FUZZY MATCHING IN CLASSIFICATION AND INTERPRETATION OF ECG SIGNALS

G. Bortolan*, R. Degani*, K. Hirota**, W. Pedrycz***

* LADSEB-CNR, Padova, Corso Stati Uniti 4, Italy

** Dept. of Instrument & Control Engineering, College of Engineering, Hosei University, Tokyo, Japan

*** Department of Electrical Engineering, University of Manitoba, Winnipeg, Manitoba, Canada R3T 2N2

Abstract

We deal with the problem of classifying electrocardiograms in the presence of uncertain fuzzy environment. Referring to the classification scheme, both the labels used to describe the features of the patterns and class assignment (membership functions) are viewed as fuzzy sets. The main classification structure consists of three stages: (i) matching the input patterns; (ii) transformation of the result of matching from the feature space to the space of class assignment. The classifier is designed (trained) with the aid of real-world data set.

Keywords

Fuzzy pattern matching, equality index, computerized ECG analysis and classification.

1. Introduction

In many real-world pattern recognition problems, a factor of uncertainty cannot be neglected. Its influence depends heavily on a complexity of classification tasks at hand. Usually two main sources of uncertainty pertaining to fuzziness are reported, namely fuzziness coming from the process of feature determination and uncertainty related to class labelling. This usually holds true in medicine or biology, that a particular pattern (object) could belong to some classes at a certain degree varying between 1 (complete belongingness) or 0 (which excludes any link of the pattern to the class being considered). The first source of uncertainty mentioned above results as a consequence of properties of cognitive processes of the human being. In existing references concerning ECG interpretation, an occurrence of these sources of fuzziness can be easily observed, cf. [1] [4]. The aim of this paper is to discuss an application of fuzzy classification scheme in the field of analysis, recognition and interpretation of ECG signals. An underlying background assumes that we deal with fuzzy labels as well as with imprecise class assignment values.

The paper is organized in three sections. First of all, we consider the structure of the classifier and discuss the relevant matching principle. In sequel we present numerical results concerning the performance of the classifier.

2. A structure of the classification procedure

Let us consider a feature space **X** consisting of "m" coordinates, i.e.

$$\mathbf{X} = \mathbf{X}_1 \times \mathbf{X}_2 \times \ldots \times \mathbf{X}_m = \mathop{\mathbf{X}}_{i=1}^{m} \mathbf{X}_i$$

For each coordinate of **X** we deal with a set of linguistic labels having a clear interpretation in a context of the specified application. The labels are represented as fuzzy sets. For $\mathbf{X}_k$ we get

$$\chi_1^{(k)}, \chi_2^{(k)} \ldots \chi_{n_k}^{(k)} : \mathbf{X}_k \to [0,1] \quad k=1,2,\ldots,m$$

Let "c" classes $\omega_1, \omega_2 \ldots \omega_c$ also be distinguished. For each pattern in a learning set a class membership function can be specified. It is put down in a vector form,

$$\Omega = \left[\Omega(\omega_1)\,\Omega(\omega_2)\ldots\Omega(\omega_2)\right]$$

where $\Omega(\omega_l)$, $l = 1,2,\ldots c$, denotes a grade of membership to the l-th class.

A general structure of the classifier is displayed as follows:

$$\boxed{\text{matching}} \longrightarrow \boxed{\text{feature-class transformation}} \longrightarrow \boxed{\text{inverse matching}}$$

A key issue which plays a significant role in the above classifier refers to matching of fuzzy quantities. Due to existence of fuzzy labels the features of two patterns match to a certain degree (partial matching). This result of matching strongly influences (via a transformation block namely feature-class transformation) a degree of matching of their class membership. As known in literature in fuzzy sets two fuzzy sets could be compared in different ways. One alternative studied in detail in [3] has a straightforward logical interpretation. Denote by "a" and "b" grades of membership which are to be compared.

The equality index $a \equiv b$ can be conveniently defined in the following ways:

(i) $a \equiv b = (a \to b) \,\&\, (b \to a)$ (1)

and

(ii) $a \equiv b = (\bar{a} \to \bar{b}) \,\&\, (\bar{b} \to \bar{a})$ (2)

In (1) and (2) implication, &, and ¯ are represented by means of pseudocomplements, t-norms and negation operator. Thus one has:

$$a \to b = \sup\left\{c \,\varepsilon\, [0,1] \,\middle|\, a\,t\,c \le b\right\},$$

$$a \,\&\, b = t(a,b)$$

$$\bar{a} = 1 - a.$$

where "t" denotes a triangular norm. Relevant plots of (1) and (2) for t-norm specified as product are contained in Figs. 1 and 2, respectively.

