The Method of Analysis Pseudo-phase Portrait in the Problem of Recognition of Biomedical Signals

Karina Khachatryan, Lyudmila Manilo, Aleksei Anisimov
Saint-Petersburg State Electrotechnical University "LETI"
St. Petersburg, Russia
karinasmbatovna@gmail.com, lmanilo@yandex.ru, aaanisimov@etu.ru

Abstract—**Today the problem of analysis of biomedical signals with chaotic properties is an important task. Its decision is important for the recognition of signals with varying severity of non-linear component. The paper presents a method for biomedical signal processing, tapping on the pseudo-phase portrait differences from the normal rhythm of some of its violations. The algorithm is intended for medical computer systems and implemented in MATLAB software environment.**

I. Introduction

Currently, there is growing interest in the problem of the analysis of biomedical signals with chaotic properties, which is especially important in recognizing conditions associated with the assessment of the dynamics of the generating systems. These tasks include the study of brain electrical activity during different states of a living organism, cardiac rhythm classification on ECG signal or rhythmogram, and others [1]. Through the methods of nonlinear dynamics, a series of indicators can be obtained. These series can be used for evaluation of properties of chaotic signals and for implementation of their recognition.

In this paper there are presented the methods for the analysis of biomedical signals, based on a study of the correlation dimension of rhythmograms signals and calculating the morphological characteristics of the pseudo-phase portrait. All experiments were carried out on real data, including recording rhythmograms of normal rhythm and its disorders. In Section II, there is the method of finding the correlation dimension using an algorithm Grassberger-Procaccia. Results for calculating the correlation dimensions of the signal's rhythmograms are presented. In Section III, the method of constructing a pseudo-phase portrait is suggested. A description of the proposed morphological signs of pseudo-phase portrait, and the results of their calculations, including statistical processing of data, are provided. In Section 4, the estimation of the division of classes using the Fisher's ratio test, derived decision rules of classification and recognition errors are estimated.

II. Analysis of Correlation Dimension Rhythmograms Signals

An analysis of the correlation signals dimension allows us to estimate the degree of complexity of their generatingsystems and to assess the dimension of pseudo-phase portrait. To conduct the study records three groups were formed, each of which contained 10 patients rhythmogram. Groups included different types of cardiac rhythm: normal rhythm (NR), ventricular extrasystoles (VE) and atrial fibrillation (AF). Rhythmograms examples are shown in Fig. 1: *a* – NR, *b* – VE, *c* - AF. The realization of the algorithm implemented in MATLAB programming environment.

To find the correlation dimension, Grassberger-Procaccia algorithm was used, which is one of the most effective methods of time series analysis [2], [3]. The essence of the algorithm lies in the calculation of the correlation integral *D(m)* for different values of *m* dimensions, depending on the construction obtained in a coordinate system with a double logarithmic scale, finding in it the linear section and calculating the angular coefficient *d*. The resulting value of d is an estimate of the correlation dimension of a regular sampling of samples.

$$x(i) = y_1, x(i+\tau) = y_2, ..., x(i+(m-1)\tau) = y_m$$

wherein data is measured at a fixed time interval τ.

In this paper we used the embedding values *m* from 1 to 9. In processing the original rhythmogram great length (250 samples) one of the m-length sequences is fixed and the distance between vectors $\mathbf{X}^{(i)}$ and $\mathbf{X}^{(j)}$ for any pair of vectors is determined

$$\left|\mathbf{X}^{(i)} - \mathbf{X}^{(j)}\right| = \left[\sum_{k=1}^{m} (x_{i-m+k} - x_{j-m+k})^2\right]^{1/2}.$$

By selecting a specific value of embedding dimension, it is possible to calculate the correlation integral:

$$C(m) = \lim_{N\to\infty} \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} H(r - \left|\mathbf{X}^{(i)} - \mathbf{X}^{(j)}\right|)$$

Where $i \neq j$, $H(\cdot)$ - Heaviside function ($H(x) = 0$ if $x<0$; $H(x) = 1$ if $x \geq 0$), *r*- threshold distance, *N* - the original sample length (rhythmogram length).

