ABOUT CLASSIFICATION OF ECG SIGNALS BASED ON HIGH-FREQUENCY WAVELET COMPONENTS

Paulina N. Podkur[a,*], Nikolay K. Smolentsev[b]

[a] Kemerovo Institute (branch) of Plekhanov Russian University of Economics,
Kuznetsk Ave. 39, Kemerovo, 650992 Russian Federation

[b] Kemerovo State University,
Krasnaya Str. 6, Kemerovo, 650043 Russian Federation

* e-mail: paulina.podkur@gmail.com

Abstract: The new design method of the automated classifying system for electrocardiograms recognition (ECG) of healthy and ill patients, based only on high-frequency components of ECG signal with the use of statistical images recognition is offered. Cardiograms of two groups of patients were studied: healthy and those who came through myocardial infarction. The first step of classification method is ECG wavelet decomposition to the 4th level and allocation of four high-frequency ECG components. The choice of the 4th decomposing level is explained by the fact that the first four high-frequency components represent high ECG frequencies from 30 to 350 Hz, and low-frequency component represents the undistorted smoothed ECG signal cleared of high-frequency oscillations. In case of more deep signal expansion the following high-frequency component has frequency spectrum to 30 Hz, and low frequency component is significantly distorted. For each of the first four components of wavelet decomposition there is a number of ECG numerical signs, including energy, entropy and frequency characteristics, 21 signs in total. During the second step reduction of the dimensionality of the feature space by using scatter matrix is made for two chosen ECG groups. It has turned out that the reduced feature space is one-dimensional. Histograms of values of this one-dimensional feature for groups of healthy and ill patients are constructed. The third step is finding of the dividing constant which is able to distinguish both groups of ECG records. For testing 96 ECG records of patients with normal cardiograms and 120 ECG records of the patients who came through myocardial infarction are used. Only three features (3%) of 96 given features values of the first group are referred by the classifier to patients group and only 20 features (< 17%) of 120 given features values of the second patients group are referred by the classifier to ECG group of healthy patients. Considering that for each patient the system determines 12 features by 12 standard assignments, testing results show well classification accuracy.

Keywords: wavelet analysis, electrocardiogram, high-frequency ECG components, statistical pattern recognition

INTRODUCTION

Heartbeat represents a complex electrochemical process. It is registered in the form of electrocardiogram by skin electrodes placed in certain places of body surface. One heartbeat cycle recorded on ECG usually consists of several bursts: P wave, then QRS complex, T wave and U wave (fig. 2). After a while this complex of PQRSTU waves repeats. The specified waves, their sizes, sort, rhythms, intervals and PR and QT segments traditionally serve for heart diseases diagnosis. The form of waves, waves and segments' complexes duration, lengths variability of various cardiosignal intervals are analyzed. For ECG studying various statistical methods, Fourier transformation and spectrum analysis are used. More modern methods are based on wavelet analysis and artificial neural networks application [1], [2], [3], [4] and [5]. In these researches it is considered that high cutoff frequency of normal (without loading) cardiosignal noticeably influencing its form does not exceed 100 Hz. That is why in such analysis ECG of higher than 100 Hz frequency is almost not considered. Moreover, for cardiogram smoothing high-frequency components are usually deleted by means of various filters. It is clear, that at that part of information registered by the cardiograph is lost. The physical origin of high frequencies of a cardiosignal is not clarified up to the end. They can include both the hardware noise and high-frequency physiological rhythms which are to a large extent consequence of heart electrical activity, as they are registered by the sensors located near heart. In the modern technical means the hardware noise are almost insignificant in comparison with physiological rhythms. That is why ECG high-frequency components reflect electrical heart activity and therefore for registration of high frequencies electrocardiographs of high resolution with sampling frequency of 5, 10, and 20 kHz are used nowadays. Effective allocation of high-frequency components is possible with the use of wavelet decomposition of signal. There are publications in which high frequencies of cardiosignal are analyzed by means of the continuous wavelet decomposition. For example, in works [6], [7] and [8] the digitized cardiosignals with sampling rate of 5, 10 and 20 kHz and properties of cardiosignal at frequencies from 25 Hz to 400 Hz were studied. In works [9] and [10]

frequency and stochastic characteristics of high-frequency cardiosignal components with the use of discrete wavelet decomposition have been analyzed.

Long ECG records of patients contain a huge amount of data. That is why detection of disease symptoms is time-consuming process which demands the detailed analysis of all ECG data length. Reliable automatic classification and system detecting ECG parameters' anomalies would guarantee more exact diagnostics and significant facilitation of cardiogram "decoding". Therefore creation of the automated system detecting ECG parameters' anomalies and ECG classification is an urgent task. In this direction certain results based on studying of low-frequency ECG characteristics are achieved, e.g. [2], [4] and [5]. Nowadays, the problem of creation of the automated systems of ECG classification considering high-frequency components is open issue. In this work the description of the automated classification system, separating ECGs of ill and healthy patients using only high- frequency ECG component is given. Decision-making process consists of three main stages: wavelet decomposition of ECG and features extraction; reduction of the dimensionality of the feature space using scatter matrix and classification by the linear classifier.

The offered method was tested on ECG data sets which belong to two groups of patients: healthy patients (96 ECG signals) and the patients who have recently come through myocardial infarction (120 ECG signals). According to the results of testing, record definition accuracy of ECG lead of healthy patients is 97%, and record definition accuracy of ECG lead of ill patients is 83%. Results of testing confirmed that the offered algorithm has potential in classification of ECG signals and can improve diagnosis of heart disease.

