Annals of the Academy of Romanian Scientists
Series on Science and Technology of Information
ISSN 2066 - 8562 Volume **3**, Number **2**/**2010**

PROCESSING ECG DATA USING MULTIVARIATE DATA ANALYSIS

Mircea Valer PUŞCĂ[1]
Horia F. POP[2], Nicolae Marius ROMAN[3], Vasile IANCU[4]

Abstract. *In this paper it is present an efficient technique for ECG data processing, based on fuzzy and non-fuzzy multivariate analysis methods. The present study shows the theoretical advantages of fuzzy algorithm in clinical utility for computer criteria in ECG studies at patients with heart disease. It can be starting point for new ECG devices with computer multivariate data analysis.*

Keywords: ECG data processing, fuzzy, clustering algorithm, multivariate analysis methods

1. Introduction

Fuzzy multivariate data analysis has been extensively used in a lot of research in chemistry and chemo metrics [3,18,19], medical sciences [13,21], health and environment [16,20], software engineering [5,6,15,24,22].

The paper is organized as follows. Section 2 covers fuzzy and non-fuzzy multivariate analysis methods useful for ECG data processing. Section 3 presents an overview of previous experiments with multivariate analysis of ECG data. The paper ends with concluding remarks.

2. Fuzzy and crisp multivariate analysis methods modelling of the proposed adjustment scheme

2.1 Fuzzy sets and fuzzy clustering

The theory of fuzzy sets was introduced in 1965 by Lotfi A. Zadeh [26] as a natural generalization of the classical set concept. Let *X* be a data set, composed of *n* data items characterized by the values of s characteristics. A fuzzy set on *X* is a mapping *A*: *X* --> [0, 1]. The value *A(x)* represents the membership degree of the data item *x* from *X* to the class *A*.

[1]Senior Researcher, Electrical Engineering Faculty, Technical University of Cluj-Napoca, Cluj-Napoca, corresponding member of the Academy of Romanian Scientists; mvpusca@yahoo.com.
[2]Prof. univ. PhD, Department of Computer Science, Babes-Bolyai University Cluj-Napoca, Cluj-Napoca, Romania, hfpop@cs.ubbcluj.ro.
[3]Prof. univ. PhD, Electrical Engineering Faculty, Technical University of Cluj-Napoca, Cluj-Napoca, Romania, marius.roman@et.utcluj.ro.
[4]Prof. univ. PhD, Electrical Engineering Faculty, Technical University of Cluj-Napoca, Cluj-Napoca, Romania, vasile.iancu@mae.utcluj.ro.

The advantage of this approach is that it allows a data item x to be a member of more classes, with different membership degrees, according to certain similarity criteria.

Clustering algorithms based on fuzzy sets have delicate situations.

The class of fuzzy clustering algorithms based on fuzzy objective functions, by James C. Bezdek [1] provides a large share of geometrical prototypes and combinations thereof, to be used according to the data substructure.

On the other hand fuzzy hierarchical algorithms provide an in-depth analysis of the data set, by deciding on the optimal subcluster cardinality and the optimal cluster substructure of the data set.

2.2 Restricted fuzzy clustering

Let us consider a set of classified objects, $X = \{x^1, \ldots, x^p\} \in \mathsf{R}^s$ and the fuzzy partition $P = \{A_1, \ldots, A_n\}$ corresponding to the cluster substructure of the set X.

Let $x^0 \in \mathsf{R}^d$ be an object that needs to be classified with respect to the fuzzy partition P.

The algorithm we are presenting here computes the optimal fuzzy partition $\widetilde{P}$ corresponding to the set $\widetilde{X} = X \cup \{x^0\}$, by using a mechanism similar to Fuzzy n-means, with the difference that the membership degrees of the objects in X to the classes A_i, $i = 1, \ldots, n$ may not be modified [Fre09].

In what follows we consider a metric d in the Euclidean space R^s.

We will suppose that d is norm induced, so $d(x, y) = (x - y)^T M (x - y), x, y \in \mathsf{R}^s$, where M is a symmetrical and positively defined matrix.

The objective function we have in mind for our problem is similar to that for the Fuzzy n-Means Algorithm:

$$\widetilde{J}(\widetilde{P}, L) = \sum_{i=1}^{n} \sum_{j=0}^{p} \left(A_i(x^j)\right)^2 d^2\left(x^j, L^i\right),$$

with the mention that $A_i(x^j)$ are kept constant for each i and for $j = 1, \ldots, p$.