Correlation dimension can be defined as the slope of the linear section the correlation integral, presented in a logarithmic scale:

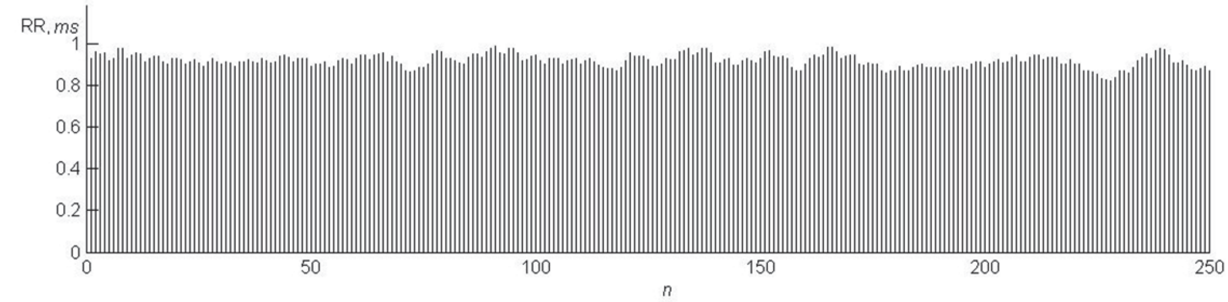
$$D(m) = \frac{\ln C(m)}{\ln(r)}$$

Calculations can be performed in decimal or binary logarithmic scale.

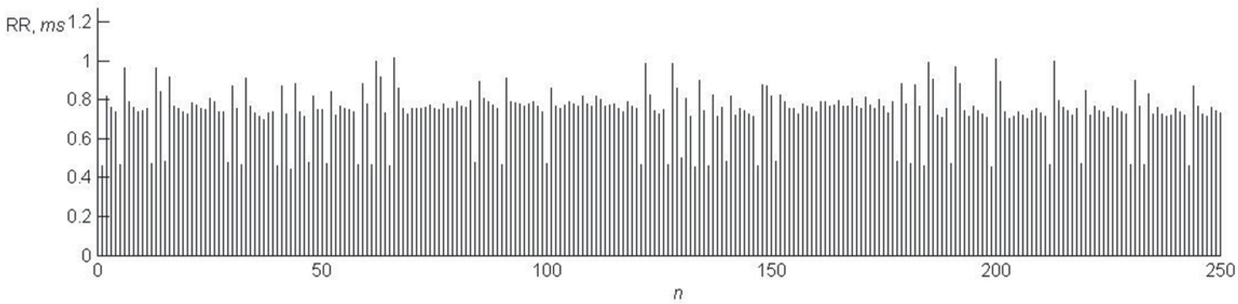
Fig. 2 shows the correlation integral curves corresponding to the different types of cardiac rhythm: *a* – NR, *b* -VE, *c* - AF.

Results of calculating the correlation dimension depending on the dimension of attachment are shown in Table I. The results are shown graphically in Fig.3: *a* – NR, *b* – VE, *c* - AF.

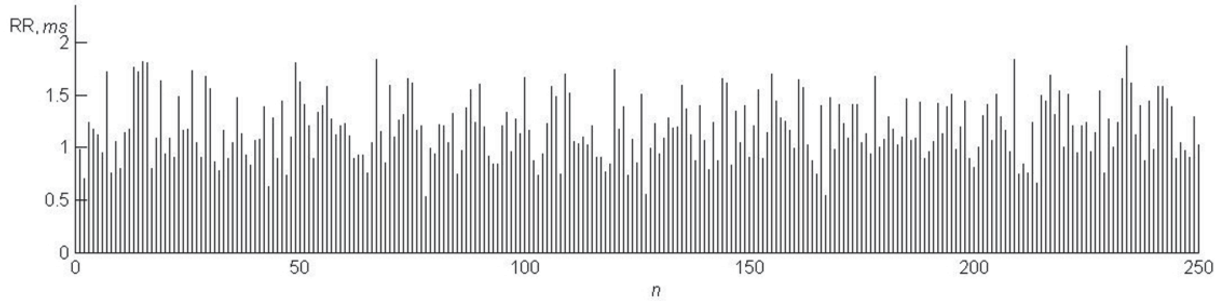
From Fig. 3, it follows that the saturation curves for NR and VE observed in value *D(m)* =4, so their correlation dimension is 4, and the dimension of the embedding space 7 and 6, respectively. When AF the correlation dimension is 5, and the dimension of the embedding space - 6. To solve the problem of classification, taking into account differences in the minimum values of *m* for the three types of heart rate, you can reduce the dimension of the embedding space to two and analyze features of pseudo-phase portrait on the plane.



