11th International Conference on Intelligent Data Processing: Theory and Applications, 10-14 October 2016, Barcelona, Spain

Multiple discriminant analysis for recognition of biosignals in frequency domain

L. Manilo, A. Nemirko

Department of Biotechnical Systems
Saint-Petersburg Electrotechnical University "LETI",
Saint Petersburg, Russian Federation

Plan of the presentation

- Introduction
- Optimization of decision rules constructing
- Method of weight functions calculation
- Application of the proposed method for dangerous arrhythmias recognition
- Conclusion

Introduction

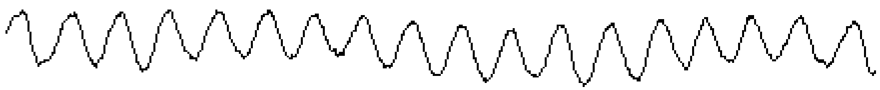
The most important tasks of ECG monitoring in cardiology are:

- detection of dangerous arrhythmias (ventricular fibrillation – VF and ventricular flutter of the heart - VF)
- recognition of rhythm disorders, that are precursors to severe patient states: (polytopic ventricular extrasystoles, paroxysmal tachycardia - PT, bidirectional ventricular tachycardia (pirueta form) – TP (torsade de pointes).
- normal rhythm and ventricular extrasystoles we will denote as the background rhythm BR

For the solution of these tasks the description of signals in the frequency domain is most often used, because of its relatively high informativity for detection of dangerous arrhythmias.

In this case, there arises the problem of decision functions developing for many classes of ECG.

Examples of ECG forms in a time interval of 5 s

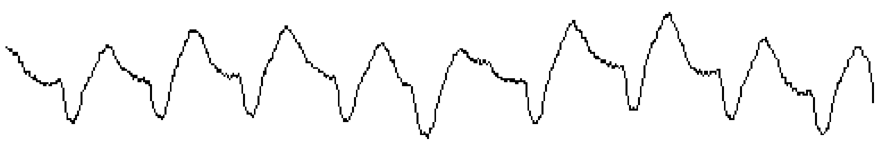


VF (ventricular flutter)

a)

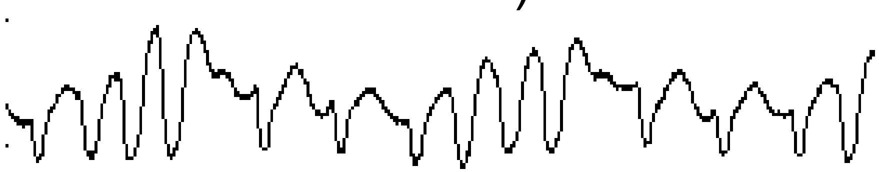
VF (ventricular fibrillation)

b)



PT (paroxysmal tachycardia)

c)



TS (torsade de pointes)

d)

ECG Recognition in the frequency domain

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 involves the necessity to determine discriminant functions in a **larger dimensional space**, that makes difficulties for the classification algorithms creation.

Possibility **to reduce feature space dimension** can be obtained by mapping the obtained description to a smaller dimensional space using **multiple discriminant analysis**.

The decision functions are created on the basis of **Fisher's linear discriminant *J***, maximizing of which provides the best choice for the separation of c-classes of the set of $(c - 1)$ signal vectors. However, its optimization doesn't provide reliable detection of signals in every case.

Fisher's criterion *J*

J criterion for assessment the separation degree of the initial signals 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 of spectral features $\mathbf{G}^{(L)} = \left(\mathbf{G}_1, \mathbf{G}_2, \ldots, \mathbf{G}_L\right)$ into the (*c*-1)-dimensional space can be obtained in the form

$$\mathbf{Y} = \mathbf{W}^T \cdot \mathbf{G}^{(\mathrm{L})},$$

where **W** is a matrix of dimension $L \times (c-1)$, obtaining of which involves maximization of *J*.

