ON THE STATISTICAL METHODS TO LOCATE THE AREAS OF A HUMAN BRAIN ACTIVITY BY THE MEG SIGNALS AND MYOGRAMS

Andrey Gorshenin
Institute of Informatics Problems,
FRC CSC RAS
Vavilova str., 44/2, Moscow, Russia
Email: a.k.gorshenin@gmail.com

Victor Korolev
Lomonosov Moscow State University
Leninskie Gory, Moscow, Russia;
Institute of Informatics Problems,
FRC CSC RAS
Email: victoryukorolev@yandex.ru

Tatiana Zakharova
Miroslav Goncharenko
Semen Nikiforov
Maxim Khaziakhmetov
Lomonosov Moscow State University
Leninskie Gory, Moscow, Russia
Email: lsa@cs.msu.ru

Alexander Zeifman
Vologda State University,
S.Orlova, 6, Vologda, Russia;
Institute of Informatics Problems,
FRC CSC RAS;
Institute of Socio-Economic
Development of Territories, RAS
Email: a_zeifman@mail.ru

KEYWORDS

Moving separation of mixtures; Smoothed EM algorithm; Mixtures of probability distributions; Dynamic component; Diffusive component; Myogram window variance; Method of evoked potentials; Magnetoencephalography; MEG signals; Myogram

ABSTRACT

The investigation of brain activity is one of the most important fields in modern medicine. One of most popular experimental techniques is the so-called method of evoked potentials: the subject repeatedly makes some movements whereas brain activity and some auxiliary signals are recorded for further analysis. The key problem is the determination of points in myogram which correspond to the beginning of the movements (we deal with the movements of the subject's finger). The precise detection of the points promotes successful processing of the magnetoencephalogram because it is a way to identify the sensors which are closest to the activity areas.

The paper proposes a statistical approach based on a special modification of the method of moving separation of mixtures of probability distributions (MSM method) to reveal start points for the finger's movements. We demonstrate the correctness of a new procedure as compared with the method based on the notion of the myogram window variance using the same experimental myogram.

INTRODUCTION

The brain is an organ that serves as the center of the human nervous system, the cerebral cortex processes all information. Therefore, the research of brain activity is one most important problem of modern medicine. To obtain new information about the structure and relationships between the functional areas in the brain, various statistical methods can be used. This allows to investigate a human mental activity.

The precise preoperative localization of irreplaceable areas in the brain is the basis for the planning of surgeries and minimization of possible postoperative aftereffects of illness for patients with brain disorders. In clinical practice, the localization of the primary motor cortex (M1) is one of most important and difficult problems, especially the problem of detection of the hands activity's area in M1. Due to various involvements of central nervous system (e.g., as a result of epilepsy), the information about configuration of the areas by the anatomical data is deficient.

Magnetoencephalography (MEG) is a method of preoperative localization of various brain areas. It is often used in a combination with magnetic resonance imaging (MRI) to localize activity in the areas of the human brain. The method of evoked potentials is one of most popular techniques to determine the exact location of the motor areas of the cortex (McGillem and Aunon, 1987; Fabiani et al., 2007).

The external actions and some movements of the subject can be considered as the events for the analysis by the method

of evoked potentials. Such methodology leads to the increase of the signal-to-noise ratio, and it allows to reveal the brain activity corresponding to the events. The noise in this case is a superposition of the physical noise (say, generated by the noises of sensors, amplifiers, analog-to-digital conversion, external signals, network interferences, etc.) and the physiological one (which is just a background brain activity). The main problem of the method relates to the detection of the starting point of a movement.

Let us briefly describe the scheme of the experiment. The subject puts his hand on the table and taps by a forefinger for a few minutes. His MEG signals and myogram are recorded during the experiment. Additionally, the information about contacts with the table surface can be registered (say, by the accelerometer, photocell button, etc.). All signals are taken synchronously and with a strictly fixed sampling frequency. MEG signals are dataset of sensors located on the subject's head, each time series is called a channel.