a

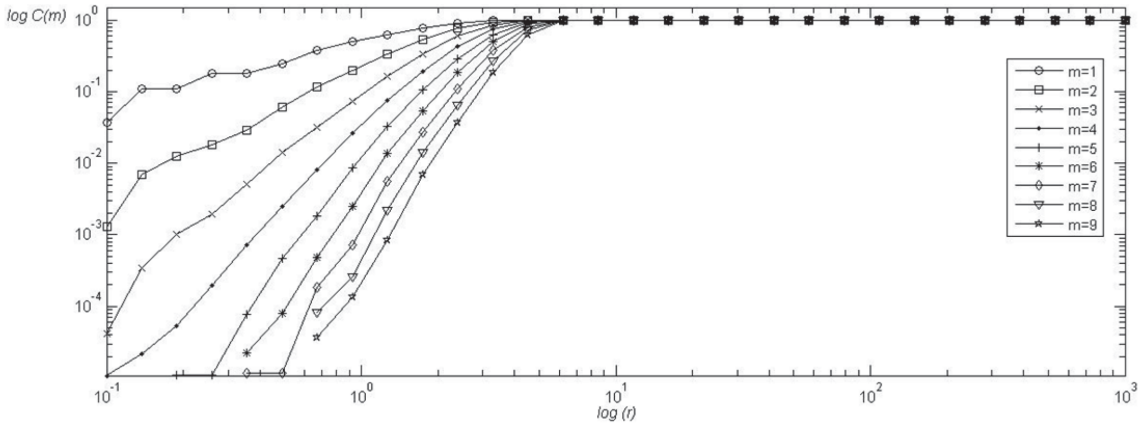


b

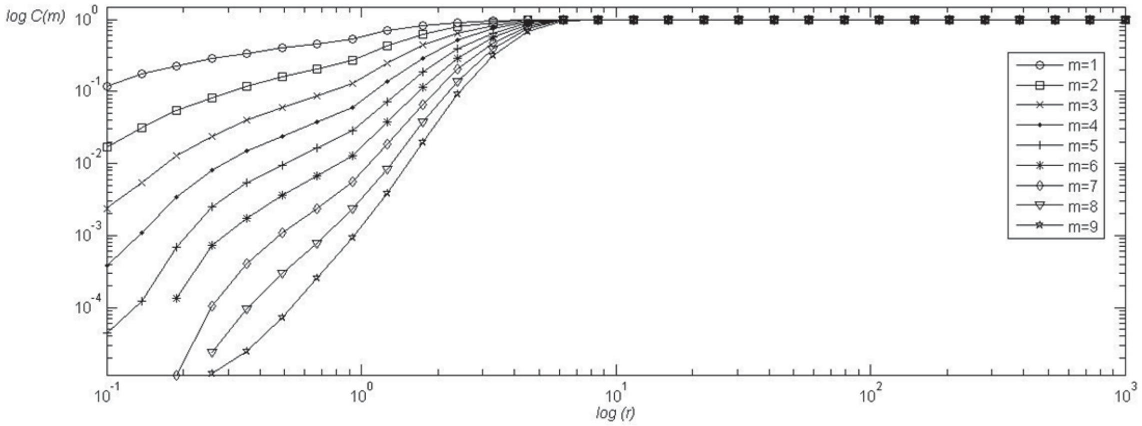


c

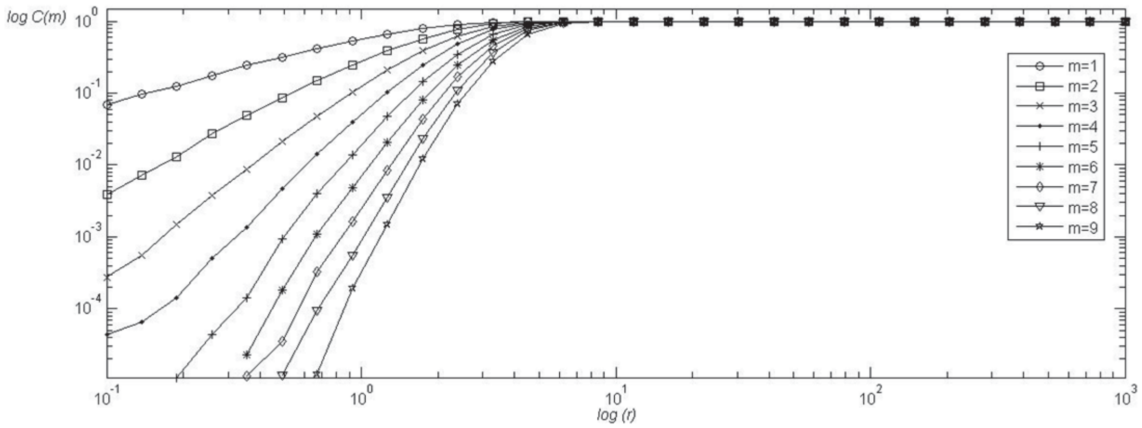
Fig. 1. Rhythmograms examples: *a* – NR, *b* – VE, *c* - AF