Disadvantage of the formula (1) is that with increasing numbers of classes criterion *J* becomes an indicator of large intergroup distances, and poorly reflects the mutual locations of closely spaced classes in frequency domain.

And this in turn limits the application of the method for the classification of difficult separable groups of objects.

Optimization of decision rules constructing (1)

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**. In this case the generalized formula for J criterion takes the form:

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

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

Determining the elements of matrix $\mathbf{W}$ comes down 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.

Optimization of decision rules constructing (2)

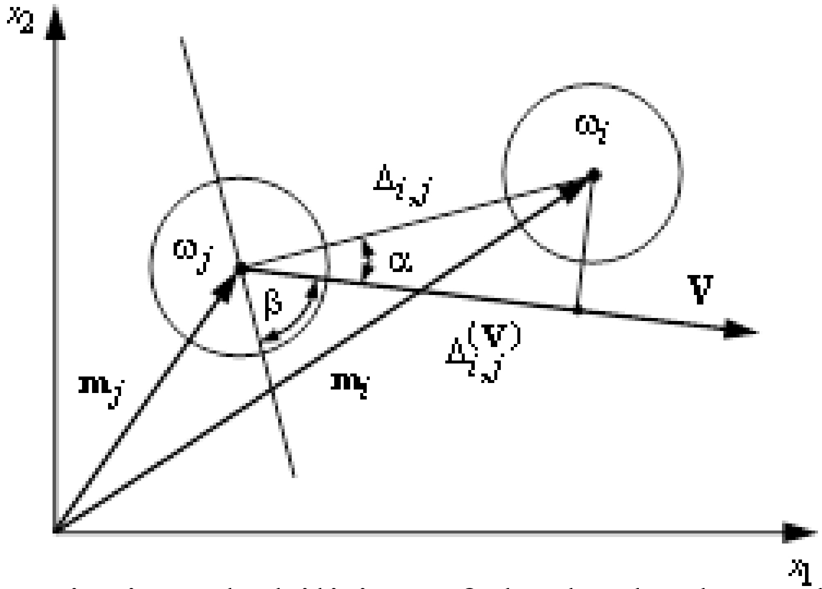
Weight function $a_{i,j}$ can be linked to the recognition error values of each pair of classes ω_i and ω_j. In the publication [Loog M., 2001] it is proposed to use weights in the form of a certain error function representation $erf\left(\frac{\eta - t}{\sigma}\right)$, where t is the decision rule 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 paper 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.

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.

Method of weight functions calculation (1)

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).



With identical a priori probabilities of the both class objects occurrences the probability of the correct recognition towards direction V is equal to

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

where *erf(.)* is the error function, $\Delta_{i,j}^{(v)} = \Delta_{i,j} \cdot \cos\alpha$, $\Delta_{i,j} = \left\|\mathbf{m}_i - \mathbf{m}_j\right\|$

Method of weight functions calculation (2)

For the case of *c*-classes with identical distributions the crite
evaluates average accuracy of the recognition, can be represented as fol

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

and the criterion *J*, which evaluates the range of divergence between t
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:

$$a_{i,j} = \gamma'_{i,j} \Big/ tr\left(\mathbf{V}^T \mathbf{S}_1^{(i,j)} \mathbf{V}\right)$$

for the case of the vectors $\mathbf{V}$ and $\mathbf{m}_{i,j} = \left(\mathbf{m}_i - \mathbf{m}_j\right)$ direction coincider

Method of weight functions calculation (3)

In this case $\alpha = 0$; $tr\left(\mathbf{V}^T\,\mathbf{S}_1^{(i,j)}\,\mathbf{V}\right) = \left(\Delta_{i,j}\right)^2$, and parameter $a_{i,j}$ is equal to

$$a_{i,j} = \gamma'_{i,j} \Big/ \left(\Delta_{i,j}\right)^2 = \frac{1}{2\left(\Delta_{i,j}\right)^2} \cdot erf\left[\frac{\Delta_{i,j}}{2\sqrt{2}}\right]$$

For a normalized spectrum (PSD), using the approximation of the error function by polynomial function provided that $\max(\Delta_{i,j}) = \sqrt{2}$, the following expression was obtained.

