# Method of Data Dimensionality Reduction in Brain-Computer Interface Systems

Rustam Fayzrakhmanov and Roman Bakunov
*Department of Information Technologies and Automated Systems, Perm National Research Polytechnic University, 29 Komsomolsky prospekt, Perm, Russia*
*fayzrakhmanov@gmail.com, bakunov_roman@mail.ru*

Keywords: Brain-Computer Interface, Information-Measuring System, Data Dimensionality Reduction, Linear Discriminant Analysis.

Abstract: The article is devoted to the problems of performance increasing of information-measuring and control systems based on brain-computer interface technology (BCI). BCI is a technology that allows communication between the brain and the external environment only on the basis of processing of the electroencephalogram (EEG). The functioning of the BCI system can be represented as a cycle. In each iteration, the EEG signal is measured and preprocessed, the characteristic features are extracted, the classification is implemented and the control action corresponding to the recognized command of the operator is generated. For the functioning of BCI systems in real-time mode it is necessary to solve the problem of the processed data dimensionality reduction (without losing significant information). The article describes the author's algorithm which is designed to use for that purpose. The algorithm is based on the use of digital signal processing and cluster analysis. Also, the results of experimental testing of the approach are described in the article. The experiments showed that proposed approach allows to significantly reduce the time required to perform operations of data dimensionality reduction. In addition, it's using has not negative affect on the clustering quality of processed sets of signals. It is experimentally confirmed that the developed algorithm effectively works in conjunction with the linear discriminant analysis (LDA), acting as a preprocessor for the LDA. At the same time, the speed of such bundle is much higher than speed of LDA without the preprocessor.

## 1 INTRODUCTION

Brain-computer interface (BCI) technology is based on the measurement of user EEG signals and the recognition of conscious brain electrical activity using digital signal processing techniques.

In BCI systems, electroencephalogram (EEG) signals analysis is often performed in the frequency domain. In this case, the set of processed signals can be represented as a set of points in N-dimensional space, where N is the number of allocated spectral components. Often this value can be equal to 32, 64, 128, etc. A large amount of coordinates in the processed vectors generates a number of problems.

Firstly, it is impossible to visualize them on a plane or in three-dimensional space for visual presentation. Secondly, the dimensionality of the data strongly influences on the computational complexity of processing operations. Thirdly, it is necessary to provide an acceptable level of clustering quality of the explored data sets, and it is often impossible when amount of coordinates is too large.

At the moment, there are methods that solve the problem of processing of multidimensional data [1], [2], [3]. One of them is a linear discriminant analysis (LDA). This is one of the fastest approaches used by researchers to reduce the dimensionality of the processed data in BCI systems [4]. However, its use requires a fairly productive computer, because it is associated with complex calculations.

Therefore, there is a need for new methods of data dimensionality reduction in BCI systems. One of these approaches is described in the article.

## 2 METHOD DESCRIPTION

The initial data are the averaged amplitude spectrums of the EEG signals measured during the preparation of the learning sample. These spectrums are considered as vectors. They form clusters that correspond to K different operator commands. Let's denote the number of vectors in each cluster by L. N is the number of components in each vector. It is proposed to represent an arbitrary cluster $\mathbf{K}_i$ in the form of a matrix (1):

$$\mathbf{K}_i = \begin{bmatrix} y_{11} & \cdots & y_{1N} \\ \vdots & \ddots & \vdots \\ y_{L1} & \cdots & y_{LN} \end{bmatrix}, i = 1...K \cdot \quad (1)$$

The rows of this matrix are the vectors included in the cluster, and the columns are the coordinates of these vectors.

The functioning of the proposed method for the data dimensionality reduction is based on the assumption that the values of the same coordinates within the current cluster have some similarity. This assumption follows from the fact that each cluster corresponds to one particular operator command. Similarly, when comparing different clusters, the same coordinates of the vectors belonging to them should differ in some way. Thus, the method is based on the idea of exploring of same columns taken from different matrices for the presence of similarities or differences.