a



b



c

Fig. 2. The correlation integral curves: *a* – NR, *b* – VE, *c* - AF

TABLE I. THE RESULTS OF CORRELATION DIMENSION D(M) CALCULATION

Rhythm types	**Embedding dimension *m***								
	1	2	3	4	5	6	7	8	9
Normal rhythm	0,669	1,374	2,150	2,873	3,565	3,630	3,979	4,069	4,095
Ventricular extrasystoles	0,241	0,577	0,977	1,428	2,901	3,789	3,873	3,987	4,019
Atrial fibrillation	0,388	0,878	1,425	2,015	3,666	4,915	5,074	5,126	5,229

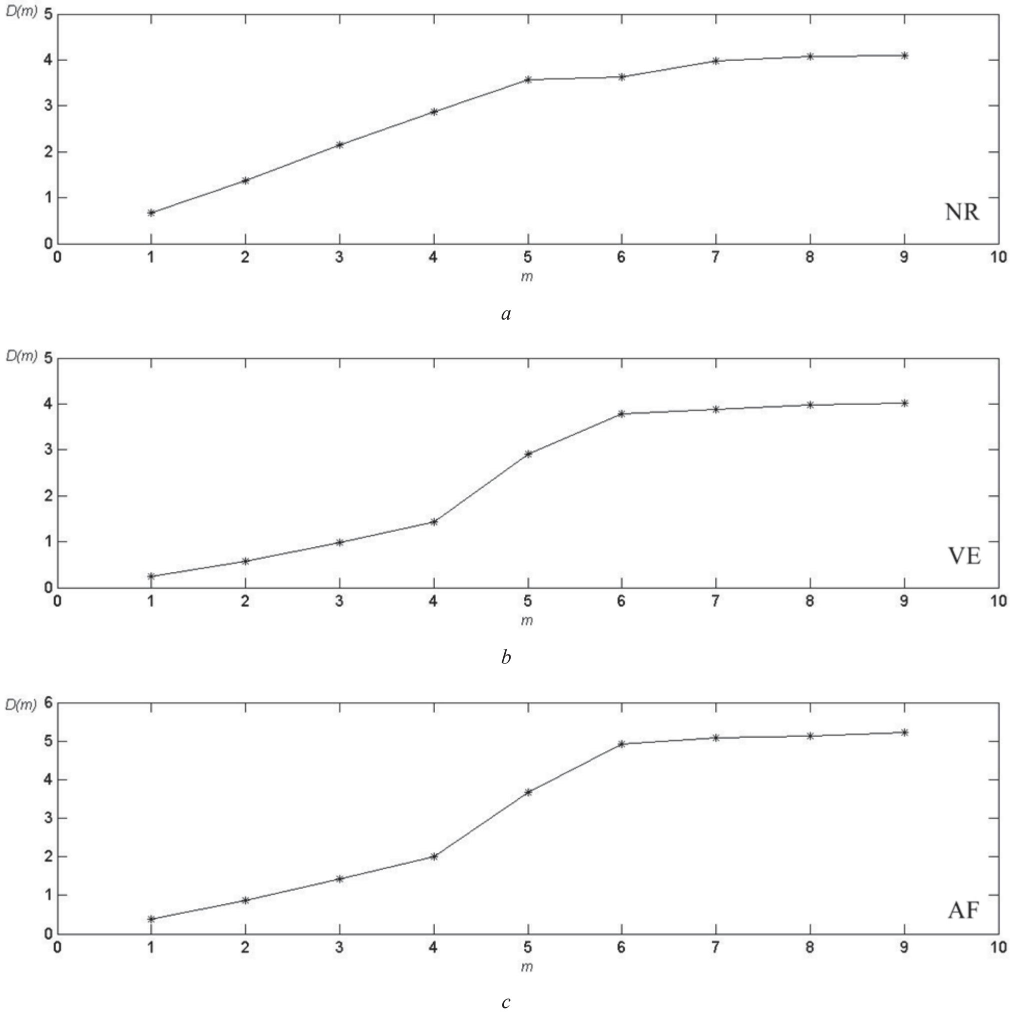


Fig.3. The dependence of the correlation dimension of the embedding dimension: *a* – NR, *b* – VE, *c* - AF

III. Analysis of morphological characters of pseudo-phase portrait

The phase plane - coordinate plane in which the axes are deposited phase coordinates, clearly defining the state of the second-order system. This plane is a special case of phase space which can have large dimensions. In the phase space of the state arbitrarily complex system represented a single point, and the "evolution" of this system - the trajectory of movement of this point [4], [5].