MATERIALS AND METHODS

Materials. For classification system creation and its testing ECG data sets of two groups are studied: healthy patients with normal ECG data and patients who recently came through myocardial infarction (MI). For the analysis digitized 30 seconds long cardiosignals made on the high-resolution cardiograph (1028 counts per second) "Cardiotekhnika – 4000, by EcgShell" were used. The cardiosignal is registered on 8 standard channels: L – left hand (+) and right hand (–), F – left leg (+) and right hand (–) and six chest leads marked as C1 – C6. Of 8 cardiograph channels L, F, C1, C2, C3, C4, C5, C6 there are 12 so called standard leads [11]: I, II, III, aVR, aVL, aVF, V1, V2, V3, V4, V5, V6 according to formulae:

$$I = L, \quad II = F, \quad III = F - L,$$

$$aVR = -(L+F)/2, \quad aVL = L - F/2, \quad aVF = F - L/2,$$

$$Vi = Ci - (L+F)/3, \qquad i = 1,2,\ldots, 6.$$

8 seconds long fragments were chosen from ECG records of each of 12 leads, wavelet decomposition and features calculation are made for them. All calculations are performed in MATLAB system [12] using wavelet analysis package MATLAB Wavelet Toolbox [10].

Functions of this wavelet analysis package provide correct processing of boundary values at filters action by symmetric signal continuation. Thus, for each patient 24 ECG fragments were studied. During the construction of classifying system ECG records of two groups of patients were used. The first group of ECG records of healthy patients contains 96 ECG fragments for four patients aged from 21 to 27. The second group of ECG records of the patients who recently came through myocardial infarction (subacute period) contains 120 ECG fragments for five patients aged from 44 to 55. For testing 96 ECG fragments of four healthy patients aged from 21 to 56 years and 120 ECG fragments for five patients aged from 45 to 57 years which recently came through myocardial infarction (subacute period) were used. All data contain only ECG records and information about patient's age. Cardiograms' analysis for the purpose of the diagnosis was not made.

Methods. In this work the automated classification of ECG signals for an assessment of patients' condition based only on high-frequency components of ECG signal received by multi-level wavelet decomposition and with the use of statistical recognition of images is offered. The first step of classification method is receiving a number of properties after wavelet decomposition of ECG data, including energy, entropy and frequency characteristics the first four components of wavelet decomposition. In the second step reduction of the dimensionality of the feature space with the use of scatter matrix is made. The third step consists in using of linear classifiers which are able to distinguish both groups of ECG signals.

Wavelet decomposition. The main operation of wavelet analysis [10] is decomposition of the studied signal $S = \{S_n\}$ in two components $D_1 = \{D_{1,k}\}$ and $A_1 = \{A_{1,k}\}$ by means of some filters. A_1 array represents smoothed part of signal and it is called approximation coefficient array. D_1 array represents details in which the initial signal S differs from its smoothed part. From the point of view of signals' analysis orthogonal wavelets represent four digital filters $\{h_n\}$, $\{g_n\}$ и $\{h_n^*\}$, $\{g_n^*\}$ [9]. Filters $\{h_n^*\}$, $\{g_n^*\}$ are used for decomposition of signal in formulae:

$$A_{1,k} = \sum_n h_n^* S_{2k-n}, \quad D_{1,k} = \sum_n g_n^* S_{2k-n}. \quad (1)$$

The result of filter $\{h_n^*\}$ action is low-frequency approximation of the signal. The result of filter $\{g_n^*\}$ action is high-frequency part of the signal. Filters $\{h_n\}$ and $\{g_n\}$ are used for signal restoration $S = \{S_n\}$ according to the formula:

$$S_n = \sum_k (h_{n-2k} A_{1,k} + g_{n-2k} D_{1,k}). \qquad (2)$$

During multi-level wavelet analysis the procedure of wavelet decomposition (1) is used many times to approximation coefficient array. It can be represented schematically as follows (Fig. 1):
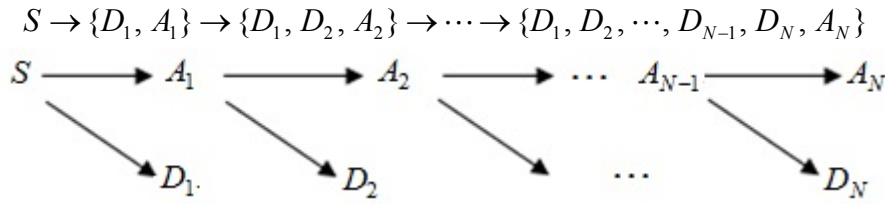
$$S \to \{D_1, A_1\} \to \{D_1, D_2, A_2\} \to \cdots \to \{D_1, D_2, \cdots, D_{N-1}, D_N, A_N\}$$

Fig. 1. Multi-level wavelet decomposition of signal S.

Restoration of the initial signal is made consistently in reverse order. If we apply restoration procedure only to one set of coefficients and all other coefficients consist of zero, then we will receive the part of a signal corresponding to one set of coefficients. We will call this part the *component* of a signal. The signal components restored only on coefficients of details $D_1, D_2, \ldots, D_N$, will be called high-frequency and will be defined as $RecD_1, RecD_2, \ldots, RecD_N$, respectively (Fig. 2). For example, $RecD_2$ is a signal component, restored on the following set of wavelet coefficients $\{0, D_2, 0, \ldots, 0\}$, where 0 means the array from zeros. Similarly, low-frequency components $RecA_1, RecA_2, \ldots, RecA_N$ came out by restoration only of one set of approximating coefficients. The sum of all received signal components $RecD_1, RecD_2, \ldots, RecD_N$ and $RecA_N$ is equal to the original signal:

$$S = RecD_1 + RecD_2 + \ldots + RecD_N + RecA_N .$$

We will note that for calculation of wavelet components of S signal the entire S signal is used. Though coefficients of wavelet decomposition have lengths which are consistently decreasing twice wavelet components of a signal have the same length as the S signal itself and in the formula stated above addition of components is coordinate wise.

Feature space. For every high-frequency decomposition components $RecD_1, RecD_2, \ldots, RecD_N$ many various statistical, frequency and stochastic characteristics can be calculated:
– maximum absolute value;
– dispersion;
– L^1- and L^2-energy;
– relative L^2-energy;
– maximum value of the power spectrum;
– frequency, where maximum value of the power spectrum is achieved;
– Shannon's entropy;
– the Hurst exponent and other characteristics of randomness;
– average value of instantaneous frequency of oscillations calculated on the basis of discrete Hilbert transform.