As a specified measure of similarity or difference, it is proposed to use the following concepts: the distance between vectors (2) and the cross-correlation coefficient between two signals (3). These concepts are described in detail in [5].

$$d(\mathbf{f}, \mathbf{g}) = \|\mathbf{f} - \mathbf{g}\| = \sqrt{\sum_{k=1}^{N} (f_k - g_k)^2}, \quad (2)$$

$$r_{fg}(j) = \frac{R_{fg}(j)}{\frac{1}{N}\sqrt{\sum_{i=1}^{N} f_i^2 \sum_{i=1}^{N} g_i^2}}. \quad (3)$$

In the given equations, **f** and **g** are signals (vectors) consisting of $N$ samples, $R_{fg}(j)$ is the cross-correlation function between signals **f** and **g** at shift $j$.

In the process of the method application a number of coordinates is excluded. The possibility of exclusion is based on a special criterion. These coordinates do not have a significant effect on the differences between the clusters. Therefore, the application of the algorithm should not lead to a downgrade of the clustering quality.

A step-by-step description of the proposed algorithm is given below.

The columns of each matrix $\mathbf{K}_i$ are considered as signals whose number of samples is equal to L. The same columns from different matrices are considered in pairs. Let P be the total number of such pairs (for columns with the same numbers). It depends on the number of clusters K and is calculated by the (4), which determines the number of edges for a complete graph with K vertices:

$$P = \frac{K(K-1)}{2}. \quad (4)$$

For each pair it is necessary to calculate the cross-correlation coefficient (at zero shift) and the distance (in accordance with the equations given earlier). The result is a vector consisting of two components and shown in the (5):

$$\mathbf{f_{ip}} = \left(r_{x_i y_i}(0); d(\mathbf{x_i}, \mathbf{y_i})\right), \quad i = 0...N-1,\ p = 0...P-1 \quad (5)$$

In this equation, $\mathbf{x}_i$ and $\mathbf{y}_i$ are signals composed of elements of same columns of matrices corresponding to two different clusters. The $\mathbf{f}_{ip}$ vectors calculated for different pairs of same columns and summed up as illustrated by (6):

$$\mathbf{f_i} = \sum_{p=0}^{P-1} \mathbf{f_{ip}} = (R_i, D_i), \quad i = 0...N-1 \cdot \quad (6)$$

The key idea of the proposed method is to begin the dimensionality reduction at the least important coordinate. Therefore, it is required an objective function, the largest value of which corresponds to the most important coordinate, and the smallest – to the least important. It is proposed to adopt the expression given in (7) as this objective function.

$$Y_i = CR_i + D_i \cdot \quad (7)$$

The coefficient C is calculated in accordance with the (8):

$$C = -\frac{\max\{D_i\}}{\max\{R_i\}} + 1, \quad i = 0...N-1 \cdot \quad (8)$$

The function Yi is calculated for each vector fi. The calculated values are sorted in ascending order. After this, the coordinate numbers i (i = 0 ... N-1) are written in the order corresponding to the increase of the values of the function Yi. As a result, an array consisting of coordinates numbers sorted in order of increasing importance (determined according to an accepted criterion) will be obtained. This array will start with the number of the least important coordinate, which can be eliminated first.

The coordinates exclusion is performed step by step in accordance with their order in the described array. At each step it is important to check how it affects on the quality of clustering. If the quality of clustering has not downgraded, then the next coordinate can be excluded. Thus, the data dimensionality is reduced step by step. The values of special indicators must be calculated at each step. As such criteria it is proposed to use the compactness and isolation index CS and the efficiency index PI. Their descriptions are given in [6], an example of practical application is described in [7].

If there is a downgrade of the clustering quality in comparison with the initial values of criteria (before the dimensionality reducing), then the process ends here. Next, the best combination of values of the criteria is selected, and the corresponding step is remembered. After that, all excluded coordinates are restored, and then, the coordinates are eliminated again (in accordance with the sorted values Yi) until the specified step (including it). Thus, the final values of the clustering quality criteria will match with the selected best values.

The algorithm scheme of the proposed method of the data dimensionality reduction is presented in Figure 1.

Continuously calculation of the clustering quality criteria CS and PI can put a heavy load on the computing device, especially if it has low performance.

Therefore, it is proposed to use in practice a slightly modified version of the developed method.