In the article [6] the geometric method of signal analysis is discussed (a pseudo-phase portrait construction), which is today a radically new technique. The author proposes the nonlinear dynamics of the signal analysis by constructing a broken line in a coordinate system where the x-axis shows the duration of received signals in the range of $x_1,...,x_{n-1}$, and the y-axis in the range of $x_2,...,x_n$. In order to evaluate the morphological features of pseudo-phase portrait of the heart rate is selected a number of parameters that allow to differentiate analyzed different types of signals. In this paper we study the seven parameters that describe the features of a pseudo-phase portrait.

As one of the parameters, the contour length (perimeter) of the phase portrait is selected. Suppose we have a sequence of N intervals. To construct a profile of pseudo-phase portrait in a convex hull in the MATLAB function is *convhull*. The input vectors are x-coordinate and y-coordinate points, and output arguments are the number of points that form the vertices of the convex hull. Non-points are the vertices of the convex hull are entered into an array *kvert*:

```
x=RR(1:h-1);

y=RR(2:h);

kvert=convhull(x,y).
```

Points forming the convex hull of a pseudo-phase portrait marked circular markers and drawn itself contour. Knowing the coordinates of points forming the contour of the phase portrait, the length can be found by calculating the length of the segments sequence. The length of the *i*-th segment constituting the circuit pseudo-phase portrait is:

$$L_i = \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2}\ , \qquad (1)$$

where (x_i, y_i), (x_{i+1}, y_{i+1}) - the coordinates of the start and end of a segment.

A parameter equal to the area bounded by the pseudo-phase portrait was introduced, because it may be different for the same length. This parametr is also calulated using *convhull* function, but this function must be called with the additional autput argument *s*, that would contain the area of a convex hull:

[kvert, s] = convhull(x, y),

where $x_i = RR_i, y_i = RR_{(i+1)}$.

Due to the fact that the signals have a different scale, introduced an additional characteristic (coefficient):

$$K = \frac{P^2}{S}\ ,$$

where P - the length of loop pseudo-phase portrait, S - the area bounded by its contour.

As a fourth parameter characterizing the morphology of the pseudo-phase portrait, the length of the broken line forming the phase portrait was selected. This parameter characterizes the extent of its spread. For the less regular signals the points that are used for constructing of the portrait are arranged far away from each other. Knowing the coordinates of points forming a polyline, you can easily find its length in accordance with the formula (1).

Also, it is advisable to make an estimate of the average length of lines, forming a pseudo-phase portrait. Dividing the length of broken line, forming a pseudo-phase portrait on the number of direct components of it, we find the average length.

The sixth parameter was proposed to use a quantitative estimate of the "leap" (transition from large to small lengths). The "leap" means the lines which differential value from the current point to the following is higher than one-third of the current value:

$$((x(i) - x(i+1)) > x(i)/3.$$

As a final morphological characteristic number of points of figures is selected, falling in the range from 0 to 45 degrees. The abscissa of the coordinate plane to determine the selected range passes through the figure lowest point, and the ordinate through the figure center of mass (Fig. 4). This angular range has been selected due to the fact there is a significant point scattering in irregular signals and this area includes pseudo-phase portrait of the rhythms that have more expressed chaotic component than NR. Using this characteristic, the normal rhythm can be easily distinguished from the VE and AF pathologies.