The main aim of the analysis of the experimental data is the detection of the area of the cerebral cortex which is responsible for the beginning of the movement. This is a particular case of the so-called inverse problem of finding the source of a signal by the characteristics of the field generated by the source. One of the simplest solutions is to find the channel with the best response by averaging parts of MEG signals over the starts of the movements. Then, the corresponding curve could be chosen to improve the signal-to-noise ratio. However, start points cannot be determined by the MEG signals due to the part of noise in channel (it equals 0.95 and more). But the beginning of the movement can be found from the myogram, and this is sufficient for the averaging of MEG. At the same time, the values of a button can be used to adjust the fact of movement in the neighborhood of some point.

In the following sections we describe some statistical methods which were designed for finding the points corresponding to the beginning of movements.

METHOD BASED ON MYOGRAM WINDOW VARIANCE

The basic algorithm (Zakharova et al., 2012) for the reference points identification employs a simple property of myogram: its window variance associated with muscle movements, exceed significantly the one related to the muscle rest period (the window width is $30 - 50$ ms). This happens due to the physiological characteristics of the human muscles. Utilizing the fact, the basic approach consists in:

1) Construction of the empirical distribution function for the myogram window variance (MWV) and identification of some high order quantile for the function.
2) Allocation of the crossing of quantile and the ascending MWV segments and determination of the base point.
3) By this base points the MWV function is split into intervals of length not less than a predefined minimum.
4) Calculation of the MWV empirical distribution function for each interval and the estimation of the high order quantile. The crossings of quantile and the ascending MWV segments are chosen (the points of eventual array are defined as preliminary points).
5) For each preliminary point "the rest interval" is identified which is the interval between the previous and the current muscle movement, reflecting only noise muscle activity. For this interval the empirical distribution function is obtained as well as the corresponding relevant high order quantile. The last moment before the preliminary point for which the MWV is less than the relevant quantile is taken as the initial movement moment.
6) The last step is the filtration for the purpose of excluding of false responses of the algorithm.

It should be noted that myogram is a cyclic non-stationary random process which can be represented as

$$\xi(t) = \sum_i ((s_i(t) + \varepsilon_i(t) + \theta_i(t))\mathbf{1}\{t_i \leqslant t < t_{i+1}\}),$$

where the random process $s_i(t)$ corresponds to the signal component related to the finger movement ($s_i(t) = 0$ out of the movement period), $\varepsilon_i(t)$ is the rest noise (it equals zero during a movement), $\theta_i(t)$ is the movement noise (it equals zero during a rest period), in neurophysiology the half-interval $[t_i, t_{i+1})$ is called an epoch (each epoch includes the rest interval before the movement and the movement itself), i is the epoch number.

The window variance obtained from the myogram is also a cyclic non-stationary process, but much less contaminated with noise. The transfer to the window variance allows to eliminate trends and to emphasize rest-to-movement transition moments which are then determined by the threshold processing.

The main advantage of this algorithm is the construction of an adaptive threshold for each epoch. We consider pre-movement time moment, when noise threshold was exceeded, as the reference point.

METHOD BASED ON MOVING SEPARATION OF MIXTURES

To reveal the changes of the structure of forming stochastic processes in time, so-called method of moving separation of mixtures (MSM method) is successfully used (Korolev, 2011), For example, the papers with results for the financial markets (Gorshenin et al., 2008), for the traffic in information systems (Gorshenin and Korolev, 2013) and for the plasma turbulence (Gorshenin, 2014) should be mentioned.

In the present work, finite location-scale mixtures of normal distributions are used as approximation to the one-dimensional distributions of stochastic processes. Some theoretical background of these models can be found in (Korolev, 2011).

To analyze the dynamics of the changes in the stochastic process, the problem of statistical estimation of unknown parameters of distributions should be successively solved for a part of sample which moves in a direction of astronomical time (i.e., the initial sample is divided into sliding or moving sub-samples often called windows). Typically, the window (sub-sample) size is fixed. Once the analyzed parameters are obtained for a current location, the window should be moved by one element of the initial sample (i.e., the method will

analyze the next sub-sample). This allows to detect all possible changes in the behavior of components.