The main result with respect to determining the fuzzy partition $\widetilde{P}$ and its representation L minimizing the function $\widetilde{J}$ is the following

Theorem. (i) The fuzzy partition $\widetilde{P} = \{A_1, \ldots, A_n\}$ has the minimum value of the function $J(\cdot, L)$ if and only if

$$A_i(x^0) = \frac{1}{\sum_{k=1}^{n} \frac{d^2(x^0, L^i)}{d^2(x^0, L^k)}}. \tag{1}$$

(ii) The set of prototypes $L = \{L, \ldots, L^n\}$ has the minimum value of the function $J(\widetilde{P}, \cdot)$ if and only if

$$L^i = \frac{\sum_{j=0}^{p} \left(A_i(x^j)\right)^2 x^j}{\sum_{j=0}^{p} \left(A_i(x^j)\right)^2}. \tag{2}$$

With this result, the optimal membership degrees for x^0 to the classes A_i will be determined using an iterative method in which $\widetilde{J}$ is successively minimized with respect to $\widetilde{P}$ and L. The process will start with the initialization of prototypes L^i to the values that correspond to the fuzzy membership degrees of the original fuzzy partition P.

The resulted algorithm, **Restrictive Fuzzy n-Means Clustering Algorithm**, follows:

1. Let us have X and P as given variables.

2. Determine the initial positions of the prototypes L^i according to value of P.

3. Determine the membership degrees $A_i(x^0)$, $i = 1, \ldots, n$, using relation (1).

4. Determine the new positions of prototypes L^i, $i = 1, \ldots, n$, using relation (2).

5. If the new positions of the prototypes L^i are close enough to the former positions, then **stop**, else return to step 3.

2.3 Hierarchic Fuzzy Clustering with incomplete data

The problem of incomplete data is extremely important [1]. If a cluster substructure of a data set with incomplete data is required, the option of ignoring the incomplete data items altogether is not realistic, because that would assume ignoring useful, available data. Hathaway and Bezdek have proposed in 2001 a few strategies to cope with the incomplete data problem. Their approach has been to extend the Fuzzy c-means algorithm in such a way as to accept incomplete data as well. We have used the approach of Hathaway and Bezdek and produced hierarchic clustering versions thereof.

The main issue is that, in computing the fuzzy membership degrees, the square distances $d^2(x^j, L^i)$ from the data item x^j to the class prototype L^i cannot be computed for incomplete data. Assuming that we use the Euclidean metric, we would need to compute

$$D_{ij} = \sum_{k=1}^{s} \left(x_k^j - v_k^i\right)^2. \tag{3}$$

However, for some values of j and k, x_k^j may be missing. In our paper we are going to use the *Partial Distance Strategy* of Hathaway and Bezdek [1]. Namely, we are going to compute this sum considering only the available values and we are going to scale the result to take into account the missing dimensions. As such, we are actually going to compute

$$\widetilde{D}_{ij} = \frac{s}{\sum_{k=1}^{s} I_{jk}} \sum_{k=1}^{s} I_{jk} \left(x_k^j - v_k^i\right)^2, \tag{4}$$

where $I_{jk} = 1$ if the value x_k^j is available, and $I_{jk} = 0$ if the value is unavailable, in which case the difference will not be computed.

Similarly, the new prototypes will have to be computed by taking into account only the available data:

$$L_k^i = \frac{\sum_{j=1}^{p} I_{jk} \left(A_i(x^j)\right)^2 x_k^j}{\sum_{j=0}^{p} I_{jk} \left(A_i(x^j)\right)^2}, \tag{5}$$

where I_{jk} have the same meaning as above.

The resulted algorithm, **Hierarchic Fuzzy Clustering with Incomplete Data**, follows [6]:

1. Let us consider the data set X and the initial fuzzy partition P.

2. Determine the positions of prototypes L^i, $i = 1, \ldots, n$, using relation (5).

3. Determine the membership degrees $A_i(x^0)$, $i = 1, \ldots, n$, using dissimilarities from (4).

4. If the new fuzzy partition is close enough to the former fuzzy partition, then **stop**, else return to step 2.

2.4 Non-binary Divisive Hierarchic Fuzzy Clustering

A quite important problem of divisive clustering is the choice for binary split at every level of the clustering hierarchy.

