Reduction of the Spectral Feature Space Dimension in the Multiclass Problem of ECG-signals Recognition

L. Manilo, A. Nemirko

Saint Petersburg Electrotechnical University "LETI", Russia

lmanilo@yandex.ru, nemirko@yandex.ru

Abstract: *The problem of multiclass recognition of ECG signals using ECG description in the frequency domain is considered. The basis for developing the recognition algorithm is the analysis of Fisher's linear discriminant. Experiments on real signals showed a high efficiency of dangerous arrhythmias recognition.*

Keywords: ECG-signal analysis in frequency domain, Fisher linear discriminant, arrhythmia recognition.

1. INTRODUCTION

The most important tasks for the cardiological monitoring of ECG signals are detection of the dangerous arrythmias (ventricular fibrillation and ventricular fluttering) at the moment of its sudden occurrence, as well as for detection of anomalies, which prognose a grave condition of a patient. These arrhythmias are: polytopic ventricular extrasystolae, paroxysmal tachycardia, bidirectional ventricular tachycardia (polytopic ventricular tachycardia – torsades de pointes). It causes the problem to create decision functions for many classes of electrocardiographic signals. The most frequently used is signal description in the frequency domain, which is highly informative for detection of the dangerous arrhythmias [1, 2].

Recognition of the electrocardiographic signals in the frequency domain as a rule is based on the spectral characteristics obtained by calculation of the power spectral density (PSD) function. This description adequately reflects the frequency-response properties of the presented groups of signals, but involves the necessity to determine discriminant functions in a larger dimensional space, aggravates creation of the classification algorithms. Possibility to reduce feature space dimension can be obtained by grouping of the spectral coefficients, as well as by mapping the obtained description to a smaller dimensional space using multiple discriminant analysis. The decision functions are created on the basis of analysis of the Fisher's linear discriminant J, maximizing of which provides the best choice for the separation of c-classes of the set of signal vectors ($c - 1$) [3]. However, its optimization doesn't provide reliable detection of signals in every case.

J criterion for assessment the separation degree of the signals' initial classes can be presented with a scalar value defined by the trace of matrix as follows

$$J = tr(\mathbf{S}_2^{-1}\mathbf{S}_1), \qquad (1)$$

where $\mathbf{S}_1$ is scattering matrix between the classes, $\mathbf{S}_2$ is generalized scattering matrix inside the classes.

For c-classes, in case of transition from the L-dimensional space formed by the spectral features $\mathbf{G} = \left(\mathbf{G}_1, \mathbf{G}_2, \ldots, \mathbf{G}_L\right)^T$ into the (c-1)-dimensional space projections of the objects can be obtained through the following matrix transformation: $\mathbf{Y} = \mathbf{W}^T \cdot \mathbf{G}$, where $\mathbf{W}$ is a matrix of dimension $L \times (c-1)$, obtaining of which involves maximization of J. Disadvantage of the formula (1) is defined by the fact that increasing of the classes' quantity turns criterion J into the indicator of big distances between the groups, which results in poor reflection of mutual locations of closely located classes in the frequency domain with the J criterion.

2. OPTIMIZATION OF THE DECISION RULES CREATION

The procedure of decision rules creation can be optimized by reducing it to a set of pairwise classification tasks and introduction of weighting coefficients to increase the impact of closely spaced classes on the J criterion [4]. In this case the generalized formula for J criterion takes the form:

$$J = \sum_{i=1}^{c-1} \sum_{j=i+1}^{c} n_i n_j a_{i,j} \cdot tr\left[\left(\mathbf{W}^T \mathbf{S}_2 \mathbf{W}\right)^{-1}\left(\mathbf{W}^T \mathbf{S}_1^{(i,j)} \mathbf{W}\right)\right], \quad (2)$$

where n_i and n_j are rates of occurrence of the objects that form classes ω_i и ω_j.