Definitions of the listed parameters will be reminded. L^1-energy of signal $X = \{x_n\}$ is the sum of elements' modules x_n, and L^2-energy of signal $X = \{x_n\}$ is the sum of squares of elements' modules x_n. Relative energy of signal component is the ratio of L^2-energy of component to L^2-energy of the entire signal.

Discrete Fourier transform $C = \text{fft}(X)$ of signal $X = \{x_n\}$ of N length is made according to the formula:

$$c_k = \sum_{n=0}^{N-1} x_n\, e^{-i\frac{2\pi}{N} k\, n}, \quad k = 0, \ldots, N-1 .$$

The received signal $C = \{c_k\}$ shows frequency properties of the array $\{x_n\}$, therefore it is called a signal spectrum $\{x_n\}$. As values $\{c_k\}$ can be complex, the so-called power spectrum of frequencies (frequency spectrum) calculated according to the formula is of interest:

$$P_k = \frac{|c_k|^2}{N}, \quad k = 0, 1, \ldots, N-1 .$$

The diagram of power spectrum P_k (Fig. 3) is usually figured for values k in the range from 0 to the middle of $N/2$ as it is symmetric and it makes sense to consider only the frequencies which are smaller than Nyquist frequencies corresponding to $k = N/2$. In Fig. 3 diagrams of power spectra for components of cardiosignals of two patients from different groups are given. As capacities of frequencies over 360 Hz are almost not noticeable in the figure, these diagrams are given in the range from 0 to 360 Hz.

Shannon's entropy characterizes dispersion of signal values $X = \{x_n\}$ and is defined by the formula:

$$E(x) = -\sum_k x_k^2 \ln(x_k^2) .$$

Components implementation of signals can be rather difficult and of randomness nature. Degree of randomness can be estimated by Hurst exponent. It represents a propensity score of process to trends. $H > 0.5$ value means that process dynamics directed to a certain side in the past most likely will lead to movement continuation in the same direction. If $H < 0.5$, then it is predicted that process will change its direction, and $H = 0.5$ means uncertainty. Calculation of Hurst exponent for the signal $\{x_n\}$ is usually made on the basis of the so-called RS *analysis*. Let's remember a classical RS method for Hurst exponent finding. Let the signal $\{x_n\}$ of length N be given. Then Hurst exponent H can be found from the ratio:

$$R/S = (N/2)^H,$$

where S is a standard deviation with selective average m_X, and R is the so-called scope, accumulated deviation from average:

$$R = \max_{1 \le n \le N} \sum_{k=1}^{n}(x_k - m_X) - \min_{1 \le n \le N} \sum_{k=1}^{n}(x_k - m_X) .$$
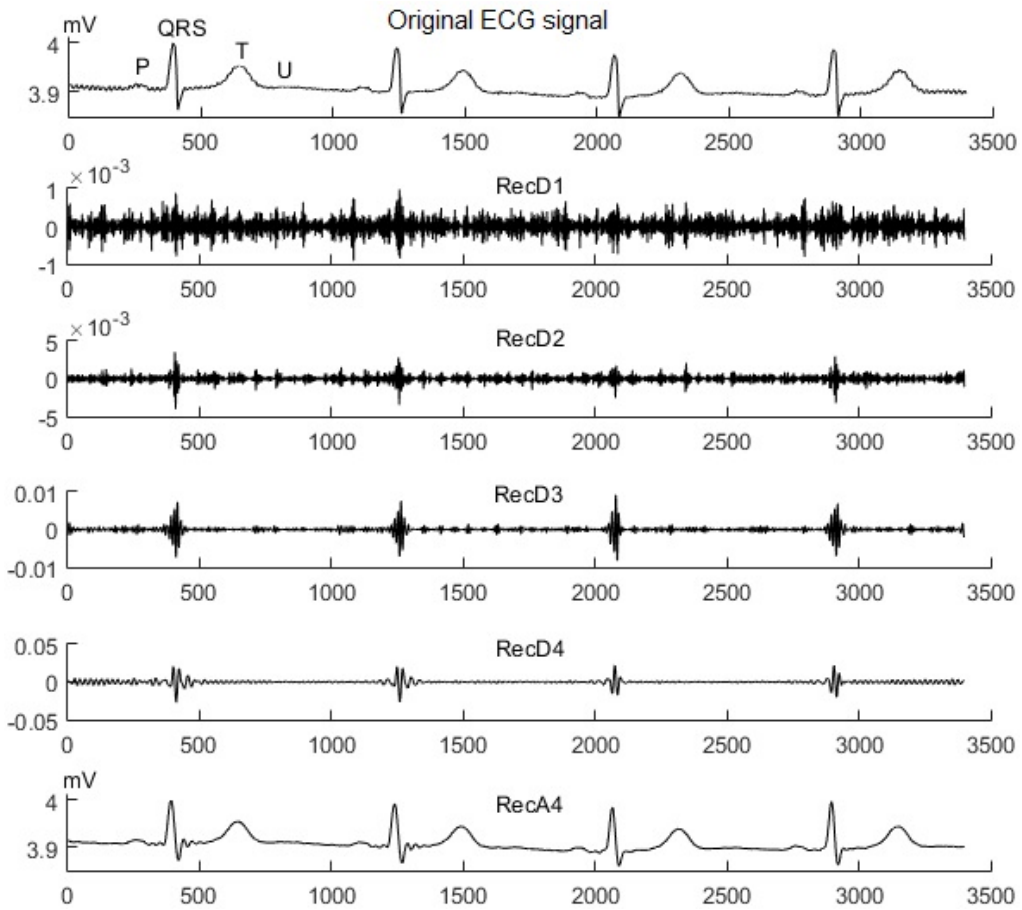
Fig. 2. Wavelet components of ECG signal during decomposition to the 4th level (across is signal counting from 1 to 3400).
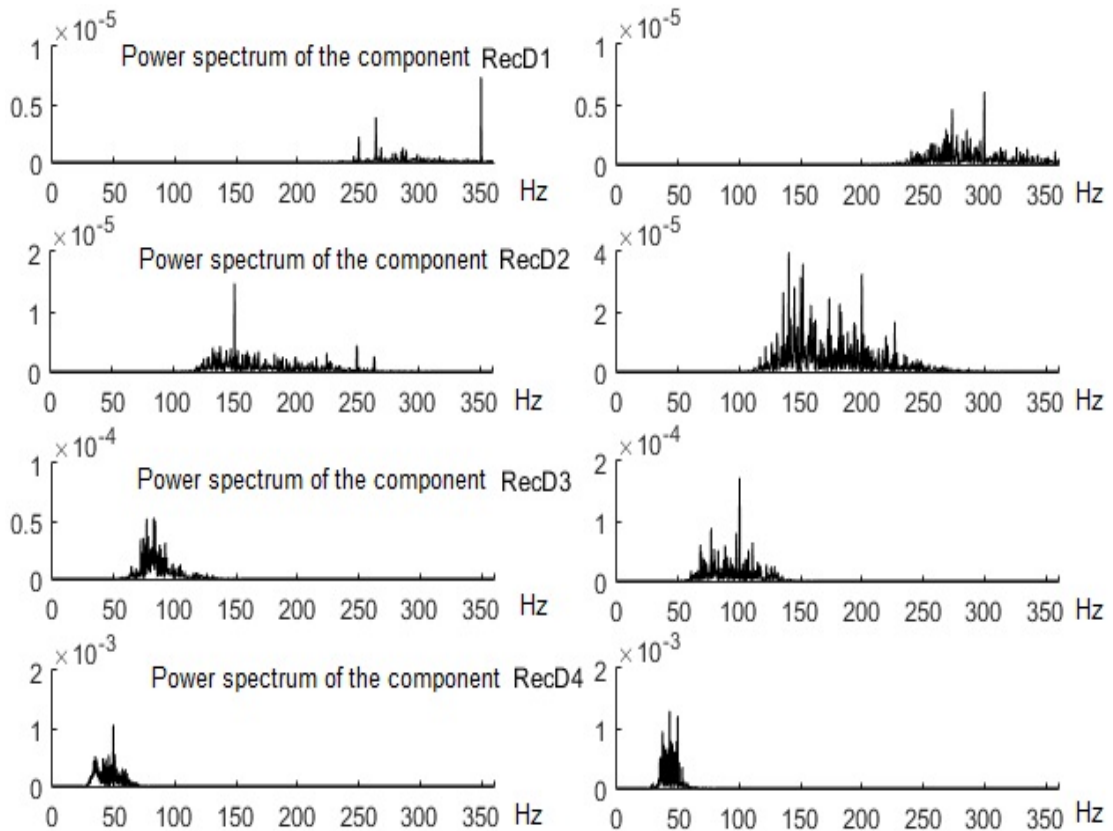