It is evident that they are asymmetrical around a point at which the equality index attains 1.0. To diminish this property it is worthwhile to take an average value of (1) and (2). Thus we have:

$$a{\equiv}b=\frac{1}{2}\left\{(a\varphi b) \wedge (b\varphi a) + [(1-a)\varphi(1-b)] \wedge [(1-b)\varphi(1-a)]\right\} \qquad (1)$$

a,b ε [0,1]. This equality index is displayed in Fig. 3.

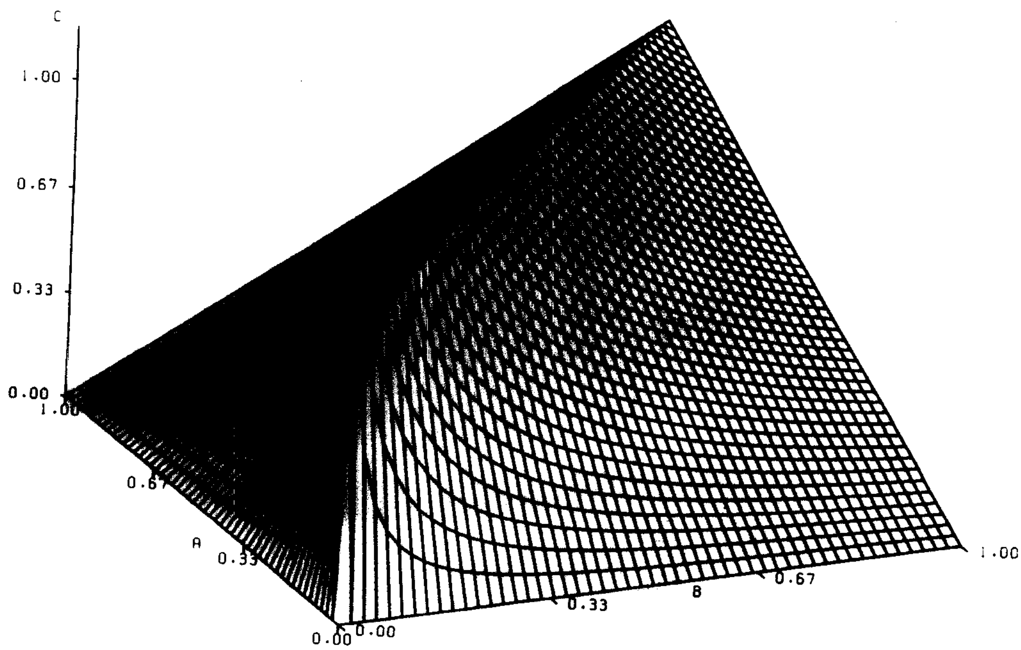

Fig. 1. Plots of the Equality Index (1)