This is a very simple, intuitive and effective approach. While it leads to desired results in most of the cases, there are situations where the data does not show a natural binary split-up. Here a different approach must be used.

A possible solution is to generalize the binary divisive approach introduced in [4].

The central point of the Fuzzy Hierarchic Divisive method is the binary polarization index.

Thus, considering a binary fuzzy partition $P = \{C_1, C_2\}$, of the fuzzy set C, the partition separation index is defined as

$$R(P) = \frac{\sum_{i=1}^{2}\sum_{k=1}^{n} C_{i,1/2}(x_k)}{\sum_{k=1}^{n} C(x_k)} \tag{6}$$

where

$$C_{i,t}(x) = \begin{cases} C_i(x) & C_i(x) > t \\ 0 & otherwise \end{cases}$$

Now, instead of considering a binary partition, let us consider one with p arity, $P = \{C_1, \ldots, C_p\}$. We define the generalized polarization index quite naturally as

$$R(P) = \frac{\sum_{i=1}^{p}\sum_{k=1}^{n} C_{i,1/2}(x_k)}{\sum_{k=1}^{n} C(x_k)}. \tag{7}$$

The key problem is then to obtain the best partition arity at every node of the classification tree. Fortunately, this is a very simple issue from a constructive point of view. Our assumption has been that the data does not display a binary structure. So, at every node of the classification tree we are going to construct both a 2-partition and a 3-partition.

The decision of which is best to adopt is based on the higher partition polarization index. By working recursively, in this way, we are going to obtain a classification tree where each node has either two or three children, as appropriate.

2.5 Fuzzy regression

Fuzzy clustering techniques are suitable for determining the optimal cluster substructure of a data set, and they suppose that such a substructure does exist. The problem at hand is, however, to be able to determine the one fuzzy set *A* and its prototype *L* that best describes the data set. In such a case, a regular fuzzy clustering algorithm will not work.

The fuzzy set that best corresponds to a data set, based on a prototype characterization of the data, is a useful notion in the search for robust regression techniques, as well as for developing data analysis techniques where the data items are considered according to their goodness of fit (i.e. their membership degree to this fuzzy set).

We consider a binary fuzzy partition, {A, Ā}, where Ā is a virtual class with a hypothetical prototype, characterized by the constant dissimilarity

$$D(x^j, \bar{L}) = \delta := \left(\frac{\alpha}{1-\alpha}\right)^{m-1}. \tag{8}$$

The optimal fuzzy set A, as defined by our problem, is determined by minimizing the following fuzzy objective function:

$$J(A, L) = \sum_{j=1}^{n} A(x^j)^m D(x^j, L) + \sum_{j=1}^{n} \bar{A}(x^j)^m \left(\frac{\alpha}{1-\alpha}\right)^{m-1}, \alpha \in (0,1). \tag{9}$$

The algorithm used to solve this problem has been called the *Fuzzy Regression* generic algorithm [12,18]:

1. Given α; Initialize $A^{(0)}(x) = 1$, l=0;
2. Compute prototype L that minimizes $J(A^{(l)},.)$;
3. Compute fuzzy set $A^{(l+1)}$ that minimizes J(.,L):

$$A^{(l+1)}(x^j) = \frac{\frac{\alpha}{1-\alpha}}{\frac{\alpha}{1-\alpha} + D(x^j, L)^{\frac{1}{m-1}}} \tag{10}$$

4. Compare fuzzy sets $A^{(l+1)}$ with $A^{(l)}$. If close enough, then stop, else increase l by 1 and goto step 2.

As an improvement, in order to assure the independence of scale, in the equation from step 3, we will replace the dissimilarity $D(x^j, L)$ with the relative dissimilarity

$$D_r(x^j, L) = D(x^j, L) / \max_{j=1,n} D(x^j, L). \tag{11}$$

This is equivalent to setting δ initially to

$$D(x^j, \bar{L}) = \delta := \left(\frac{\alpha}{1-\alpha}\right)^{m-1} \max_{j=1,n} D(x^j, L). \tag{12}$$

Let us now suppose that X is a data set, and A and L are the optimal fuzzy set and its prototype representation, respectively. The following properties are valid [Pop96]:

1. (Maximal membership degree) A(x)=1 => D(x,L)=0
2. (Minimal membership degree) A(x)=α => D_r(x,L)=1
3. (Membership degree interval) A(x) ∈ [α, 1] for all x in X
4. (Empty fuzzy set) α = 0 => A(x) = 0 for all x in X
5. (Degenerate fuzzy set) α = 1 => A(x) = 1 for all x in X
6. (Strict monotony) A(x) < A(y) => D(x,L) < D(y,L)
7. (Equality) A(x) = A(y) => D(x,L) = D(y,L)

Based on these properties we remark that the constant α is an input parameter that has the role of setting the polarization of the fuzzy partition {A, $\bar{A}$}. The best results appear to be obtained with α = 0.10.

Being based on the Fuzzy c-Lines algorithm, the linear version of this algorithm (Fuzzy Linear Regression) uses as dissimilarity the square distance to the line, as opposed to the y-distance, used by the classical Least Squares algorithm. Due to the use of fuzzy sets, the algorithm is efficient in all testing conditions, and is better than most methods it has been compared with [12].

The algorithm allows the detection of the type of data sets, i.e. homoscedasticity, heteroscedasticity, presence of outliers, or any combination thereof, with or without any other irregularities. This is done through repeated runs by analysing the graphical representation of the surface made by the coefficients vectors of the linear prototypes, as determined for α varied continuously in the interval (0, 1). In the case of a two-dimensional data set, the curve is defined through the points (a_0, a_1), where $y=a_0+a_1 x$ is the linear prototype of the data set defined.

2.6 Principal components analysis

Principal component analysis (PCA) is a favourite tool in environ metrics for data compression and information extraction. PCA finds linear combinations of the original measurement variables that describe the significant variation of data.

PCA is designed to transform the original variables into new, uncorrelated variables (axes) called the principal components, that are linear combinations of the original variables. The new axes lie along the directions of maximum variance. PCA provides an objective way of finding indices of this type so that the variation in the data can be accounted for as concisely as possible [10,25].

PCA is based on eigenanalysis of the covariance or correlation matrix. Let us consider a data set $X = \{x^1, \ldots, x^p\}$ and its covariance matrix M, given by:

$$M_{ij} = \frac{1}{p-1}\sum_{k=1}^{p}\left(x_i^k - \overline{x_i}\right)\left(x_j^k - \overline{x_j}\right), i, j = 1,\ldots,n. \tag{13}$$

Let us also consider the orthonormal eigenvectors e^i of the matrix M, and the corresponding eigenvalues λ_i ($i = 1, \ldots, n$).

The principal components of the data set X appear as linear combinations of the original variables in the form

$$PC_i = e_1^i y^1 + e_2^i y^2 + \ldots + e_n^i y^n, \tag{14}$$

where y^i represents the i-th original variable and e^i_j represent the j-th element of the eigenvector e^i of the matrix M.

A constraint that all eigenvectors e^i have unit norms is imposed. This constraint is introduced in order to ensure that the variance of a principal component cannot be increased by simply increasing any of the e^i_j values.

From the orthonormality of $e^1, e^2, \ldots, e^n$ it follows that

$$\begin{gathered}
\text{eiT ei} = 1 \text{ for any } i \in \{1,\ldots,n\} \\
\text{eiT ej} = 0 \text{ for any } i, j \in \{1,\ldots,n\}, i \neq j \\
\text{eiT M ei} = 1 \text{ for any } i \in \{1,\ldots,n\} \\
\text{eiT M ej} = 0 \text{ for any } i, j \in \{1,\ldots,n\}, i \neq j
\end{gathered} \tag{15}$$

and

$$M = \lambda_1 e^1 e^{1T} + \lambda_2 e^2 e^{2T} + \ldots + \lambda_n e^n e^{nT},$$

where T denotes the transposing operation.

The basic property of the new variables is their lack of correlation.

We have that

$$\text{Var}(e^i X) = \lambda_i \text{ for } i = 1, \ldots, n$$

and

$$\mathrm{Cov}(e^i X, e^j X) = 0 \text{ for } i, j = 1, \ldots, n;\ i \neq j.$$

The first principal component PC1 is that linear combination of sample values for which the "scores" have maximum variation.

The second component PC2 has scores that are uncorrelated with the scores for PC1. Among the many linear combinations with this property we select the one which has maximum variation among its scores.