Fig. 3. Power spectra diagrams of wavelet components of a ECG signal when decomposition is to the 4th level on the interval from 0 to 360 Hz. On the left is for healthy patient, on the right – for ill patient.

In MATLAB there is `wfbmesti` function which assesses the fractal Hurst index of H signals. Hurst exponent is one of characteristics of randomness degree. Besides, according to the available one-dimensional data it is possible to construct dynamic system in multivariate phase space for which the observed variable will be one of coordinates, and system tracks lies on some set having fractal structure and fractional dimension. Therefore for the assessment of randomness of wavelet-coefficients and ECG signal component it is also possible to use such characteristics as: phase space dimension, fractal dimension and correlation dimension. It is possible to read about it in more detail in the work [10] where these characteristics were used for the analysis of high-frequency components of ECG signal, see also [13] where stochastic parameters were used for research EEG of healthy patients in comparison with EEG of patients with epilepsy.

We will remind that Hilbert transform $y(t) = H(x(t))$ of the function $x(t)$ is defined by the formula:

$$y(t) = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{x(\tau)}{t - \tau} d\tau \, ,$$

of course, if this integral exists in sense of a principal value. One of the basic properties of Hilbert transform is that $H(H(x)) = -x$. Then complex function $z(t) = x(t) + i\, y(t)$ is eigen for Hilbert transform: $H(z) = -i\, z$. Let us write the function $z(t)$ in complex form $z(t) = x(t) + i\, y(t) = A(t)\, e^{i\theta(t)}$. Amplitude $A(t)$ is defined as function module $z(t)$, and instantaneous frequency of oscillations – by the formula $\omega = d\theta/dt$, where $\theta(t) = \arctan(y/x)$. In MATLAB there is a function $z = \text{hilbert}(x)$ for Hilbert transform of discrete signal X. Then formula

$$\text{instfreq} = \text{Fs}/(2*\text{pi})*\text{diff(unwrap(angle(z)))};$$

gives instantaneous frequency of the signal.

For ECG studying it is possible to use all listed numerical signs or to choose some of them. Selected numerical characteristics of ECG decomposing components form a vector of features $Y = [y_1, y_2, …, y_n]$ in n-dimensional feature space.

Reduction of the feature space. Let us suppose that wavelet decomposition is done and feature vectors $Y = [y_1, y_2, …, y_n]$ are made for some set of ECG records. Features space can have too big n dimension that complicates creation of classifiers, as it assumes working with high order matrixes. It is desirable to somehow reduce its dimension without essential loss of information. It is accepted to carry out decrease of dimension by means of linear display of all space of features on some smaller subspace [14, ch. 9] and [15, ch. 10]. It is performed by A matrix, in which the number of columns m is less than number of lines n so that the initial vector Y after linear decomposition $Z = A^T Y$ is projected onto the vector Z, which dimension m is significantly less (for example, 2 or 3, then it is possible to visualize classifiers in two- or three-dimensional space).

The reduction matrix A can be defined in several various ways. The main idea of these methods is in defining the direction in which dispersion of features' vector Y is the biggest by means of covariance matrix analysis. This direction is considered to be the most informative [14, ch. 9] and [15, ch. 10].