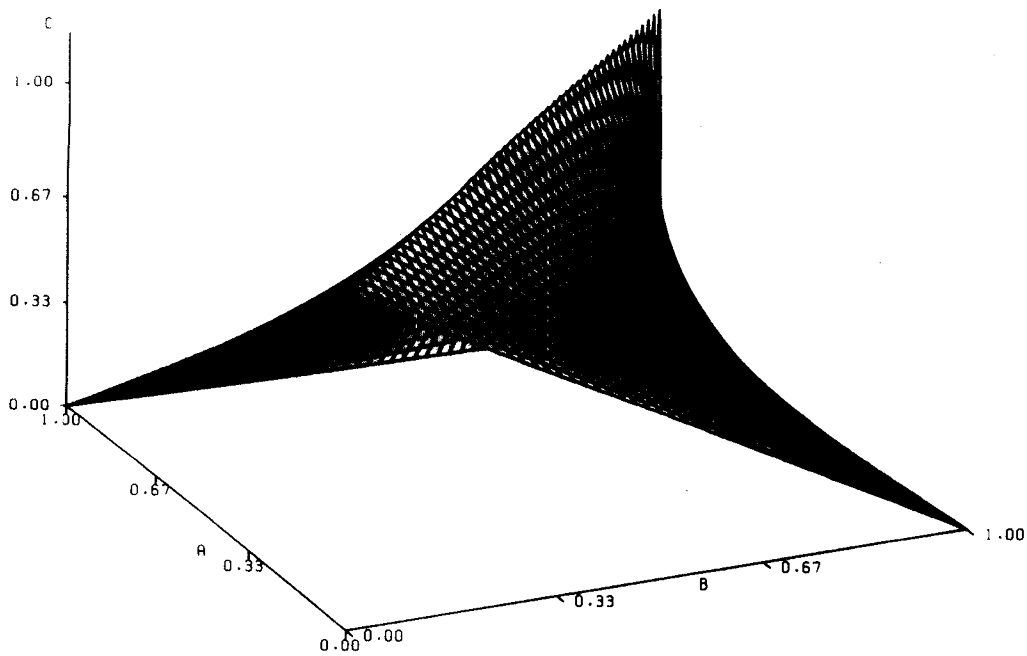

Fig. 2. Plots of the Equality Index (2)

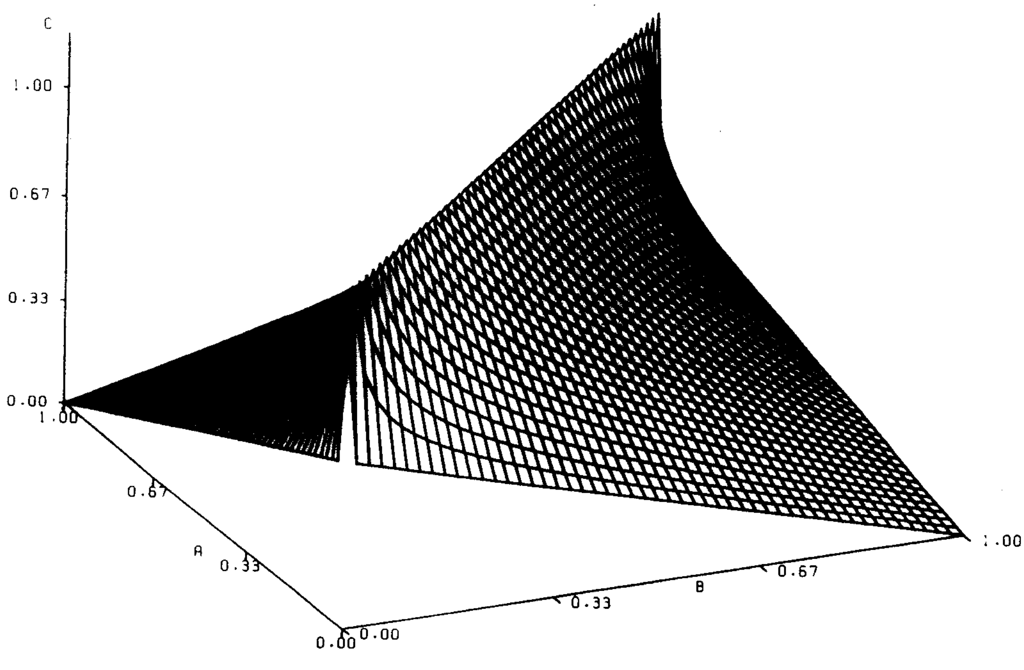

Fig. 3. Plots of the Equality Index (3)

An interesting problem arises when one is going to determine all grades of membership x satisfying an inequality

$$a \equiv x \geq \gamma \tag{4}$$

for a and γ provided. It has been proven in [3] that (4) always possesses a nonempty set of solutions forming a subinterval of [0,1]. For γ=1.0 it reduces to one point (equal to {a}) while for γ=0 we get the entire unit interval.