We suppose that the cumulative density function for corresponding moment of time (location of a window) can be represented as

$$F(x) = \sum_{i=1}^{k} \frac{p_i}{\sigma_i \sqrt{2\pi}} \int_{-\infty}^{x} \exp\left\{-\frac{(t - a_i)^2}{2\sigma_i^2}\right\} dt, \quad (1)$$

where

$$\sum_{i=1}^{k} p_i = 1, \; p_i \geqslant 0. \quad (2)$$

(for all $x \in \mathbb{R}$, $a_i \in \mathbb{R}$, $\sigma_i > 0$, $i = 1, \ldots, k$). The model (1) is called a finite location-scale normal mixture. Parameters $p_1, \ldots, p_k$ are weights under conditions (2). The parameter k is the number of mixture components. The parameters $a_1, \ldots, a_k$ are associated with the dynamic component of the volatility (variance) of the process, the parameters $\sigma_1, \ldots, \sigma_k$ are associated with the diffusion one, see (Korolev, 2011).

The main idea of detection of movements by the MSM method is as follows. First of all, the dynamic and diffusive components should be estimated. In most situations, various modifications of EM algorithm are involved (Korolev, 2011). As applied to a myogram, it was found that for both dynamic and diffusive components of the variance, there is one main component and the noise (mainly due to the computational errors). To suppress the influence of the noise, we propose a special algorithm, namely, the smoothed EM algorithm. All of EM algorithms are an iterative procedures. So, at the first step the initial values for the parameters a_i, σ_i, p_i in the ratio (1) should be determined in a way. The problem concerns the well-known fact about the dependence of EM algorithm on initial values: it finds local maxima (i.e., the estimators are closest to the initial values). Usually, to avoid the problem, initial values are chosen randomly. For the smoothed EM algorithm we use a more complicated procedure. The classical EM algorithm starts in a standard mode for some location of window (for example, $5 - 10$ movements). Then the estimators of the classic method is interpreted as the initial value for the smoothed EM algorithm. Moreover, the initial values in the next step are the estimators obtained at the current step. As a result, we obtain the unique smooth component without a noise.

One of key problems in detection of movements by the MSM method is a possible time delay. The analysis is based on sub-samples, and the influence of a new moment of time (one element of the sample) could appear with a delay. But we are interested in the precise detection of movements. The solution is based on idea of double sample processing by MSM method: in forward and backward directions. Comparing the probable points for both directions, we can find right location of movements.

APPLICATION TO REAL DATA

To detect the points of movements we use the myogram as the initial data (sample). To avoid trends in the myogram (and in the dynamic volatility component), the differences of successive elements of the sample should be found. We tested various sizes of window (e.g., 20, 30, 50) to compare the character of components. In the paper we present the results for sub-samples of size equal to 50 elements in each position of a window.

The algorithm described above was applied to the myogram. The results for both directions are represented on Fig. 1 and 2. The x-axis corresponds to the time of experiments (in milliseconds), the y-axis demonstrates the corresponding values of components. The thin solid lines are components (above is the dynamic component and below is the diffusive one).

We can see the balance statement of the subject at the beginning of components approximately until the point 3000. Note that it is very important to use the diffusive component to determine the mode due to possible ambiguities in the dynamic component. We exploit the balance statement of the subject to estimate the bounds to be used for movements detection. For example, the w-sigma rule (for the normal distribution with parameters a and σ) can be used for this purpose:

$$\mathbb{P}\left(|X - a| < w\sigma\right) = \sqrt{\frac{2}{\pi}} \int_{0}^{w} e^{-t^2/2} \, dt.$$

We can specify a sufficient probability and then find the corresponding value of w. It is the way to fix the bounds.

The probable points of movements are determined by the crossings of the bounds. Next we need to group the probable points of both directions. It seems reasonable to classify groups with the help of metric related with the window size. For example, all of probable points which are located inside the interval with length $j \cdot size_{window}$, should be classified as a group. And the starting point of a movement is the average of their time locations.