If Y data consist of several classes (for example, ECG features of ill and healthy patients), then it is necessary to choose such subspace of the most informative features which is the most effective from the point of view of classes' divisibility. Discriminant analysis procedure which at the same time gives dimensions and divides classes is based on scatter matrixes [14, ch. 9] and [15, ch. 10].

It can be described as follows: Let us suppose that we have N vectors of n-dimensional features Y_i, $i = 1, …, N$, as $Y_i = [Y_{i1}, Y_{i2}, …, Y_{in}]$. It is also supposed that elements of this data set can be divided into some number of c classes. In this case n is a number of extracted features after the wavelet analysis of ECG signal and N is a number of decomposed fragments of ECG data. The number of c classes in this case is two (ECG of normal patients and patients who had IM). In other words, all set of basic data $\{Y_i, i = 1, …, N\}$ can be divided into c subsets $\{Y_i^{(k)}, i = 1, …, N^k, k = 1, …, c\}$.

Scatter matrix S_w within the classes shows dispersion of features concerning vectors of expected value of classes:

$$S_w = \sum_{k=1}^{c} P_k E\{(Y^{(k)} - M_k)(Y^{(k)} - M_k)^T\} = \sum_{k=1}^{c} P_k \Sigma_k \,, (3)$$

where P_k is probabilities of getting into a certain class within all data set, $E\{\}$ is the operator of the expected value, $M_k = E\{Y^{(k)}\}$. In practice these values are approximated by selective estimates:

$$M_k = \frac{1}{N_k} \sum_{j=1}^{N_k} Y_j^{(k)} \,, \quad \Sigma_k = \frac{1}{N_k} \sum_{j=1}^{N_k} (Y_j^{(k)} - M_k)(Y_j^{(k)} - M_k)^T \,, \quad P_k = \frac{N_k}{N}, \quad k = 1, …, c \,. \tag{4}$$

Scatter matrix S_b between classes shows vectors dispersion of expected values around average value of mixture and is defined as follows:

$$S_b = \sum_{k=1}^{c} P_k (M_k - M_0)(M_k - M_0)^T \,, M_0 = \sum_{k=1}^{c} P_k M_k \,. (5)$$

For receiving criterion of classes divisibility and the choice of optimum features, some number which increases at increase of dispersion between classes or at reduction of dispersion within the class is linked to these matrixes. There are different approaches which consider these two requirements. In this article the following criterion is accepted [14, ch. 9]:

$$J_1 = tr\{S_w^{-1} S_b\} = \sum_{i=1}^{m} \lambda_i, \tag{6}$$

where $tr\{\}$ is a track of square matrix and λ_i are

eigenvalues of the matrix $S_w^{-1}S_b$. Now it is necessary to choose such subspace of features which maximizes criterion J_1. Let us consider the matrix $A = [\Psi_1 \Psi_2 \ldots \Psi_m]$ created as a set of columns $\Psi_1, \ldots, \Psi_m$ which are eigenvectors of matrix $S_w^{-1}S_b$, which correspond to m the biggest eigenvalues of this matrix: $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m \geq \lambda_{m+1} \geq \cdots \geq \lambda_n$. Then transition from full space of features to the given space of features is made by projection $Z = A^T Y$ on vectors $\Psi_1, \ldots, \Psi_m$ with the help of matrix A. Herewith some information is lost. Every eigenvector Ψ_i bears the amount of information which corresponds to the value of the corresponding eigenvalue λ_i. Therefore the relative measure of the saved information can be calculated as follows: $\sum_{i=1}^{m} \lambda_i / \sum_{i=1}^{n} \lambda_i \cdot 100\%$. Results of reduction of dimension with index of informational content which are higher than 85% are considered as satisfactory.

Despite some loss of information, reduction of dimension improves divisibility between classes and facilitates the task of classification. Besides, reduction of dimension allows to visualize results. It is very difficult to analyze any classifier in 21-dimensional space of features (as in this work) and it is almost impossible to present it.

During reduction $Z = A^T Y$ of features space the reducing matrix A of dimension of *n*-on-*m* mapping each of vectors Y sets of basic data on the corresponding *m*-dimensional vector Z. Therefore, vectors Z are also divided into c subsets $\{Z_i^{(k)} = A^T Y_i^{(k)}, i = 1, \ldots, N^k, k = 1, \ldots, c\}$. Now the task is in creation of classifiers, i.e. of functions which divide all these subsets.

Linear classifiers. The algorithm of classification assumes the division of patients into two groups (healthy and those who had MI). Therefore it is possible to use linear classifiers. Let us remind their construction [15, ch. 4, 10]. The linear classifier has the form of linear heterogeneous function $h(Z) = V^T Z + v_0$, where Z is *m*-dimensional vector of data received after dimension reduction, V is vector of coefficients and v_0 is a constant term. If $h(Z) > 0$, then Z belongs to the first class ω_1, and if $h(Z) < 0$, then Z belongs to the second class ω_2.

Expected values and dispersions of function $h(Z)$ for each class ω_i are set by formulae:

$$\eta_i = E\{h(Z)|Z \in \omega_i\} = V^T E\{Z|Z \in \omega_i\} + v_0 = V^T M_i + v_0,$$
$$\sigma_i^2 = Var\{h(Z)|Z \in \omega_i\} = V^T E\{(Z-M_i)(Z-M_i)^T|Z \in \omega_i\}V = V^T \Sigma_i V. \tag{7}$$

For finding optimum values V and v_0 criterion in the form of some function is usually used $f(\eta_2, \eta_2, \sigma_1^2, \sigma_2^2)$, its critical points define required optimum values. One of widespread criteria is [15, ch. 4, 10]:

$$f = \frac{P_1 \eta_1^2 + P_2 \eta_2^2}{P_1 \sigma_1^2 + P_2 \sigma_2^2}, \tag{8}$$

This function measures the dispersion between classes (around zero) normalized by dispersion within a class. Then optimum values V and v_0 turn out in [15, ch. 4]:

$$V = [P_1 \Sigma_1 + P_2 \Sigma_2]^{-1}(M_2 - M_1) \text{ and}$$
$$v_0 = -V^T (P_1 M_1 + P_2 M_2)) \tag{9}$$

RESULTS AND DISCUSSION

For creation of the classifying system ECG records of two groups of patients were used. The first group of ECG records of healthy patients contains 96 ECG fragments for four patients aged from 21 to 27. The second group of ECG records of patients who have recently came through myocardial infarction (subacute period) contains 120 ECG fragments for five patients aged from 44 to 55. Let us remind that the initial cardiosignals were 30 seconds long with sampling rate of 1028 counts per second. For wavelet decomposition and calculation of features two fragments of signal 8 seconds long each were chosen from each ECG record.