Returning to the general scheme of pattern classification let us discuss its components. Matching is performed with the aid of the equality index (3) knowing the grades of membership of the pattern and the prototype to different linguistic lables specified in the coordinates of the feature space. The result of matching is a vector η for which a number of coordinates is equal to the number of the features, while the i-th coordinate is equal to an averaged value of the equality index of the consecutive grades of membership taken with respect to the existing linguistic labels. In sequel η is transformed into an equality index for each element of the class assignment space, so we deal with "c" different transformation function $F(\cdot,\omega_1), F(\cdot,\omega_2),\ldots, F(\cdot,\omega_c)$. Hence $\xi_1 = F(\eta,\omega_1)$, $\xi_2 = F(\eta,\omega_2)$, ... $\xi_c = F(\eta,\omega_c)$. They specify how strongly the class membership of discussed pattern resembles the class membership of the imposed prototype. In sequel we make use of inverse matching. This stage is based on the solution of the inverse problem. Then, not aggregating x after the prototype has already been selected, a resolution of the inverse problem is formulated in such a framework: for each ω_i determine solutions of inequalities $b_i \equiv x \geq \xi_i$ i=1,2,...,c where b_i denotes a grade of membership of the prototype at the class ω_i. Hence the result is obtained in the form of the interval $[b_{i-} \; b_{i+}]$. Its width as well as its distribution in the unit interval would serve as a good means for graphical displaying of the precision of the matching. The determination of the transformation functions could be conveniently realized in a form of a neural network, cf. [2].

An overall structure of the system of ECG classification is presented in Fig. 4.

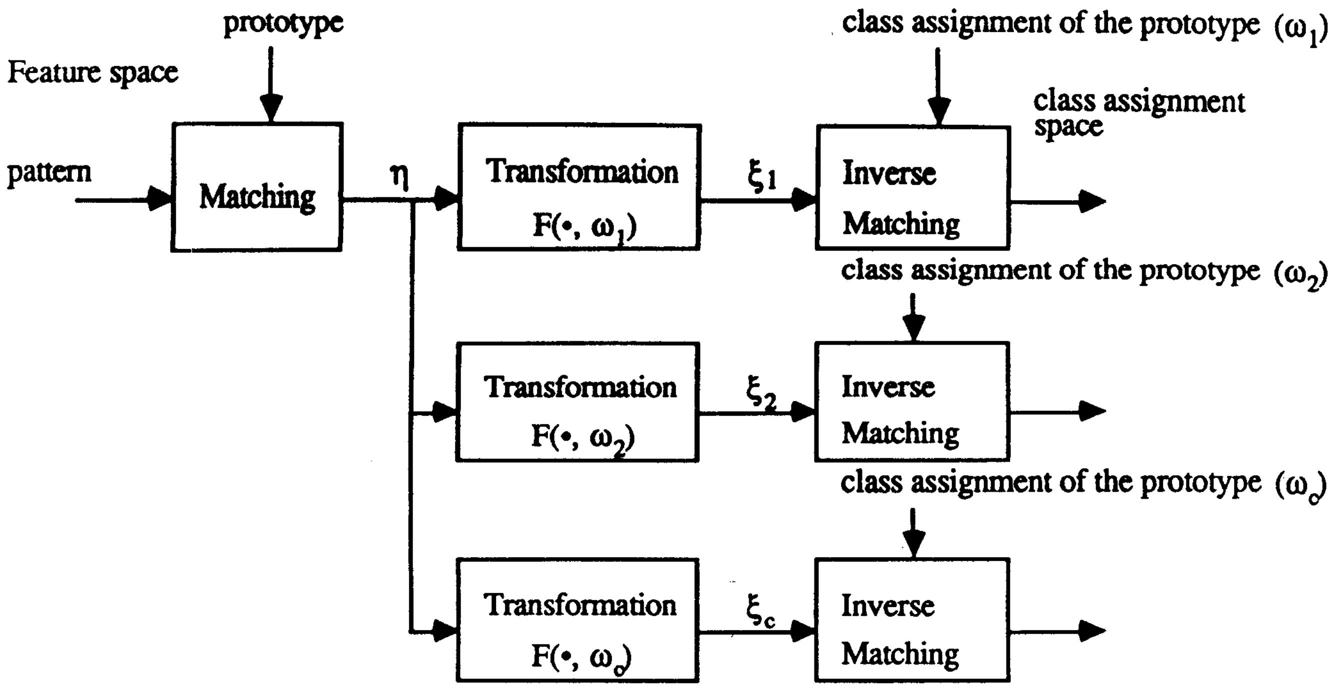

Fig. 4. The Structure of the Classification System

3. Numerical Studies

The standard 12-lead electrocardiogram is a group of 12 signals coming from leads named I, II, III, aVR, aVL, aVF, V1, V2, V3, V4, V5, V6. They represent the electrical activity of the heart recorded from the body surface. These signals convey an important information concerning the electrical behavior of the heart. The features are related to durations, amplitudes, areas and shapes of the various waves in each lead. Despite a long tradition of electrocardiography (over 70 years) the definitions of the basic paramters are not clearly stated. This is of particular interest in computerized classification system. To overcome inadequate standardization of definitions and of measurement rules, a European project called Common Standards for Quantitative Electrocardiography (CSE) has been launched in 1980 [4].