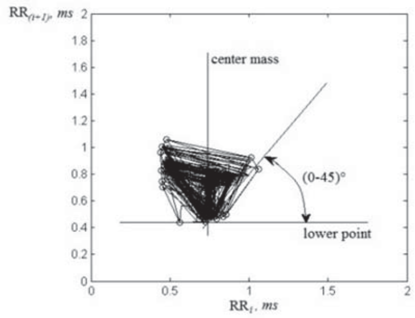
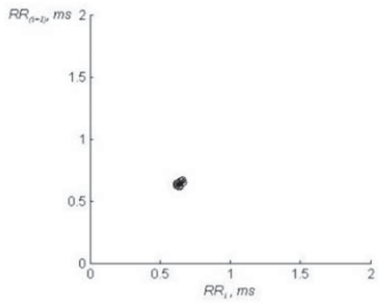
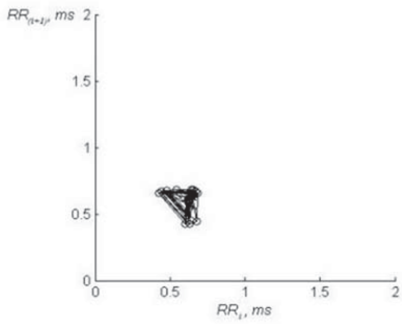


Fig. 4. Example of selection an angle

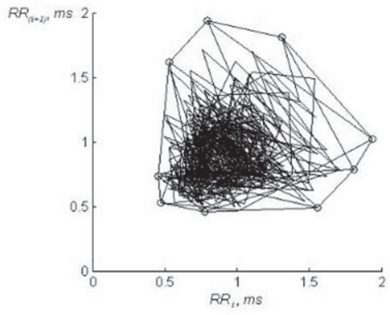
Examples of the pseudo-phase portraits for each rhythmogram shown in Fig. 5: a – NR, b – VE, c - AF. The figure also shows the contour bounding the phase-portrait.



a



b



c

Fig. 5. The obtained pseudo-phase portraits: *a* – NR, *b* – VE, *c* – AF

The results of statistical processing of the measured characteristics numerical values are shown in Table II.

After analyzing the data, we can conclude that all values except the space are clearly different for the rhythmogram types analyzed. The more chaotic components in the signal (typical for AI), the more scatter of phase points on the plane, and greater the perimeter and the length of the phase portrait loop. This analysis is useful for the study of the morphological features of biosignals at pseudo-phase plane and assess the recognition possibility of certain types of arrhythmias by the means of pseudo-phase portrait construction.

TABLE II. THE VALUES OF MORPHOLOGICAL CHARACTERS FOR THE THREE TYPES OF HEART RHYTHM

Number of sign	Parameters pseudo-phase portrait	Types of rhythm		
		Normal rhythm	Ventricular extrasystoles	Atrial fibrillation
1	The perimeter of the loop	0,3±0,1	1,0±0,2	2,9±0,3
2	Contour area	0,2±0,1	0,3±0,0	0,5±0,1
3	Coefficient	1,8±0,5	4,6±1,2	17,4±3,5
4	The length of the broken line	1,1±0,2	8,6±2,1	14,8±2,4
5	The average length of lines	0,05±0,02	0,28±0,04	0,57±0,10
6	The number of " leap "	0,00±0,00	9,20±3,74	47,3±8,81
7	The number of points in the range (0-45)°	0,00±0,00	26,7±6,21	52,1±12,21

IV. CLASSIFICATION OF DATA OBTAINED BY FISHER'S RATIO TEST

For a two-class problem of linear discriminant function (LDF) or the separating hyperplane, which determines the recognition algorithm is as follows:

$$\mathrm{D}(\mathrm{x}) = \mathbf{W}^{\mathbf{T}}\mathbf{X} - \alpha = 0$$

$$\text{or} \quad \mathbf{W}^{\mathbf{T}}\mathbf{X} = \alpha ,$$

W- where the weight vector of unit length, α - scalar threshold.

Recognition algorithm in this case is as follows:

$$\text{if } \mathbf{W}^{\mathbf{T}}\mathbf{X} < \alpha \text{, то } \omega_1 \text{ (1 class),}$$

$$\text{if } \mathbf{W}^{\mathbf{T}}\mathbf{X} \geq \alpha \text{, то } \omega_2 \text{ (2 class).}$$

To determine the LDF must find **W** and α. One way of solving this problem is to implement the following algorithm.

1) Find **W** as a best position of this vector in n-dimensional space, for which the projection of points of the classes on the W direction is best separated;
2) To design both points of classes on the straight line defined by the vector **W**;
3) Solve the one-dimensional problem of finding the value of α, for example, by the criterion of minimum average classification errors [7].

Solution of the 1st step is given by

$$\mathbf{W} = \mathbf{S}_{\mathrm{W}}^{-1}(\mathbf{M}_1 - \mathbf{M}_2),$$

where $\mathbf{M}_1$ and $\mathbf{M}_2$ - vector of mean values two classes, $\mathbf{S}_{\mathrm{W}} = \mathbf{S}_1 + \mathbf{S}_2$ - total scatter matrix for the two classes.