Wavelet decomposition. In the wok the orthogonal Meyer wavelet dmey is used, which is derived from Meyer wavelet [10] of infinite impulse response by truncation of its filter to 102 members. It has the carrier on the interval [0,101] and central frequency $Fr = 0.6634$ Hz. The choice of this wavelet is explained by well localization of frequency spectra of signal components. The point is that this wavelet has the widest frequency spectrum among orthogonal wavelet with the compact carrier. In it the frequencies which are in rather large surrounding area of its center frequency are equally provided 0.6634 Hz. For this reason it provides well expansion of the signal into the items corresponding to certain frequency bands. As ECG sampling rate is 1028 counts per second, in the spectrum of the digitized signal frequencies up to 514 Hz will be provided. Therefore in case of the first level of Meyer wavelets' decomposition it will single out signal elements with the highest frequencies close to center frequency of the first level of decomposition, equal to $Fr_1 = 0.6634*514 = 340.99$ Hz. In case of the second and following levels these frequencies decrease sequentially twice.

Wavelet decomposition. Decomposing of ECG signal to the 4th level is made: $S \rightarrow \{D_1, D_2, D_3, D_4, A_4\}$ (fig. 2). For signal components we have:

Fragment = $RecD_1 + RecD_2 + RecD_3 + RecD_4 + RecA_4$.

Frequency spectrum of power of the first component $RecD_1$ is concentrated within the limits of 220 to 350 Hz, for the second component $RecD_2$ within the limits of 120 to 200 Hz, for the third component $RecD_3$ – of 60 to 90 Hz and for the fourth one – of 25 to 70 Hz (fig. 3). The choice of the 4th level of decomposition is explained by the fact that low-frequency component $RecA_4$ represents the undistorted smoothed ECG signal cleared of high-frequency oscillations (fig. 2). In case of more deep expansion of

the signal the following high-frequency component $RecD_5$ has frequency spectrum to 30 Hz, and the low-frequency component $RecA_5$ is significantly distorted. Let us remind that one of the purposes of this work is to show that high-frequency numerical characteristics can be successfully used in ECG classification. Wavelet decomposition of a signal fragment is made by the following MATLAB command:

[c,l] = wavedec(Fragment,4,'dmey');

As a result there is the structure [c,l], which contains a set of wavelet-coefficients $\{D_1, D_2, D_3, D_4, A_4\}$, where D_1, D_2, D_3, D_4 are the coefficients of details and A_4 is the approximating coefficients. For restoration function wrcoef MATLAB Wavelet Toolbox is used. It allows to restore both high-frequency $RecD_i$ and low-frequency $RecA_i$ components of the signal on the structure [c,l] of wavelet components $\{D_1, D_2, D_3, D_4, A_4\}$ received earlier [10]:

for s=1:4
RecD(s,:)=wrcoef('d',c,l,w,s);
End.

Feature space and its reduction. As it was noted in section Feature space, for every high-frequency ECG decomposing component $RecD_1$, $RecD_2$, $RecD_3$, $RecD_4$ it can be calculated to 10 various statistical, frequency and stochastic characteristics. In total there are 40 features for ECG. In the course of work those features which influence was very little (i.e. the corresponding elements of reduction matrix A are small) and those features using of which gave bad results of groups' division have been removed. As a result of such analysis it has turned out that the most suitable feature set for classification aims is the following:

– maximum absolute component value;
– L^2-energy of the component;
– maximum value of the power spectrum of the component;
– frequency, where maximum value of the power spectrum is achieved;
– Shannon's entropy;
– Hurst exponent (only for $RecD_4$ component).

These features are calculated for every component $RecD_1$, $RecD_2$, $RecD_3$, $RecD_4$. As a result there are 21 features for one ECG record. The specified features form the vector of 21-dimensional space of features. Because of such high dimension it is inconvenient to conduct ECG researches taking into account all features. The reduction procedure described in section 1.2 will be applied. Let us calculate scatter matrixes S_w and S_b according to the formulae (3) – (6), then find eigenvalues λ_i, $i = 1, \ldots, 21$ and eigenvectors λ_i, $i = 1, \ldots, 21$ of $S_w^{-1}S_b$ matrix. For this procedure in MATLAB there are [12] function:

[Psi,Lambda] = eig(inv(Sw)*Sb).

Eigenvalues Lambda are arranged in descending. Eigenvectors Psi are normalized. According to the procedure stated above it is necessary to choose the subspace of features which is formed by eigenvectors $\Psi_1, \ldots, \Psi_m$ with the greatest values of eigenvalues λ_i, $i = 1, \ldots, m$.

As a result of calculations it has turned out that value of the first eigenvalue λ_1 is approximately equal to 3.4124, and other eigenvalues have the exponent 10^{-13} and lower. That is why the reduced space of features is one-dimensional and formed by eigenvector (column) Ψ_1. Reduction matrix A consists of one column, $A = [\Psi_1]$ and affects the full vector of features $Y = [y_1, y_2, \ldots, y_{21}]^T$ as projection $Z = A^T Y$ to the one-dimensional space generated by vector Ψ_1 of single length. Relative measure of the saved information $\lambda_1 / \sum_{i=1}^{21} \lambda_i \cdot 100\%$ at dimension reduction makes approximately 100%.