Fig. 1 and 2 demonstrate the points of movements obtained by the method based on myogram window variance and by the method based on moving separation of mixtures: the vertical red dashed lines and the green triangles, respectively. We can see a good compliance between the results of the two statistical methods. The only exception is the triangle near the mark 6200 ms (the second triangle from left). There are two reasons for that point. First, at the beginning of experiment, the subject adjusts to the conditions and can move not only his finger but his head or leg as well. But this influences the myogram. Second, for the method based on myogram window variance the information from photocell button was available. It allows to omit some false moments. For the MSM method, we had no data but the myogram.

We compared the results of the method for the components graphs. But we analyzed differences of myogram. So, it needs to compare the starting points of movements by both methods for initial myogram as is demonstrated in Fig. 3.

The x-axis corresponds to the time of experiments (in ms), the y-axis demonstrates the corresponding values of the myogram. The thin solid line demonstrates the signals of myogram, the vertical red solid lines and the green triangles are the movement points for two methods respectively. There are no changes for the conclusions mentioned above.
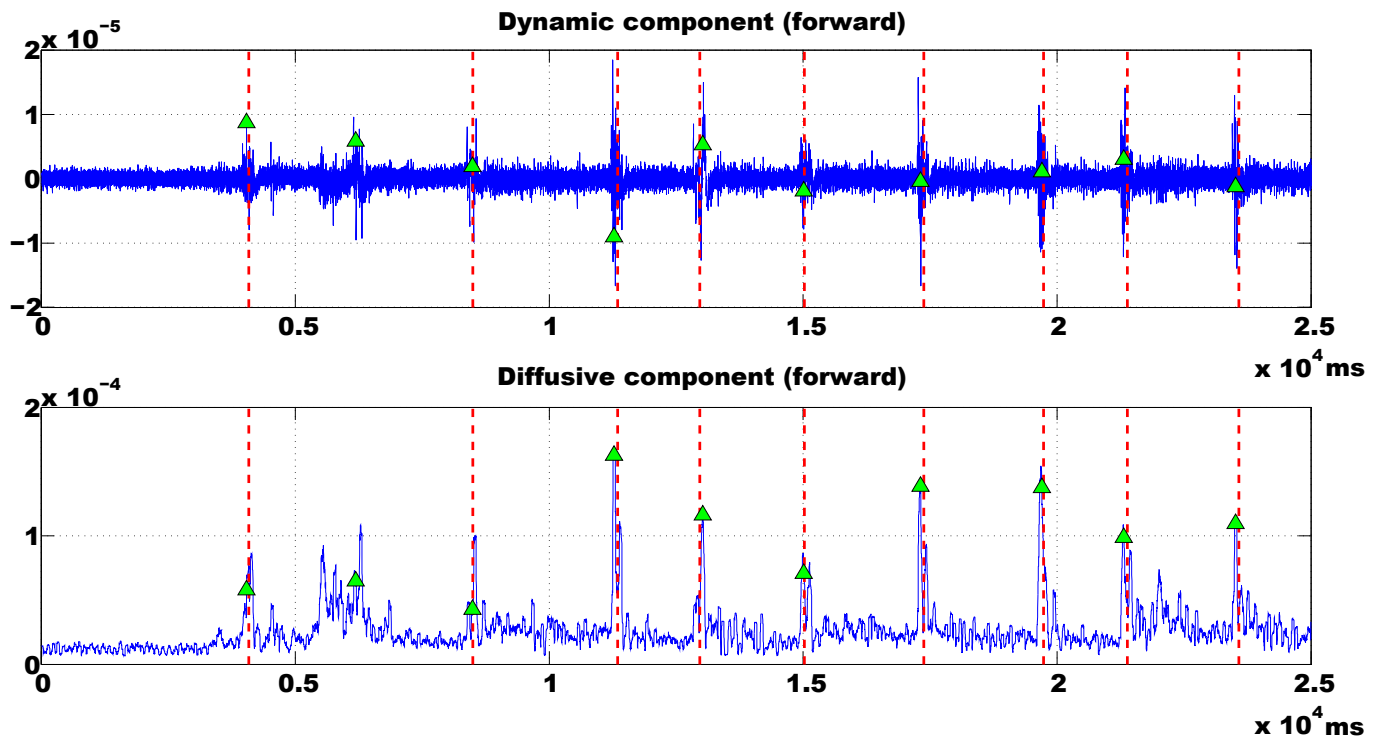
Fig. 1. The forward direction of analysis. The thin solid lines are components (a dynamic one is located on top of the figure, a diffusive component is arranged on below). The points of movements by method based on myogram window variance and by method based on moving separation of mixtures: the vertical red dashed lines and the green triangles respectively.
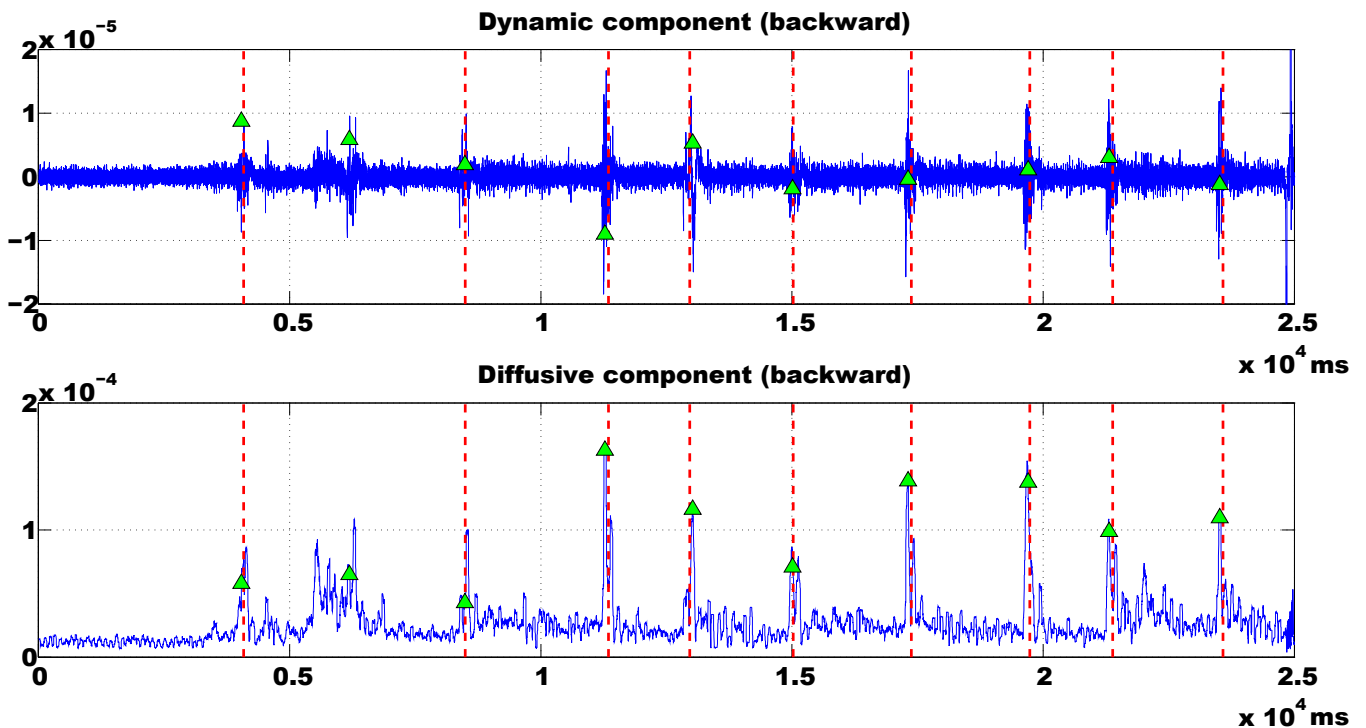