The number of calculations can be significantly reduced if desired data dimensionality, which should remain after the algorithm operation, is specified before starting. In that way the required number of the least important coordinates will be excluded.

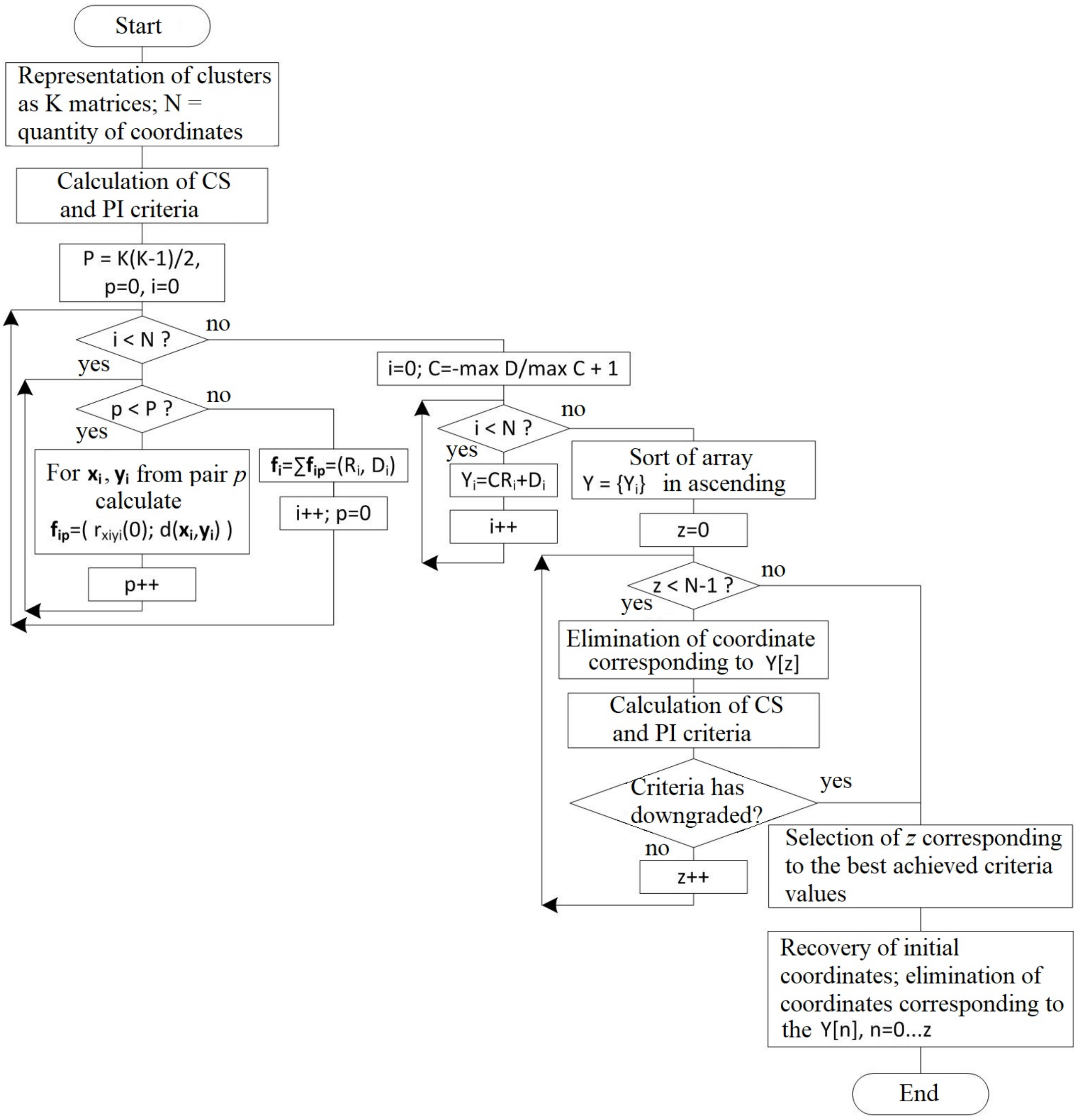


Figure 1: The scheme of the sequential elimination of the least important coordinates.