During projection of features vectors to one-dimensional space it has turned out that value Z accepts values ranging from $-2.6 \cdot 10^{-4}$ to $-0.97 \cdot 10^{-4}$. Values' distribution of Z feature is shown on histograms (Fig. 4) separately for the first group of healthy and the second group of ill patients.
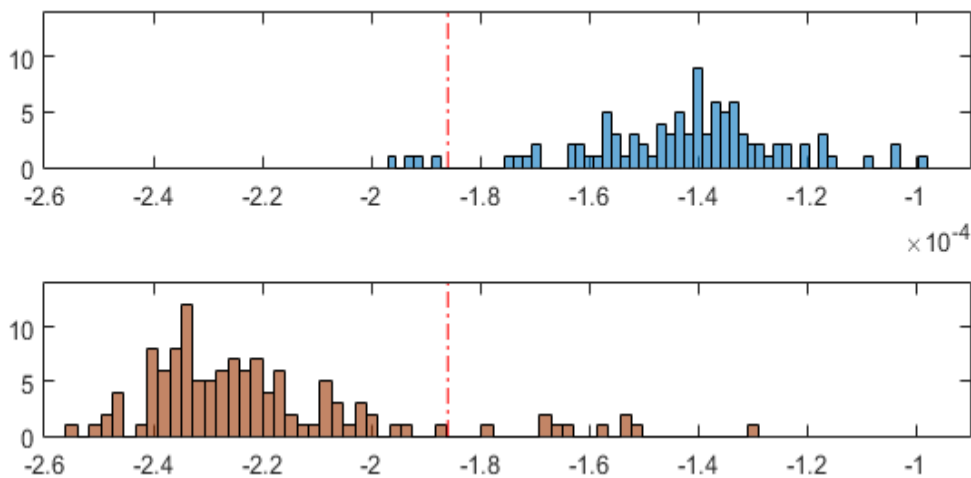

Fig. 4. Values histograms of Z feature of examined groups of healthy (top) and ill (bottom) patients in the given one-dimensional space of features. The vertical line is determined to be the dividing constant $z_0 = -1.86 \cdot 10^{-4}$ of linear classifier.

Linear classifier construction. As it was determined in the previous section, the reduced feature space is one-dimensional and can be derived by projection $Z = \Psi_1^T Y$ in one-dimensional space generated by vector Ψ_1. The reduced vector of features Z is a scalar. That is why the linear classifier $h(Z) = V^T Z + v_0$ takes the form of $h(Z) = VZ + v_0$, where V is coefficient and v_0 is constant term. If $h(Z) > 0$, then Z belongs to the first class ω_1, and if ω_1, then Z belongs to the second class ω_2. As it comes only to the sign $h(Z)$ it is convenient to study this classifier as

$$h(Z) = Z + v_0/V = Z - z_0,$$

where $z_0 = -v_0/V$. For criterion $f = (P_1\eta_1^2 + P_2\eta_2^2)/(P_1\sigma_1^2 + P_2\sigma_2^2)$ by formulae (4), (7) and (9) optimum coefficients V and v_0 are calculated and divining constant z_0 is found. As a result, $z_0 = -0.0001859$. In figures 4 and 5 this constant is represented by the vertical red dash-dotted line.

Testing of the classifying system. For testing 96 ECG fragments of four healthy patients aged from 21 to 56 and 120 ECG fragments of five ill patients aged from 45 to 57 who recently came through myocardial infarction (subacute period) are used. For these groups features vectors are created and space reduction of features to one-dimensional space is made, formed by eigenvector Ψ_1 for eigenvalue λ_1. Relative measure of the saved information $\lambda_1 / \sum_{i=1}^{21} \lambda_i \cdot 100\%$ during reduction of the dimensionality is approximately 100%. For classification the dividing value $z_0 = -0.0001859$ received in the previous section is used.

During vectors reduction of features to one-dimensional space it has turned out that value Z take values in the range from $-2.5 \cdot 10^{-4}$ to $-0.8 \cdot 10^{-4}$. Values distribution of Z sign for groups of the tested patients is shown on histograms in Fig. 5.

The data submitted in fig. 5 show that only three features (3%) of 96 given features values of the first group are carried by the classifier to group of ill patients (there are values which are smaller than the dividing constant z_0) and only 20 signs (<17%) of 120 given features values of the second group are carried by the classifier to ECG group of healthy patients (there are values which are bigger than the dividing constant z_0). Let us remind that during ECG record of the patient about 12 assignments are registered. For each assignment the classifier defines whether the ECG assignments belong to ECG-healthy type of patients, or it has some properties typical of the patients who had MI. When testing it has turned out that some healthy patients have assignments features of which are closer to features of ill patients ($h(Z) < 0$). However the average value of all 12 leads of the classifying function $h(Z)$ for all healthy patients participating in testing is higher than zero. The same with ill patients, average value of the classifying function $h(Z)$ of all 12 leads is lower than zero for all ill patients participating in testing. In this sense the classifier is accurate. Besides, it defines "problematic" leads where $h(Z)$ differs from the majority of values in other leads. For every patient the set of 12 classifier values $h(Z)$ in every lead can serve as additional characteristics of patient's condition.

For example, values of qualifier $h(Z) = Z - z_0$ are calculated for 12 standard leads of the personal cardiogram of the second author of this work (the values multiplied by 10^3 are given below):

−0.0561 −0.2441 −0.7162 0.5211 −0.0255
−0.5415 0.1347 −0.1841 −0.0558 −0.4035 0.5141
−0.3891.

