CSIO Chandigarh, INDIA.

[14] Anisheh S.M. , “Adaptive Segmentation of EEG Signal”,Master of Science Thesis, Anooshiravan university, Mazandaran, Babol, Iran.

[15] Teli M.N. (2007) “DIMENSIONALITY REDUCTION AND CLASSIFICATION OF TIME EMBEDDED EEG SIGNALS”, Master of Science THESIS, Colorado State University, Fort Collins, Colorado.

[16] “Genetic Algorithm,” http://en.wikipedia.org/wiki/Eeg, last visited on 15/01/2012.

[17] J. Culpepper, “Discriminating Mental States Using EEG Represented by Power Spectral Density”, Department of Computer Science Harvey Mudd College.

[18] E. Niedermeyer, F.H.L.D. Silva, “Electroencephalography: basic principles, clinical applications, and related fields”, Lippincott Williams & Wilkins, 2005.

[19] Yan T. Jingtian T. Andong G. (2008) “Multi-Class EEG Classification for Brain Computer Interface based on CSP”، International Conference on BioMedical Engineering and Informatics.

[20] J. Lehtonen (2002) “EEG-based Brain Computer Interfaces”, Master of Science Thesis ,HELSINKI UNIVERSITY OF TECHNOLOGY Department of Electrical and Communications Engineering.

[21] E. Acar (2011) “Classification of Motor Imagery Tasks in EEG Signal and its Application to Brain-Computer Interface for Controlling Assistive Environmental Devices”, Thesis, The Graduate School of Natural and Applied Sciences of Middle East Technical University.

[22] Balakrishnama S. Ganapathiraju A. “Linear Discriminant Analysis- A Brief Tutorial ”، Institute For Signal and Information Processing Department of Electrical and Computer Engineering Mississippi State University.

[23] Kolodziej M. Majkowski A. Rak R.J. (2012) “Linear Discriminant Analysis as EEG Features Reduction Technique for Brain-Computer Interfaces”، Electrical Review، ISSN 0033-2097، R. 88 NR 3a.