After that it will be possible to conclude whether it is permissible to use the desired dimensionality of the data or it needs to be increased. This conclusion is based on the values of the CS and PI criteria. If the desired data dimensionality reduction has led to a downgrade of the clustering quality, then it is required to restore the eliminated coordinates one by one and checking the values of the CS and PI criteria in each step of restoring.

The developed approach demonstrates the greatest efficiency in the role of a preprocessor in conjunction with LDA. Schematically, such data processing sequence is illustrated in Figure 2. Experiments have shown that using of this scheme makes it possible to effectively reduce the dimensionality of the processed data faster than in the case when LDA is used without a preprocessor. In addition, the proposed approach allows the use of all practical benefits offered by LDA, which has proven itself in information-measuring systems [8].

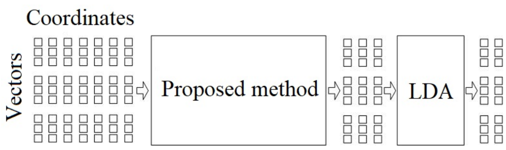


Figure 2: Proposed data processing sequence.

# 3 EXPERIMENTAL PART

The results of experimental testing of the developed method in conjunction with the LDA are given below.

EEG signals were measured by the NeuroSky MindWave Mobile system (the sampling frequency is 512 Hz).

Software for sampling and digital signal processing was created in the LabVIEW environment [9].

A computer with the following characteristics was used for the experiments:

- processor: Intel Core i7 with a clock speed of 2.2 GHz;
- RAM: 8 GB DDR3.

Analysis of the developed method was carried out in accordance with the following sequence of steps:

- Registration of EEG signals, their primary processing and grouping into clusters corresponding to specific commands of the operator.
- Splitting of each signal into segments consisting of the number of $N_{SEG}$ samples, performing a fast Fourier transform (FFT) for the segments and calculation of the averaged amplitude spectrum for each signal.
- Calculation of compactness and isolation index $CS_0$ and efficiency index $PI_0$ for clusters consisting of averaged amplitude spectrums.
- Application of LDA to the clusters. Measurement of LDA processing time $t_{LDA}$.
- Calculation of the $CS_{LDA}$ and $PI_{LDA}$ criteria for clusters modified after the application of LDA.
- Processing of the original (not affected by LDA) clusters using the developed algorithm (in conjunction with the LDA). Measurement of processing time $t_{CORR}$.
- Calculation of the $CS_{CORR}$ and $PI_{CORR}$ criteria for clusters modified after the application of the proposed algorithm.
- Analysis of the calculated indicators ($t_{LDA}$ и $t_{CORR}$, $CS_{LDA}$ и $CS_{CORR}$, $PI_{LDA}$ и $PI_{CORR}$).

This sequence of steps was repeated many times with a certain variability (the EEG signals were measured again, different operator commands were used, the sizes of the segments were changed, etc.).

Experiments have shown that the developed approach is many times faster than “pure” LDA, especially for large dimensionality of the original feature space. Graphically, this is illustrated in Figure 3.

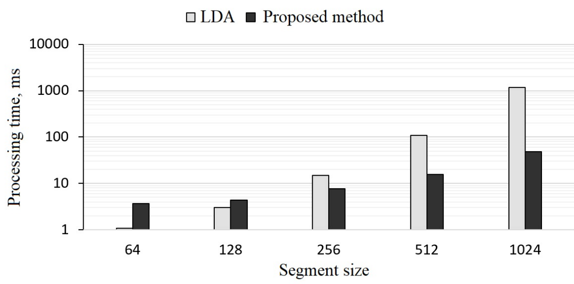


Figure 3: Performance comparison of the developed method and LDA.

Figures 4 and 5 show graphs illustrating the changes in the CS and PI criteria. These criteria are given in relative form.

Both methods improve the values of the CS and PI criteria. It should be noted that smaller values of the CS criterion are preferable because they correspond to more compact, isolated clusters. In the case of the PI criterion, larger values are preferable.