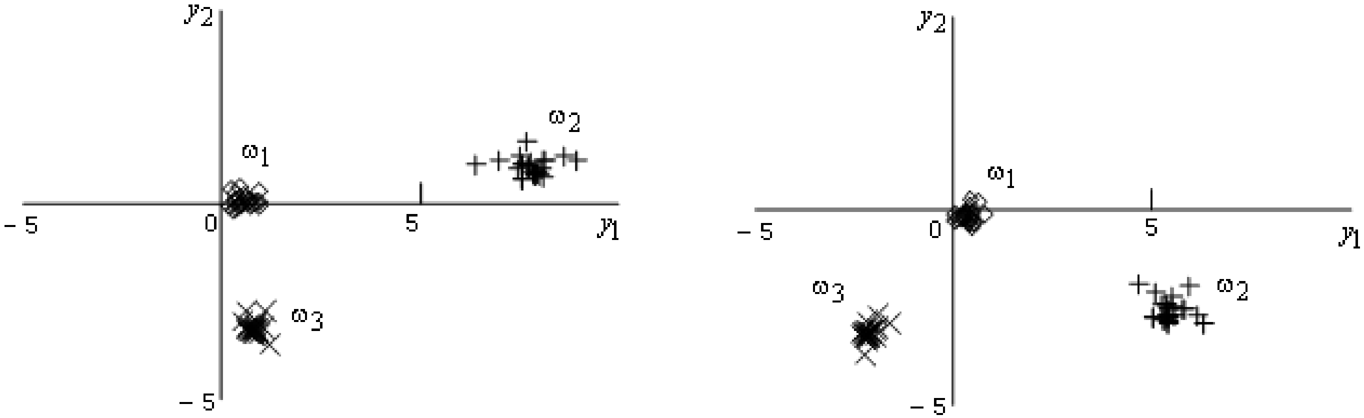
$$a_{i,j} \approx \frac{1}{8\sqrt{\pi}\,x_{i,j}}\left(1 - \frac{x_{i,j}^2}{3} + \frac{x_{i,j}^4}{2!5}\right) \qquad (5)$$

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

Then it is necessary to find the Euclidian distance between the centers of the corresponding classes $\Delta_{i,j}$ for each class pair $\left(\omega_i, \omega_j; i, j = 1, \ldots, c; i \ne j\right)$ 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.

Experiments on real data



Mapping of 3 classes of ECG: ω_1 (VF), ω_2 (PT), ω_3 (BR) – in the transformed space of spectral features (y_1, y_2), when optimizing by criterion $\boldsymbol{J}$ – (*a*), and using the weight functions – (*b*).

This example illustrates the effect of weight functions on the choice of vectors $\mathbf{W}_1$, $\mathbf{W}_2$ on the transformation of the original 38-dimensional space of spectral signs.

The use of the weights in accordance with (2) leads to a more uniform objects grouping on the plane.

Application of the proposed method for dangerous arrhythmias recognition (1)