The third component PC3 is defined to be that linear combination which has the maximum variation among all those combinations whose scores are uncorrelated with the scores of the first two components. Subsequent components are defined analogously.

Principal component analysis, as any other multivariate statistical methods, is sensitive to outliers, missing data, and poor linear correlation between variables, due to poorly distributed variables. As a result, data transformations have a large impact upon PCA [25].

2.7 Fuzzy Principal Components Analysis

One of the best approaches to robustify PCA appears to be the fuzzification of the matrix data by diminishing in this way the influence of the outliers.

The major problem of the PCA algorithm rests, as always, with the isolated points. As a first possible way to handle this, we will take into account the points isolated with respect to the first principal component only.

Fuzzy membership degrees are introduced according to the distance to the first principal component. As such, a scheme similar to the fuzzy regression algorithm is used, to determine the first fuzzy principal component and the corresponding fuzzy membership degrees. The method is called Fuzzy PCA (first component) [13,20].

The traditional covariance matrix is thus replaced by the fuzzy covariance matrix, given by

$$C_{ij} = \frac{\sum_{k=1}^{p} A(x_k)^m \left(x_i^k - \overline{x_i}\right)\left(x_j^k - \overline{x_j}\right)}{\sum_{k=1}^{p} A(x_k)^m}, i, j = 1, \ldots, n. \tag{16}$$

This procedure involves as well the computation of the membership degrees corresponding to the fuzzy set A, as they will help determining the contribution of each of data items to the fuzzy covariance matrix. Since the first principal

component is, actually, the major linear representative of the whole data set, this is exactly the fuzzy linear regression problem.

As such, the value of the fuzzy membership degree of the farthest outlier is an input parameter of the problem, denoted with α. We could, however, loop with values for α between 0 and 1, with a certain step, and get, the value of α that makes the fuzzy set A if not optimal then at least very good with respect to minimizing the fuzzy linear regression error criterion.

Once the „optimal" fuzzy set A has been constructed, the method works as in the original case, by determining the eigenvectors and eigenvalues of the fuzzy covariance matrix. The major advantage is that the first principal component will count the merits of each data item; as such, will consider the isolated points with less significance.

The Fuzzy PCA algorithm described before fuzzifies only the first component. In order to get a most effective method, we have to deal with the problem of fuzzifying all the components. The main idea is to use a different approach: by projecting the data in smaller-sized spaces. After the first fuzzy eigenvector is determined, all data is projected to the hyperplane rectangular on it. The eigenvectors corresponding to the projected data will be orthogonal to the eigenvector determined above. As such, the second largest eigenvector of the original data will correspond to the largest eigenvector of the projected data. This projection mechanism continues further on, etc. Finally, the eigenvectors are rebuilt in the original space.

More specifically, let us consider the given data set X. The PCA method actually leads to a rotation of the axes such that the data set will be represented using the scores Y, and the new axes are the fuzzy principal components. So, in the coordinates system given by the principal components, the *i*-th data item y^i will be represented as $(y^i_1, y^i_2, \ldots, y^i_n)$. We are projecting the set Y on the first fuzzy principal component. The projection of the i-th data item will be represented as $(y^i_2, \ldots, y^i_n)$ and will be an object in the n-1 dimensional space.

Let us call this new data set Y". All the fuzzy principal components of the

set Y" are orthogonal on the first fuzzy principal component of the original data set X. As such, the first fuzzy principal component of the set Y" is, actually, the second fuzzy principal component of the set X. But, the vector determined so far is represented in the n-1 dimensional space.

We will extend it, by prefixing it with 0 (for the first coordinate value), so that we could have a representation in the n-dimensional rotated space (with the scores). And we need to „unrotate" it, such that we now have a representation in the original space.

The third fuzzy principal component is determined in a similar manner: it is the first fuzzy principal component of a data set projected onto the first two principal components. The process continues until all fuzzy principal components will have been determined.

The computation of the other fuzzy components of the data is reduced to the computation of the first fuzzy component of smaller-sized data, i.e. reducing the problem to a problem that has already been solved.

This method, called Fuzzy PCA (all components) has the advantage that all the fuzzy principal components are determined in an autonomous manner, with fuzzy membership degrees independently determined for each individual component. In this way, we take into account that data items may be outliers with respect to a principal component, but mainstream items with respect to another principal component.