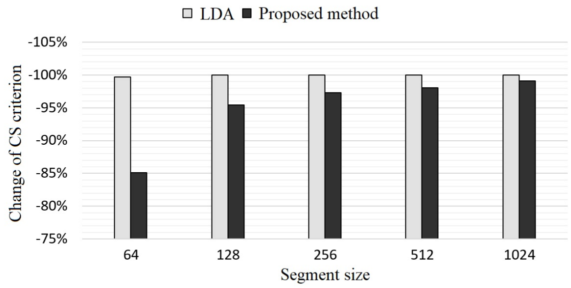


Figure 4: Comparison of the developed method and LDA by criterion CS.

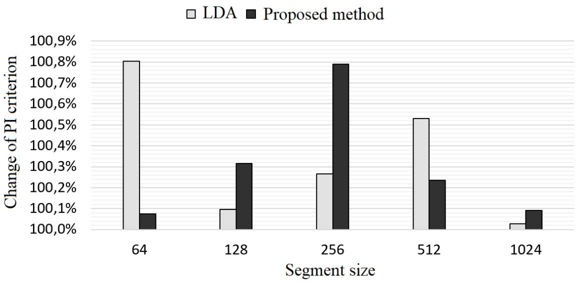


Figure 5: Comparison of the developed method and LDA by criterion PI.

Thus, the developed method of data dimensionality reduction shows a higher performance than the LDA without a preprocessor. At the same time, there are no major differences in the influence of the methods on the clustering quality criteria.

# 4 CONCLUSIONS

The described algorithm has several advantages. Its functioning does not depend on the statistical characteristics of the processed EEG signals. In addition, it is invariant to such problematic factors as the individual characteristics of the user, the design features of the BCI, the details of the control and feedback organization.

Implementation of the proposed approach does not require large computing resources. The application of the described method in conjunction with the algorithms of the learning sample reduction [10] will allow to design BCI systems based on microprocessor devices with low performance, small size of memory, low power consumption and long battery life.

The proposed method can be used not only in BCI systems. The algorithm can be implemented in any information-measuring and control systems which functioning is associated with the processing of multidimensional data.

# REFERENCES

[1] Dzh. Tu, R. Gonsales, “Principles of pattern recognition,” Mir, 1978.

[2] I. Gajdyshev, “Data analysis and processing: special reference book,” Piter, 2001, 752 p.

[3] S. A. Ajvazjan, V. M. Buhshtaber, I. S. Enjukov, L. D. Meshalkin, “Applied Statistics: Classification and Dimension Reduction,” Finansy i statistika, 1989, 607 p.

[4] T. Lan, L. Black, J. Van Santen, D. Erdogmus, “A comparison of different dimensionality reduction and feature selection methods for single trial ERP detection,” Annual international conference of the IEEE engineering in medicine and biology society (EMBC'10), 2010, pp. 6329-6332.

[5] E. C. Ifeachor, B. W. Jervis, “Digital Signal,” Processing: A Practical Approach, Second Edition, Izdatel'skij dom «Vil'jams», 2004, 992 p.

[6] A. A. Barsegjan, M. S. Kuprijanov, I. I. Holod, M. D. Tess, S. I. Elizarov, “Data and process analysis,” Tutorial, Third edition, revised and enlarged, BHV-Peterburg, 2009, 512 p.

[7] R. A. Fajzrahmanov, R. R. Bakunov, O. A. Kashin, “Application of criteria of the clustering quality in information-measuring systems,” Nauchnoe obozrenie, no. 8, pp. 231-238, 2015.

[8] R. A. Fajzrahmanov, R. R. Bakunov, “About improving of the quality of signals clustering in technical systems using linear discriminant analysis,” Elektrotehnika, no. 11, pp. 55-59, 2016.

[9] Dzh. Trjevis, Dzh. Kring, “LabVIEW for everyone,” Fourth edition, revised and enlarged, DMK Press, 2011, 904 p.

[10] R. A. Fayzrakhmanov, R. R. Bakunov, “The reduction of learning sample in information-measuring and control systems based on brain-computer interface technology,” 2nd International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM) : Proc., Chelyabinsk, Russia, May 19-20, 2016. [Online] Available: http://ieeexplore.ieee.org/document/7911544/