Weight function $a_{i,j}$ can be linked to the recognition error values of each pair of classes ω_i and ω_j. In the publication [4] it is proposed to use weights in the form of a certain error function representation $erf\left(\dfrac{\eta - t}{\sigma}\right)$, where t is the decision rule limit, while η and σ are parameters of the distribution, calculated for specified groups of the objects based on the assumption of the normal distribution law with equal covariance matrices. This approach seems to be effective due to the possibility to approximate the criterion J to the evaluation of object detection reliability by summing the probabilities of correct solution within pairwise classification. This publication develops the idea of the criterion J approximation to the object classification accuracy in the space of the spectral features which are presented by the standardized values of the PSD.

Determining the elements of matrix $\mathbf{W}$ is reduced to the eigenproblem for the following matrix:

$$\mathbf{S}_2^{-1} \cdot \sum_{i=1}^{c-1} \sum_{j=i+1}^{c} n_i n_j a_{i,j} \cdot \mathbf{S}_1^{(i,j)}$$

and taking columns of the matrix $\mathbf{W}$ with dimension $L \times d$, ($d = c - 1$) equal to d eigenvectors that correspond to d maximum eigenvalues λ_i of them.

It is proposed to solve the task of determining $a_{i,j}$ in (2) as described below.

3. WEIGHT FUNCTION CALCULATION METHOD

Let us consider two classes of objects ω_i and ω_j with the normal law of distribution and unity matrices of covariation within the two-dimensional space (x_1, x_2). If the distance between the centers of these classes is designated as $\Delta_{i,j} = \|\mathbf{m}_i - \mathbf{m}_j\|$ where $\mathbf{m}_i$ and $\mathbf{m}_j$ are vectors of average values, then projection of these classes on the new direction $\mathbf{V}$ changes the distance between them depending on the angle α between the direction which joins centers of the classes, and the $\mathbf{V}$ vector. This dependence can be represented as $\Delta_{i,j}^{(v)} = \Delta_{i,j} \cdot \cos\alpha$. With identical a priori probabilities of the both class objects occurrences the probability of the correct recognition is determined as follows:

$$\gamma_{i,j} = \frac{1}{2} + \gamma_{i,j}' = \frac{1}{2} + \frac{1}{2} \cdot erf\left[\frac{\Delta_{i,j}^{(v)}}{2\sqrt{2}}\right],$$

where $erf(.)$ is the error function.

Then, for the case of c-classes with identical distributions the criterion $J^{(\gamma)}$, which evaluates average accuracy of the recognition, can be represented as follows:

$$J^{(\gamma)} = \sum_{i=1}^{c-1} \sum_{j=i+1}^{c} n_i n_j \gamma_{i,j} \qquad (3)$$

and the criterion J, which evaluates the range of divergence between the classes can be represented as follows:

$$J = \sum_{i=1}^{c-1} \sum_{j=i+1}^{c} n_i n_j a_{i,j} \cdot tr\left[\left(\mathbf{V}^T \mathbf{S}_1^{(i,j)} \mathbf{V}\right)\right] \qquad (4)$$

Comparison of the (3) and (4) gives weight values $a_{i,j}$ as follows:

$$a_{i,j} = \frac{\gamma_{i,j}'}{tr\left(\mathbf{V}^T \mathbf{S}_1^{(i,j)} \mathbf{V}\right)}$$

for the case of the best mutual disposition of the two classes (ω_i, ω_j) which corresponds to the case of the vectors $\mathbf{V}$ and $\mathbf{m}_{i,j} = (\mathbf{m}_i - \mathbf{m}_j)$ direction coincidence. Herewith, $\alpha = 0$; $tr\left(\mathbf{V}^T \mathbf{S}_1^{(i,j)} \mathbf{V}\right) = (\Delta_{i,j})^2$, while parameter $a_{i,j}$ in the area of small values, $\Delta_{i,j}$, can be defined through error function approximation by polynomial function as follows:

$$a_{i,j} \approx \frac{1}{8\sqrt{\pi} x_{i,j}}\left(1 - \frac{x_{i,j}^2}{3} + \frac{x_{i,j}^4}{2!5}\right), \qquad (5)$$

where $x_{i,j} = \left(\frac{\Delta_{i,j}}{2\sqrt{2}}\right)$, $\Delta_{i,j} \leq \sqrt{2}$, $x_{i,j} \leq 0,5$.