2.8 Factor analysis and discriminant analysis

Factor analysis is used to describe variability among observed variables in terms of a potentially lower number of unobserved variables called factors. In other words, it is possible, for example, that variations in three or four observed variables mainly reflect the variations in a single unobserved variable, or in a reduced number of unobserved variables. Factor analysis searches for such joint variations in response to unobserved latent variables.

The observed variables are modeled as linear combinations of the potential factors, plus "error" terms. The information gained about the interdependencies between observed variables can be used later to reduce the set of variables in a dataset. Factor analysis originated in psychometrics, and is used in behavioral sciences, social sciences, marketing, product management, operations research, and other applied sciences that deal with large quantities of data.

Factor analysis is related to principal component analysis (PCA) but not identical. Because PCA performs a variance-maximizing rotation of the variable space, it takes into account all variability in the variables. In contrast, factor analysis estimates how much of the variability is due to common factors ("communality").

The two methods become essentially equivalent if the error terms in the factor analysis model (the variability not explained by common factors, see below) can be assumed to all have the same variance

Discriminant analysis is used in statistics and machine learning to find a linear combination of features which characterize or separate two or more classes of objects or events. The resulting combination may be used as a linear classifier, or, more commonly, for dimensionality reduction before later classification.

Discriminant analysis is closely related to ANOVA (analysis of variance) and regression analysis, which also attempt to express one dependent variable as a linear combination of other features or measurements. In the other two methods however, the dependent variable is a numerical quantity, while for LDA it is a categorical variable (i.e. the class label).

LDA is also closely related to principal component analysis (PCA) and factor analysis in that both look for linear combinations of variables which best explain the data. LDA explicitly attempts to model the difference between the classes of data.

PCA on the other hand does not take into account any difference in class, and factor analysis builds the feature combinations based on differences rather than similarities.

Discriminant analysis is also different from factor analysis in that it is not an interdependence technique: a distinction between independent variables and dependent variables (also called criterion variables) must be made.

2.9 Multidimensional scaling

This family of projection methods aim to reduce the data dimensionality is to optimize the representation in the lower-dimension space so that the distances between points in the projected space are as similar as possible to the distances between the corresponding points in the original space [9].

We describe here a class of methods known as *multidimensional scaling* (MDS). The aim of these methods is to project data from a pseudo-metric space (i.e. characterised by a dissimilarity measure) onto a metric space.

Such methods are especially useful for pre-processing non-metric data in order to use them with algorithms only valid with metric input.

The first MDS method is the *metric MDS*, characterized by minimizing the squared error cost function:

$$E_M = \sum_{k \neq l} \left(d(k,l) - d'(k,l) \right)^2 , \tag{17}$$

where, for the original items x_k and x_l, $d(k,l)$ is their dissimilarity, and $d'(k,l)$ is the distance between the corresponding vectors from the projected metric space.

If the components of the data vectors are expressed on an ordinal scale, a perfect reproduction of the Euclidean distances may not be the best goal.

In such a situation, only the rank order of the distances between the vectors is meaningful.

The error function is defined as

$$E_N = \frac{\sum_{k \neq l} \left(f(d(k,l)) - d'(k,l) \right)^2}{\sum_{k \neq l} \left(d'(k,l) \right)^2}, \tag{18}$$

where *f* is a monotonically increasing function that acts on the original distances and always maps the distances to such values that best preserve the rank order.

Another non-linear mapping method, the *Sammon's mapping*, is closely related to the metric MDS. The only difference is that the errors in distance preservation are normalized with the distance in the original space.

Thus, preservation of small distances is emphasized. The error function is defined as:

$$E_S = \sum_{k \neq l} \frac{\left(d(k,l) - d'(k,l) \right)^2}{d(k,l)}. \tag{19}$$

3. Overview of experiments with multivariate analysis of ECG data

The paper [11] considers the problems pertinent to signal processing in exercise ECG analysis and provides an overview of algorithms employed by research groups as well as manufacturers. The clinical utility of computer measurements and criteria for ECG changes in patients with suspected coronary artery disease is treated.

The paper [23] used exercise ECG tests as the most widely used noninvasive method of evaluating myocardial ischemia. To determine whether it is possible to enhance its diagnostic power, the paper describes an experiment with computerized ECG measurements on 118 nonischemic patients and 38 ischemic patients with or without myocardial infarctions.