Scatter matrix for the j-th class (analogue of correlation matrix) is defined by the formula:

$$\mathbf{S}_j = \sum_{i=1}^{n_j} (\mathbf{X}_i^{(j)} - \mathbf{M}_j)(\mathbf{X}_i^{(j)} - \mathbf{M}_j)^T \quad j = 1, 2 .$$

The determined weight vector W corresponds to the maximum Fisher's ratio test J:

$$J(\mathbf{W}) = \frac{(m_1 - m_2)^2}{s_1^2 + s_2^2},$$

where m_1 and m_2 – average values; s_1^2 и s_2^2 the value of the scatter of the projection on the axis solutions (y=$\mathbf{W}^{\mathbf{T}}\mathbf{X}$) samples.

As can be seen, this criterion provides maximum ratio the scatter between classes to the "average" spread within the classes. For the vector **W**, found by the criterion J, classes, projected on the direction **W**, the maximum distance from each other.

In practical calculations, an averaged covariance matrix can be used instead $\mathbf{S}_{\mathrm{W}}$ matrix.

To assess the quality classification of NR, VE and AF, 3 groups of data were formed, each containing 7 morphological characters received pseudo-phase portrait of ten patients. At NR's recognition in the alternative class of pathology (P) were integrated implementation of two classes VE and AF. At the second stage the same way solved the problem of classification of objects of class P combined into two subclasses: VE and AF.

Evaluation results of three classes rhythmograms distributions are presented in Fig. 6: *a* - class of normal rhythm and pathology; *b* - the class of ventricular extrasystoles and atrial fibrillation. They are presented as histograms objects projection distributions on vector **W**, and where *a* - the separation threshold.

Based on the data, the following critical classification rules were derived.

Recognition of NR and pathology classes:

$$\mathbf{W}_1^{\mathbf{T}}\mathbf{X}_i \geq a_1 \rightarrow \text{Normal rhythm}$$

$$\mathbf{W}_1^{\mathbf{T}}\mathbf{X}_i < a_1 \rightarrow \text{Pathology},$$

where $a_1 = 2,6$.

The weighting vector $\mathbf{W_1}$ takes the form:

$\mathbf{W_1}$= (1,3361; 3,7152; -0,0521; 0,1466; -0,0275; -0,2784; -0,0075; -0,0031).

Experimental error classification passes pathology and false alarms are equal α = 0 and β = 0.

Recognition VE and AF classes:

$$\mathbf{W}_2^{\mathbf{T}}\mathbf{X}_i \geq a_2 \rightarrow \text{Ventricular extrasystoles}$$

$$\mathbf{W}_2^{\mathbf{T}}\mathbf{X}_i < a_2 \rightarrow \text{Atrial fibrillation},$$

where $a_2 = 3,5$.

The weighting vector $\mathbf{W_2}$ takes the form:

$\mathbf{W}_2$= (3,9862; -2,6938; -0,2813; -0,1135; 0,0229; 1,0453; 0,0385; -0,0148).

Experimental error classification passes pathology and false alarms are equal α = 0 and β = 0.

The obtained data shows that the use of morphological features of pseudo-phase portrait allows for effective recognition procedure three types of rhythm: normal rhythm, ventricular extrasystoles and atrial fibrillation.