This method of determining the weighting functions $a_{i,j} = a(\Delta_{i,j})$ can be used to solve the multi-class task assuming that each c-class has matrix of intra-group scattering defined by the generalized scatter matrix $\mathbf{S}_2 = \sum_{i=1}^{c} n_i \cdot \Sigma_i$, where Σ_i is sample covariance matrix of the i-th class. Then it is necessary to find the Euclidian distance between the centers of the corresponding classes $\Delta_{i,j}$ for each class pair $(\omega_i, \omega_j; i,j = 1,\ldots,c; i \neq j)$ in the initial L-dimensional space of the spectral signs and determine the weights $a_{i,j}$ using expression (5). Maximization of the criterion J (2) leads to the procedure of the eigenvectors $\mathbf{W}_i, i = 1,\ldots,c-1$ finding and to the analysis of the object groups distribution within the feature space of reduced dimension.

4. APPLICATION FOR DANGEROUS ARRHYTHMIAS RECOGNITION

Experimental research was aimed at the creation of reliable methods to detect ventricular fibrillation, but during this research was also analyzed the possibility of recognition of such arrhythmia in the early stages of its manifestation, namely, at the time when paroxysmal tachycardia turns into ventricular flutter, which is accompanied, as a rule, with a short period of the polytopic (torsades de pointes) tachycardia. In particular, research was aimed to solve the problem of recognition of 3 classes of dangerous arrhythmias: ω_1 - ventricular fibrillation and ventricular flutter, ω_2 - paroxysmal tachycardia and basic rhythm, presented by different forms of extrasystole, ω_3 - polytopic tachycardia (torsades de pointes).

Application effectiveness of the reviewed procedures was assessed on the basis of results of the experiments performed with real data, including more than 20 minutes long electrocardiographic signal records. All ECG realizations are obtained from the standard MIT-BIH data base of ECG signals.

An ordered set of 28 spectral features obtained in the frequency domain limited by 15 Hz using overlapped segments was taken as the initial description of the objects represented by 2-second fragments of electrocardiographic signal [5]. Uncorrelated PSD evaluations are calculated within a spectral window of $\Delta f = 0.976$ Hz, but the spacing along the frequency axis was half of this value, i.e. 0.488 Hz. This case provides the possibility to retain features of the analyzed signal spectrum shape with relatively stable estimates of the PSD. Distribution of objects of three classes ($c=3$) represented in a space with dimension $L=28$ is evaluated in a space with dimension $d=2$, obtained by traditional approach (1) using weight functions (2). Comparative analysis of the two distributions shows that application of the criterion (2) results in the smoother grouping of the objects on the plane.

During experiments dividing functions were created, limits of the solution areas were defined and classification errors were determined, which are the reliability criterion for arrhythmias detection. Result of the linear

discriminant analysis showed that application of the weight functions can decrease the average classification error from 8.2% down to 4.6%, which is an indicator of the effectiveness for this optimization procedure. Analysis of the objects which were situated in the intersection zones of the received solutions has shown that they are controversial in terms of classification, since they can be included in one of the alternative classes. This mainly concerns the intersection zones ω_1 and ω_2 with the intermediate class ω_3. Important result is the unmistakable recognition of classes ω_1 and ω_2, which ensures reliable detection of ventricular fibrillation at the stage of its persistent manifestation.

5. REFERENCES

[1] Clayton R.H., Murray A., Campbell R.W.F. Frequency Analysis of Self–Terminating Ventricular Fibrillation // IEEE Computer Society Press. – 1994. – P. 705 – 708.

[2] Clayton R.H., Campbell R.W.F., Murray A. Time–Frequency Analysis of Human Polymorphic Ventricular Tachycardia // Computers in Cardiology. – 1997, vol 24. –P. 97 – 100.

[3] Duda Z., Hart P. Pattern classification and Scene Analysis. Translated from English. – M.: Mir, 1976. 511 p.

[4] Loog M., Duin R.P.W., Haeb-Umbach R. Multiclass Linear Dimension Reduction by Weighted Pairwise Fisher Criteria. IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol.23, No. 7, 2001, pp. 762-766.

[5] Manilo L.A. Spectral features harmonization by empirical estimates of distances between groups in biosignals classification problems. – Izvestia vuzov Rossii. Radioelectronika. Vyp.3, 2006 – p.20 – 29.