The paper [7] studies the spatial properties of ischemic changes induced by prolonged percutaneous transluminal coronary angioplasty (PTCA). It has also been analyzed how they are related to different indexes measured on the ECG (ST level, T-wave amplitude and position, QT interval and QRS duration) or derived from it [integrated measures based on the Karhunen-Loeve (KL) transform and applied to different ECG intervals]. The variations during the occlusion period of the different indexes were used in a multivariate discriminant analysis to determine which indexes showed the best discrimination of the three occlusion sites. Occlusions in the CIR were the most difficult to classify. The use of different ECG measurements (from different intervals) on quasi-orthogonal leads

has permitted the identification of the occluded artery in patients undergoing PTCA and may be extended for more general uses.

The paper [2] investigates an approach in human identification. For this purpose, a standard 12-lead electrocardiogram (ECG) recorded during rest is used. Selected features extracted from the ECG are used to identify a person in a predetermined group. Multivariate analysis is used for the identification task. Experiments show that it is possible to identify a person by features extracted from one lead only. Hence, only three electrodes have to be attached on the person to be identified. This makes the method applicable without too much effort.

The paper [8] presents an application of principal component analysis (PCA) to ECG processing. For this purpose the ECG beats are time-aligned and stored in the columns of an auxiliary matrix. The matrix, considered as a set of multidimensional variables, undergoes PCA. Reconstruction of the respective columns on the basis of a low dimensional principal subspace leads to the enhancement of the stored ECG beats.

A few modifications of this classical approach to ECG signal filtering by means of a multivariate analysis are introduced. The first one is based on replacing the classical PCA by its robust extension. The second consists in replacing the analysis of the whole synchronized beats by the analysis of shorter signal segments. This creates the background for the third modification, which introduces the concept of variable dimensions of the subspaces corresponding to different parts of ECG beats.

The experiments performed show that introduction of the respective modifications significantly improves the classical approach to ECG processing by application of principal component analysis.

4. Concluding remarks

The monitoring of ECG signal and ECG beats recognition are very important tasks in the coronary intensive unit. The classification of the ECG beats is essential tool for a fast and good diagnosis. Till now, it had been developed for recognition and classification of ECG. In the same time, the ECG wave analysis, the detection of QRS complex, which is associated with electrical ventricular activation are very important in the detection of cardiac anomalies.

In this paper we have studied the application of the fuzzy multivariate data analysis in medical field, especially for ECG recognition and classification.

We showed the main mathematical aspects of fuzzy as fuzzy clustering, hierarchic fuzzy clustering with incomplete data, fuzzy regression, fuzzy principal components analysis, multidimensional scaling etc.

In the future research work, we will try to apply this method in analysis of the ECG data base existed in hospital. The acquisition of ECG signal will be doing with our own telemetric system based on Microchip microcontroller interconnected with a PC.

The analysis of the ECG signal in real time gives to the doctor the possibility to find out the best medical decisions and more real chances to the patient to survey.

R E F E R E N C E S

[1] J.C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, New York, 1981.

[2] L. Biel, O. Pettersson, L. Philipson, P. Wide, *ECG analysis: a new approach in human identification, IEEE Transactions on Instrumentation and Measurement*, 50 (3), 2001, 808-812.

[3] T.R Cundari, C. Sârbu, H.F. Pop, *Robust Fuzzy Principal Component Analysis (FPCA). a comparative study concerning interaction of carbon-hydrogen bonds with molybdenum-oxo bonds,* Journal of Chemical Information and Computer Sciences 42, 6 (2002), 1363-1369.

[4] D. Dumitrescu, *Hierarchical pattern classification*, Fuzzy Sets Syst., 28, 1988, 145-162.

[5] M. Frențiu, H.F. Pop, *Tracking mistakes in software measurement using fuzzy data analysis,* In The 4-th International Conference RoEduNet Romania (Sovata, Târgu-Mures, Romania, May 20-22 2005), pp. 150-157. Organized by the Romanian Education Network RoEduNet and the ``Petru Maior'' University, Târgu-Mures.

[6] M. Frențiu, H.F. Pop, *Overview of fuzzy methods for effort estimation by analogy,* Studia Universitatis Babes-Bolyai, Series Informatica 1, 2009, 68-74.