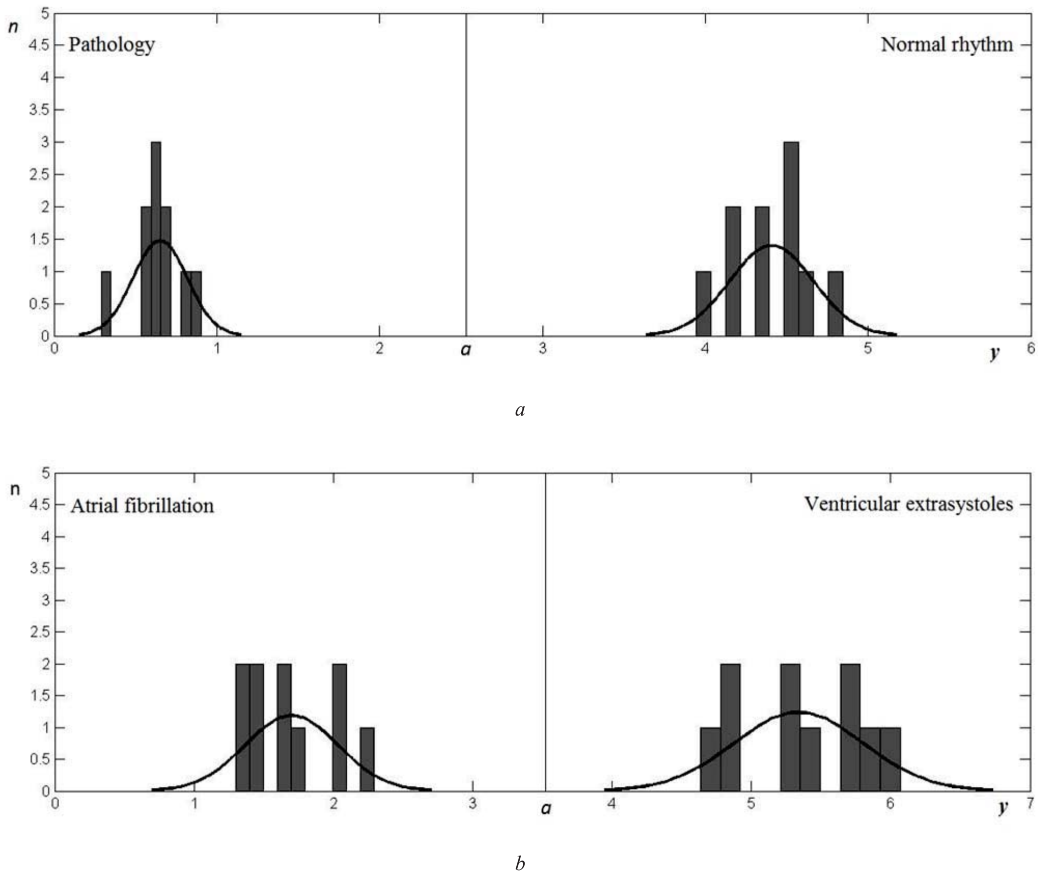


Fig. 6. Distribution by Fisher's ratio test: *a* - class of normal rhythm and pathology; *b* - the class of ventricular extrasystoles and atrial fibrillation

V. CONCLUSION

The article used the Grassberger-Procaccia algorithm to assess the correlation dimension of rhythmograms of normal rhythm, ventricular extrasystoles and atrial fibrillation. The results showed that with the increase of the chaotic component of the signal increases the complexity of its pseudo-phase description. Evaluation of morphological characteristics of pseudo-phase portrait gave a similar result. For classification of normal rhythm, ventricular extrasystoles and atrial fibrillation on rhythmogram, linear discriminant analysis was used based on Fisher's ratio test. Experimental results on real data have shown that the use of morphological characters of pseudo-phase portrait provides an effective procedure of recognition of the three rhythm types.

REFERENCES

[1]. D.A. Nikolaeva, *Primenenie metoda ocenki korreljacionnoj razmernosti dlja analiza JeJeG cheloveka s zabolevaniem jepilepsija, Journal Differential Equations and Control Processes.*, vol.2, Feb.2009

[2]. Y.A. Danilov, *Lekcii po nelinejnoj dinamiki*. Moscow: Book House "LIBROKOM", pp. 2011.

[3]. V.S. Anishhenko, T.E. Vadivasova, V.V. Astahov *Nelinejnaja dinamika haoticheskih i stohasticheskih sistem. Fundamental'nye osnovy i izbrannye problem*.

[4]. G.G. Malinetskii, A.B. Potapov and A.V. Podlazov, *Nelinejnaja dinamika*. Moscow: Book House "LIBROKOM", 2011.

[5]. V. S. Anishhenko *Znakomstvo s nelinejnoj dinamikoj*. Moskva: Izdatel'stvo URSS, 2008

[6]. A.P. Gavrilushkin and A.P. Maslyuk *"Teoreticheskie i prakticheskie aspekty nelinejnyh haoticheskih kolebanij ritma serdca. Medlennye kolebatel'nye processy v organizme cheloveka. Teoreticheskie i prikladnye aspekty nelinejnoj dinamiki, haosa i fraktalov v fiziologii i medicine"*. Proceedings of the 3rd All-Russian Symposium May 21-25, 2001. Novokuznetsk, 2001.

[7]. H. Duda, P. Hart *Pattern Classification and Scene Analysis*. New York: Wiley, 1976.