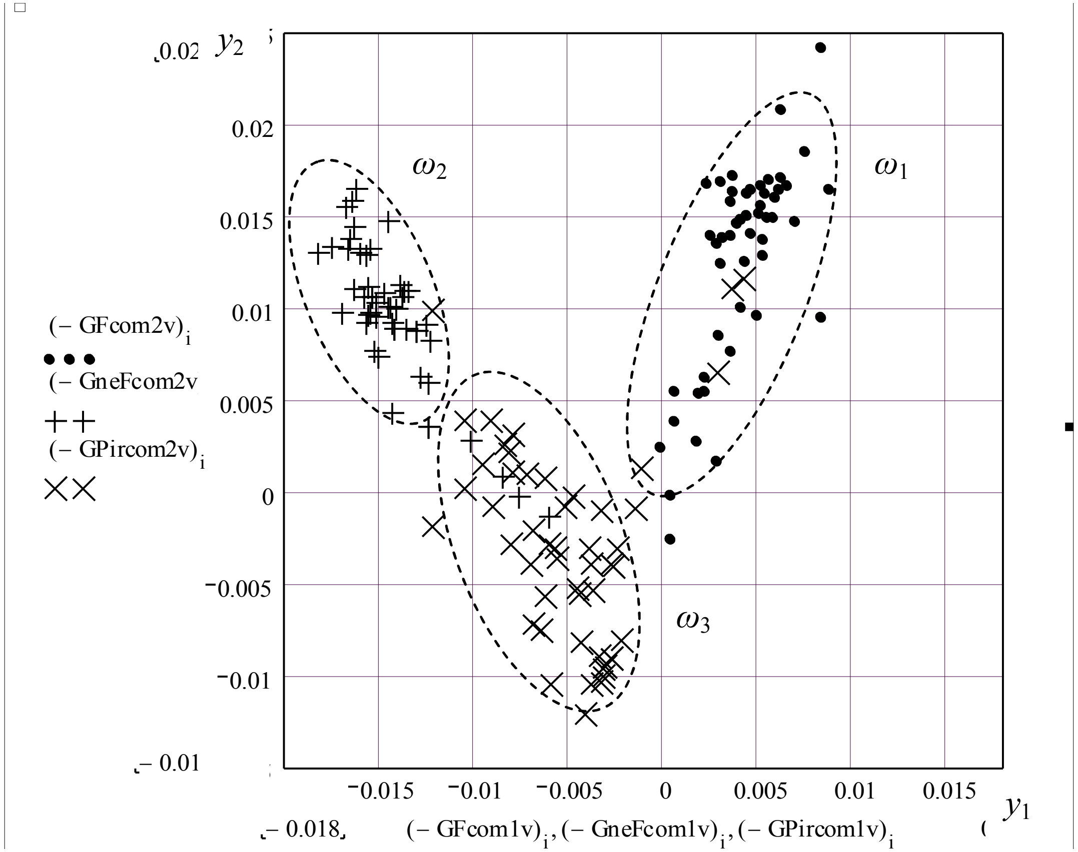
Problem: recognition of 3 classes of dangerous arrhythmias ω_1 (VF), ω_2 (PT), ω_3 (TS).

Experimental data: ECG recordings with total duration of more than 20 minutes, obtained from a standard computer ECG database (MIT-BIH).

Input data: ECG fragments with a duration of 2 s, represented by a set of 28 spectral coefficients obtained in the frequency range 0 – 15 Hz with the use of overlapping segments.

Optimization of two-dimensional space drawn according to the criterion ***J*** with the use of weight functions.

Application of the proposed method for dangerous arrhythmias recognition (2)



Representation of 3 classes of dangerous arrhythmias ω_1 (VF), ω_2 (PT), ω_3 (TS) in the space of the vectors $\mathbf{W}_1$, $\mathbf{W}_2$, obtained with the use of weight functions.

Analysis of the experimental results

Separating functions are developed. Boundaries of the decisions regions are defined and classification errors are found.

It is established:

- the use of weight functions reduces the average error of classification from 8.2% to 4.6%,
- objects caught in the intersection zone of the obtained decision regions are debatable in terms of classification. This applies mostly to crossing areas ω_1 and ω_2 with an intermediate class ω_3 formed the initial form of dangerous arrhythmias,
- correctly recognized classes ω_1 and ω_2, which ensures reliable detection of ventricular fibrillation in the stage of stable symptoms.

Conclusion

1. It is shown that multiple discriminant analysis using the weight functions improves the conditions for recognizing closely spaced classes of signals.
2. We have obtained an analytical expression for weight coefficients calculation when analyzing of signals in frequency domain.
3. The example of dangerous arrhythmias recognition showed the possibility of applying multiple discriminant analysis to practical medical problems.

Thank you