[7] J. Garcia, G. Wagner, L. Sornmo, P. Lander, P. Laguna, *Multivariate discriminant analysis of ECG-based indexes to identify the occluded artery in patients undergoing PTCA*, Engineering in Medicine and Biology Society, 1998. Proceedings of the 20th Annual International Conference of the IEEE, 29 Oct - 1 Nov 1998, vol. 1, 182-185.

[8] M. Kotas, *Application of projection pursuit based robust principal component analysis to ECG enhancement*, Biomedical Signal Processing and Control 1-4 October 2006, 289-298.

[9] J.B. Kruskal, M. Wish, *Multidimensional Scaling*, Sage Publications, Beverly Hills, CA, 1977.

[10] M. Mellinger, *Multivariate Data Analysis: Its Methods*, Chemom. Intell. Lab. Syst. 2, 1987, 29-36.

[11] O. Pahlm, L. Sörnmo, *Data processing of exercise ECG's*, IEEE Trans. Biomed. Eng. 34, 1987, 158–165.

[12] H.F. Pop, C. Sârbu, *A New Fuzzy Regression Algorithm*, Anal. Chem. 68, 1996, 771-778.

[13] H.F. Pop, *Principal Components Analysis based on a fuzzy sets approach*, Studia Universitatis Babes-Bolyai, Series Informatica 46, 2, 2001, 45-52.

[14] H.F. Pop, T.L. Pop, C. Sârbu, *Assessment of heart disease using fuzzy classification techniques*, The Scientific World Journal 1 (2001), 369-390. ISSN 1537-744X.

[15] H.F. Pop, M. Frentiu, *Applications of principal components methods,* in Complexity in Artificial and Natural Systems, Proceedings (2008), Editura Universității „Petru Maior", Târgu-Mureș, pp. 267-270.

[16] H.F. Pop, J.W. Einax, C. Sârbu, *Classical and fuzzy principal component analysis of some environmental samples concerning the pollution with heavy metals*, Chemo metrics and Intelligent Laboratory Systems 97, 1 (2009), 25-32.

[17] C. Sârbu, H.F. Pop, *Fuzzy robust estimation of central location*, Talanta, 54, 2001, 125-130.

[18] C. Sârbu, H.F. Pop, *Fuzzy classification and comparison of some Romanian and American coals*. MATCH - Communications in Mathematical and in Computer Chemistry, 44 (October 2001), 387-400.

[19] C. Sârbu, H.F. Pop, *Fuzzy soft-computing methods and their applications in chemistry*, In Reviews in Computational Chemistry, K. B. Lipkowitz, R. Larter, and T. R. Cundari, Eds., vol. 20. VCH Publishers, New York, 2004, pp. 249-332.

[20] C. Sârbu, H.F. Pop, *Principal Component Analysis versus Fuzzy Principal Component Analysis. A case study: The quality of Danube water* (1985-1996), Talanta 65, 2005, 1215-1220.

[21] C. Sarbu, H.F. Pop, R.S. Elekes, G. Covaci, *Intelligent disease identification based on discriminant analysis,* Revista de Chimie 59, 11 (2008), 1237-1241.

[22] C. Serban, A. Vescan, H.F. Pop, *A new component selection algorithm based on metrics and fuzzy clustering analysis.* in Proceedings of the 4th International Conference on Hybrid Artificial Intelligence Systems. *Lecture* Notes in Artificial Intelligence, LNCS 5572/2009 (Salamanca, Spain, June 10-12 2009), pp. 621-628.

[23] H. Sievänen, L. Karhumäki, I. Vuori, J. Malmivuo, *Improved diagnostic performance of the exercise ECG test by computerized multivariate ST-segment/heart rate analysis*, Journal of Electrocardiology 24 (2), April 1991, 129-143.

[24] A. Vescan, H.F. Pop, *Automatic configuration for the component selection problem*, in CSTST '08: Proceedings of the 5th international conference on Soft computing as transdisciplinary science and technology (Cergy-Pontoise, France, October 27-31 2008), ACM, New York, NY, USA, pp. 479-483.

[25] S. Wold, *Principal Component Analysis*. Chemom. Intell. Lab. Syst., 2, 1987, 37-52.

[26] A. L. A. Zadeh, *Fuzzy sets*, Inf. Control, 8, 1965, 338-353.