A data set at hand consists of 200 patterns (ECG signals) assigned to three different class, ω_1, ω_2 and ω_3 namely, Normal, Left Ventricular Hypertrophy and Right Ventricular Hypertrophy. Each pattern is characterized by a matrix containing degrees of possibility of each feature (16 different features have been taken into account) with respect to three fuzzy labels (quantities) defined in every coordinate of the feature space. These degrees specify to which extent the particular feature fits any of these catagories (the categories were labelled as NORMAL, BORDERLINE and ABNORMAL). The set of grades of membership to the classes taken into account consists of the following expressions: complete exclusion (0.0); fair membership (0.33); strong membership (0.66); and complete belongingness (1.0). The membership functions of different subclasses are summarized in Table 1.

no. of the subclass	membership function	no. of elements
1	[0 0 0.33]	1
2	[0 0 0.66]	2
3	[0 0 1]	4
4	[0 0.33 0]	1
5	[0 0.33 0.33]	2
6	[0 0.66 0]	7
7	[0 1 0]	10
8	[0.33 0 0]	9
9	[0.33 0 0.33]	2
10	[0.33 0.33 0]	2
11	[0.33 0.33 0.33]	2
12	[0.66 0 0]	11
13	[0.66 0 0.33]	1
14	[0.66 0.33 0]	2
15	[1 0 0]	44

Table 1. Membership functions of different subclasses

In the general scheme already discussed in the previous section, the linear one-layer neural net is implemented, viz. the equation of the input-output relationships are expressed by

$$\xi_t = w_t^T \eta, \quad t=1,2,3 \qquad \xi = \left[\xi_1 \, \xi_2 \, \xi_3\right]^T$$

. The classification rule for any input pattern is based upon results of its matching with the prototype of each class. Thus the following sequence of steps should be realized: match the input pattern with the prototype of each class. This yields η and in sequel $w^T\eta=\xi$. Assign the pattern to the class for which a sum of coordinates of ξ attains maximum. Simultaneously ξ could serve as a measure of goodness of fit. It is noticeable to discuss some categories of classification errors (here Ω refers to the original vector of membership while $\hat{\Omega}$ denotes the result of classification):

(i) lack of any misclassification error;

(ii) the fuzzy sets Ω and $\hat{\Omega}$ specify exactly the same class or classes but with different degrees of membership;

(iii) the fuzzy sets Ω and $\hat{\Omega}$ have a non-empty intersection. It expresses a situation in which different classes are indicated, however, there are also some classes for which both Ω and $\hat{\Omega}$ have nonzero grade of membership;

(iv) complete misclassification.

The overall statistics concerning classification errors with the categories distinguished is summarized in a table below,

number of errors of the category	**i**	**ii**	**iii**	**iv**
patterns 1-100 (learning set)	40	22	24	14
patterns 101-200 (testing set)	29	41	19	11

Considering the fourth category of error, we easily observe that the classification error is equal to 14/100=14% for the learning set, 11/100-11% for the testing set and 25/200=12.5% for the whole data set. These numbers are not high bearing in mind the number of classes distinguished (15) and the number of all the patterns available.

References

1. R. Degani, G. Bortolan, Computerized electrocardiogram diagnosis: fuzzy approach. In: Encyclopedia of Systems and Control, M.G. Singh (ed.), Pergamon Press, Oxford, 1987, 760-769.

2. G.A. Carpenter, S. Grossberg, A massively parallel architecture for a self-organizing neural pattern recognition machine, Computer Vision, Graphics, Image Processing, 37, 1987, pp. 54-116.

3. W. Pedrycz, Comparison of fuzzy sets and inverse problem, J. of Franklin Institute, to appear.

4. J.L. Willems (ed.), Common standards for quantitative electrocardiography, CSE Progress Reports, 1981-1987, Leuven.