Fig. 2. The backward direction of analysis. The thin solid lines are components (a dynamic one is located on top of the figure, a diffusive component is arranged on below). The points of movements by method based on myogram window variance and by method based on moving separation of mixtures: the vertical red dashed lines and the green triangles respectively.
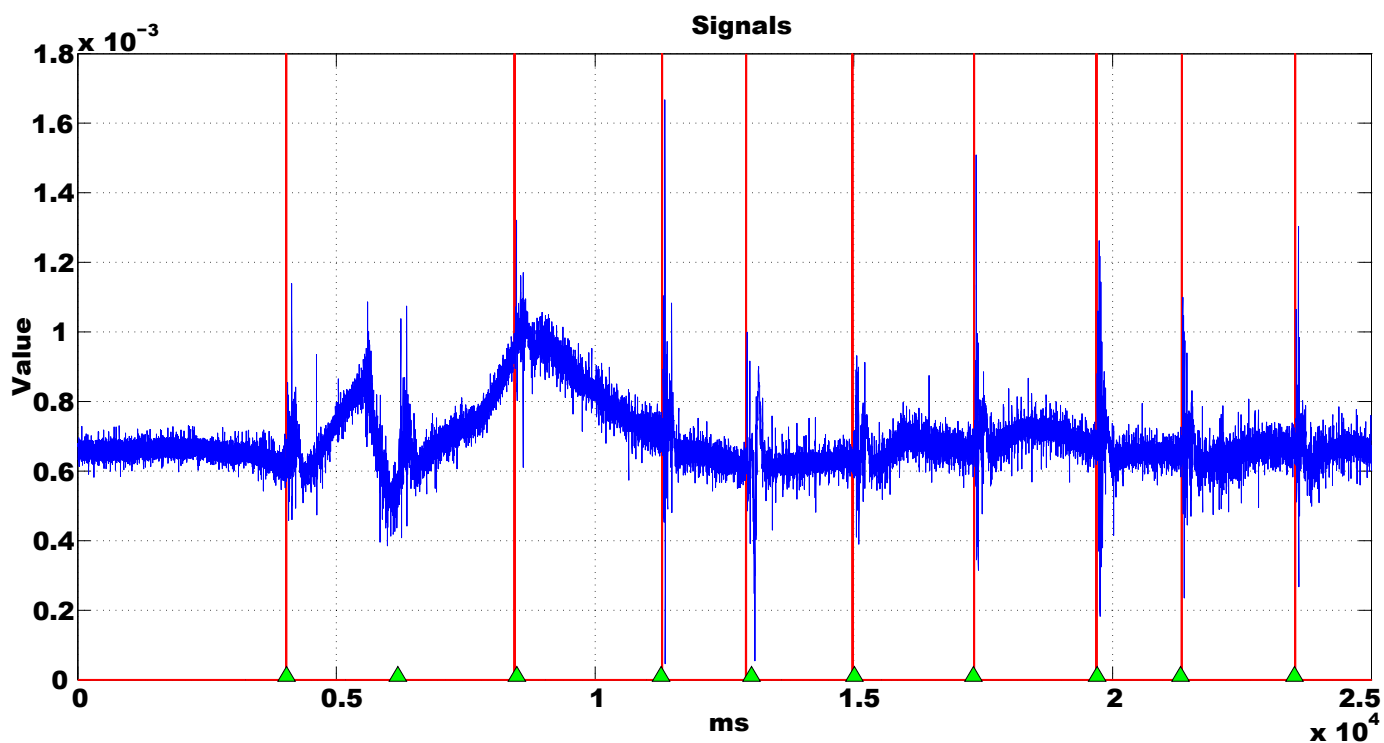
Fig. 3. The myogram. The thin solid line demonstrates signals of myogram, the vertical red solid lines and the green triangles are the movement points for two methods respectively.