The classifier defines ECG problems in 9 leads. Therefore this patient (the second author) does not belong to the group of healthy patients, and it is true. At the same time, the average value equal to $-0.1205 \cdot 10^{-3}$ is close to zero, therefore there are no bases to put this patient into the second group of the patients who had MI. Also, the cardiogram of the first author is studied. In this case values of the classifying function for all 12 leads of ECG are positive. The first author is attributed by the qualifier to "healthy" group, and it is also true.
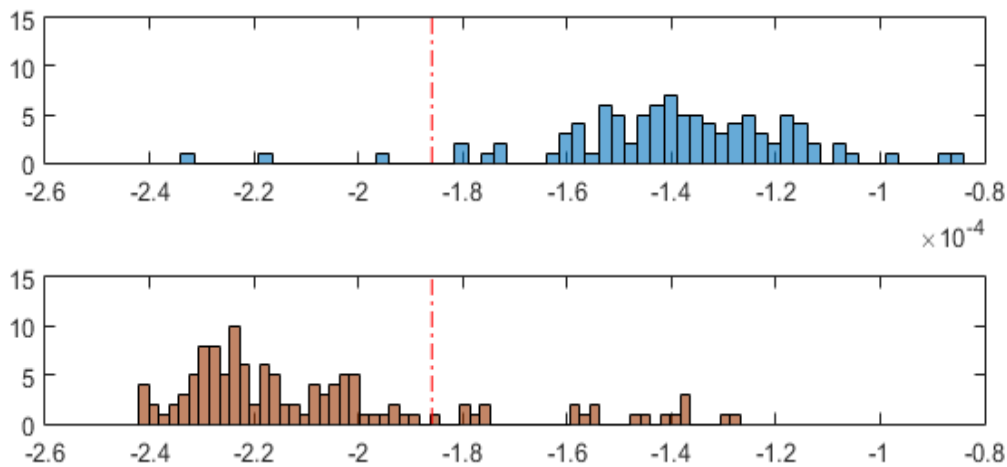

Fig. 5. Values histograms of Z feature of patients when testing. At the top is features histogram of healthy patients. At the bottom is features histogram of ill patients. The vertical line is the dividing constant $z_0 = -1.86 \cdot 10^{-4}$ of linear classifier.

CONSCLUSION

The classifying system which divides the chosen groups of patients rather reliably is constructed. Feature of this classification system is that it uses only high-frequency ECG components. Positive results of testing show that high-frequency ECG components carry essential diagnostic information concerning ECG. Such system can be used as addition to classification systems on the basis of the analysis of waves complex and PQRSTU segments for more exact and informative division of patients.

REFERENCES

1. Aldroubi A., Unser M. *Wavelets in Medicine and Biology*. CRC Press, 1996. 640 p.
2. Nagendra H., Mukherjee S., Kumar V. Application of Wavelet Techniques in ECG Signal Processing: An Overview. *International Journal of Engineering Science and Technology (IJEST)*, 2011, vol. 3, no.10, pp. 7432–7443.
3. Pavlov A.N., Khramov A.E., Koronovskiy A.A., Sitnikova E.Yu., Makarov V.A., Ovchinnikov A.A. Veyvlet-analiz v neyrodinamike [Wavelet analysis in neurodynamics]. *UFN* [UFN], 2012, vol. 182, no. 9, pp. 905–939.
4. Sambhu D., Umesh A. C. Automatic Classification of ECG Signals with Features Extracted Using Wavelet Transform and Support Vector Machines. *IJAREEIE*, 2013, vol. 2, no. 1, pp. 235–241.
5. Shahbahrami A., Kiani M. Classification of ECG arrhythmias using discrete wavelet transform and neural networks. *IJCSEA*, 2012, vol.2, no.1, pp. 1–13.
6. Ishikawa Y., Mochimaru F. Wavelet Theory–Based Analysis of High-Frequency, High-Resolution Electrocardiograms: A New Concept for Clinical Uses. *Progress in Biomedical Research*, 2002, vol. 7, no. 3, pp. 179–184.
7. Ishikawa Y. *Wavelet Analysis for Clinical Medicine. Chapter 6: SAECG (Signal Averaged ECG) which was seen from Wavelet Analysis – Supplement – original color images*. Available at: http://www.uinet.or.jp/~ishiyasu/ch6/index.html.
8. Mochimaru F., Fujimoto Y. Detecting the Fetal Electrocardiogram by Wavelet Theory-Based Methods. *Progress in Biomedical Research*, 2002, vol. 7, no. 3, pp. 185–193
9. Podkur P.N. O vysokochastotnykh komponentakh EKG [About high frequency components of ECG]. *Materialy sed'moy vserossiyskoy nauchno-praktieskoy konferentsii "Novye dostizheniya v razvitii elektrokardiografii"* [Materials of the seventh all-Russian scientific-practical conference "New achievements in the development of electrocardiography"]. Tyumen, 2005, pp. 123–126.
10. Smolentsev N.K. *Osnovy teorii veyvletov. Veyvlety v MATLAB* [Fundamentals of the theory of wavelets. Wavelets in MATLAB]. Moscow: DMK Press, 2013. 628 p.
11. Vorobyov A. S. *Elektrokardiografiya. Noveyshiy spravochnik* [Electrocardiography. The newest guide]. Moscow: Publ. house Eksmo; Saint Petersburg: Sova Publ., 2003. 560 p.
12. Smolentsev N.K. MATLAB: *Programmirovanie na Visual S#, Borland JBuilder, VBA* [MATLAB: programming in Visual C#, Borland JBuilder, VBA]: Moscow: DMK Press, Saint Petersburg: Piter Publ., 2009. 456 p.
13. Adeli H., Ghosh-Dastidar S., Dadmehr N. A Wavelet-Chaos Methodology for Analysis of EEGs and EEG Sub-Bands to Detect Seizure and Epilepsy. *IEEE Transactions on Biomedical Engineering*, 2007, vol. 54, no. 2, pp. 205–211.
14. Fukunaga K. *Introduction to Statistical Pattern Recognition*. Moscow: Nauka Publ., 1979. 368 p.
15. Fukunaga K. *Introduction to Statistical Pattern Recognition*. Academic Press, Boston, 1990. 592 p.

Paulina N. Podkur

Cand.Sci.(Phys.-Math.), Associate Professor, Associate Professor of the Department of Higher and Applied Mathematics, Kemerovo Institute (branch) of Plekhanov Russian University of Economics, Kemerovo, Russian Federation.

Nikolay K. Smolentsev

Dr.Sci.(Phys.-Math.), Professor, Professor of the Department of Fundamental Mathematics, Kemerovo State University, Professor.