CONCLUSIONS

The paper demonstrates an efficiency of two different statistical methods for solving important medical problem. The outcomes are very close, it means the results are valid. Both methods could be adjusted. For example, the method based on MSM algorithm should involve the processing of additional data (from accelerometer and photocell button) to precise location of its points. In addition, using model of probability mixture, we can obtain a convenient tool for further theoretical researches in the important field of the modern medicine. Both methods could be used for processing various types of signals of human brain.

It should be noted that one more approach for solving of the inverse problem can be found in (Bening et al., 2014).

ACKNOWLEDGEMENTS

REFERENCES

Bening, V.E., M.A. Dranitsyna, T.V. Zakharova and P.I. Karpov. 2014. "Independent component analysis for inverse problem in multi dipole model magnetoencephalogram's sources". *Informatics and Applications*, 8, No.2, 79-87.

Fabiani, M., G. Gratton and K. Federmeier 2007. "Event-Related Brain Potentials: Methods, Theory and Applications". In *Handbook of Psychophysiology 2007*, J.T. Cacioppo, L.G. Tassinary, G.G. Berntsonn (Eds.). Cambridge: Cambridge University Press, 85-119.

Gorshenin, A.K. 2014."The information technology to research the fine structure of chaotic processes in plasma by the analysis of the spectra". *Systems and Means of Informatics*, 24, No.1, 116-125.

Gorshenin, A.K., V.Yu. Korolev and A.M. Tursunbayev. 2008. "Median modification of EM- and SEM-algorithms for separation of mixtures of probability distributions and their application to the decomposition of volatility of financial time series". *Informatics and Applications*, 2, No.4, 12-47.

Gorshenin, A. and V. Korolev. 2013. "Modeling of statistical fluctuations of information flows by mixtures of gamma distributions". In *Proceedings of 27th European Conference on Modelling and Simulation* (May 27-30, 2013, Alesund, Norway). Digitaldruck Pirrot GmbHP, Dudweiler, Germany, 569-572.

Korolev, V.Yu. 2011. *Probabilistic and Statistical Methods of Decomposition of Volatility of Chaotic Processes.* Moscow University Publishing House, Moscow.

McGillem, C.D. and J.I. Aunon. 1987. "Analysis of Event-Related Potentials". In *Methods of Analysis of Brain Electrical and Magnetic Signals: EEG Handbook. 1987*, A.S. Gevins, A. Remond (Eds.). Amsterdam: Elsevier Science Publishers, 131-169.

Zakharova, T., S. Nikiforov, M. Goncharenko, M. Dranitsyna, G. Klimov, M. Khaziakhmetov and N. Chayanov. 2012.

"Signal processing methods for localization of nonrenewable brain regions". *Systems and Means of Informatics*, 22, No.2, 157-175.

AUTHOR BIOGRAPHIES

ANDREY GORSHENIN is Candidate of Science (PhD) in physics and mathematics, senior scientist, Institute of Informatics Problems, Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences; associate professor, Institute of Information Technology, Moscow State Technical University of Radioengineering, Electronics and Automation. His e-mail address is: a.k.gorshenin@gmail.com

VICTOR KOROLEV is Doctor of Science in physics and mathematics, professor, Head of the Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M.V. Lomonosov Moscow State University; leading scientist, Institute of Informatics Problems, Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences. His e-mail address is: victoryukorolev@yandex.ru

TATIANA ZAKHAROVA is Candidate of Science (PhD) in physics and mathematics, senior lecturer, Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M.V. Lomonosov Moscow State University.

MIROSLAV GONCHARENKO is PhD student, Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M.V. Lomonosov Moscow State University.

SEMEN NIKIFOROV is PhD student, Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M.V. Lomonosov Moscow State University.

MAXIM KHAZIAKHMETOV is PhD student, Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M.V. Lomonosov Moscow State University.

ALEXANDER ZEIFMAN is Doctor of Science in physics and mathematics; professor, Dean of the Faculty of Applied Mathematics and Computer Technologies, Vologda State University; senior scientist, Institute of Informatics Problems, Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences; leading scientist, Institute of Socio-Economic Development of Territories, Russian Academy of Sciences. His e-mail address is: a_zeifman@mail.ru