Proceedings of the 7th WSEAS International Conference on Fuzzy Systems, Cavtat, Croatia, June 12-14, 2006 (pp39-42)

Measuring Feature uncertainty by using similarity

BARNA CORNEL
"Aurel Vlaicu" University
Arad
ROMANIA

Abstract: - This article proposes a measurement of the uncertainty used in feature labeling which takes into consideration the similarity between the analyzed characteristic and a prototype, and the dissimilarity between the same characteristic and the nearest neighboring prototype. Using the similarity and not the equivalent relationship, the method is more general, the transitivity not being requested.

Key-Words: Fuzzy relation, similarity, rough set, information fusion, uncertainty, featutes extraction

1 Introduction:

Determining the degree of confidence for the estimated characteristics is one of the most important issues in feature discovery and labeling. It has been underlined many times [1], [2] that crisp interpretation of classification results implies loosing an important part of the initial information, and for this reason it is preferable to use a fuzzy form for out coming data, or to establish a continence measure. In these cases, an uncertainty factor can be used. Uncertainly is formalized in a various number of methods: by using probability, fuzzy membership, possibility measurement, by using Demster-Shaffer masses, or rough sets. Most of them are using an interval to define an uncertain domain, in which a crisp affirmation about the studied feature or characteristic cannot be strictly determined. Even for the fuzzy membership Atanassov proposed a degree of unbelongingness [3], as a contra part for Zadeh's fuzzy membership. Thus, in posibilistic approach, the limits of the uncertainty are the necessity and the possibility of belonging to a characteristic [4]. The main relationship is the numerical equivalence between the degree of possibility and the degree of membership, as defined in fuzzy logic:

$$\pi(x) = \mu(x) \qquad (1)$$

From this relationship, the measure of possibility is deduced:

$$\Pi(X) = \max \pi(x) \qquad (2)$$

where $x \in X$ is an element of the definition subset. It's dual, the measure of necessity is:

$$N(A) = 1 - \Pi(\neg A) \qquad (3)$$

These two measurements represent the extreme situations. The first represents any possible belonging to the set, and the second one represents the absolute positive situation.

One interesting extension of the uncertain measurement using possibilities is presented in [5], where the limits are established on intervals, creating, in this way, a granular reconstruction of a characteristic function. For each interval, the upper border of the reconstruction is used as the possibility measurement:

$$\hat{R} = \sup[a \in [0,1] \mid a \otimes A(x) \le \lambda] \quad (4)$$

where

$\otimes$ is a t-norm, and λ is the possibility measure of the *A(x)* attribute.

The relationship for the lower border has the form:

$$\tilde{R} = \inf[a \in [0,1] \mid a \oplus (1 - A(x)) \geq \mu] \quad (5)$$

where

$\oplus$ is the used co-norm, and μ is the necessity measure of the *A(x)* attribute

In the Demster-Shaffer theory, the limits are the credibility and the plausibility [6], defined by

$$\text{Bel(A)} = \Sigma_{B \subseteq A} \text{m(B)} \quad (6)$$

$$\text{Pl(A)} = \Sigma_{B \cap A \neq \phi} \text{m(B)}$$

They are based on masse evaluation, representing the maximal, respectively the minimal belief.

Pawlak's rough set theory defines the lower $\underline{R}$ and the upper $\overline{R}$ approximation [6] in the description of an object by using a set of attributes.

$$\underline{R}(X) = \{x \in U \mid [x]_R \subseteq X\} \quad (7)$$

where $X \subseteq U$,

$$\overline{R}(X) = \{x \in U \mid [x]_R \cap X \neq 0\} \quad (8)$$

The lower approximation represents the certain inclusion of the related elements in the considered subset *X* while the upper approximation represents all elements that have at least a common point with the subset. The union of all lower approximations determines a certain belonging to the subset, called the positive region

$$POS(X) = \bigcup_{X \in U} \underline{R}X \quad (9)$$

while the negative region includes all the regions which certainly have no common points with the subset *X*.

$$NEG(X) = U - \bigcup_{X \in U} \overline{R}X \quad (10)$$

The zone between the positive and the negative regions is called the boundary region. This region corresponds to the uncertain zone, in which the membership of the points to the subset cannot be determined.

$$BND(X) = \bigcup_{X \in U} \overline{R}X - \bigcup_{X \in U} \underline{R}X \quad (11)$$

One of most important observations, which is connected to the rough sets is the fact that *R* is creating an equivalence partition inside the *X* domain. The equivalence relationship determines an indistinctibility of the attributes inside each class.

2 Approach

All these formulations about uncertainty have a very strict delimitation for the bounders. Nevertheless, it must be noticed that, in practice, a flexibility of these limits exists. Some extreme points can be overlooked. The uncertain region in feature labeling is sometimes more related to the relationship between the characteristics, and less related to exception points. This article presents an uncertainty measurement based on similarity to the prototype and on dissimilarity from the nearest neighbor prototype.

Equivalence classes imply three main proprieties: reflexibility, symmetry and transitivity. In numerous applications, the transitivity doesn't hold (see Luce paradox), fact determined by the approximation of the observation (two characteristic measurements seam to be equal, but in reality there is a little difference between them). For this reason, the use of a more general relationship like similarity, in which the transitive propriety is not required, is more adequate. This is especially true for granular computing and fuzzy logic, where the observations are rough. In this article a fuzzy similarity relationship proposed in [8] will be used. This relationship is obtained by overlaying two fuzzy relationships, that of inclusion

and that of dominance. The first one implies a fuzzy implication relationship, and establishes the degree of inclusion of a variable *x* in an interval boarded by a constant value *a*. The second one implies a fuzzy implication between a boarded constant *b* and the variable *x*. Thus, the similarity relationship is:

$$sim(x,a,b) = (a \to x) \otimes (x \to b) \quad (12)$$

If the fuzzy implications have the form :

$$a \to x = \sup[c \in [0,1] \mid a \otimes c \le x] \quad (13)$$

then, for a t-norm that is considered to be a product, the relationship become:

$$a \to x = \begin{cases} 1\_daca\_x \le a \\ \dfrac{a}{x}\_daca\_x > a \end{cases} \quad (14)$$

or, for a Lukasiewicz t-norm it becomes :

$$a \to x = \begin{cases} 1\_daca\_x \le a \\ 1 - x + a\_daca\_x > a \end{cases} \quad (15)$$

The graphical representations of the product t-norm similarity relationships ispresented in fig.1

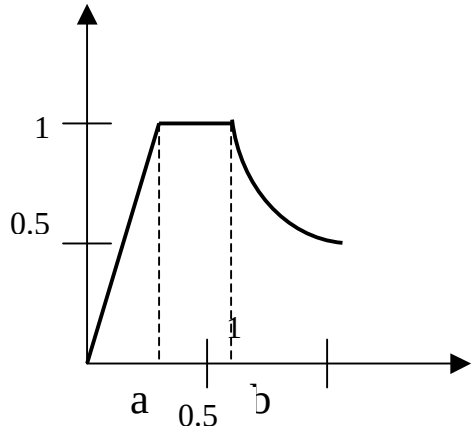

Fig.1

And the graphical representations of the Lukasiewicz t-norm similarity relationships is presented in fig.2

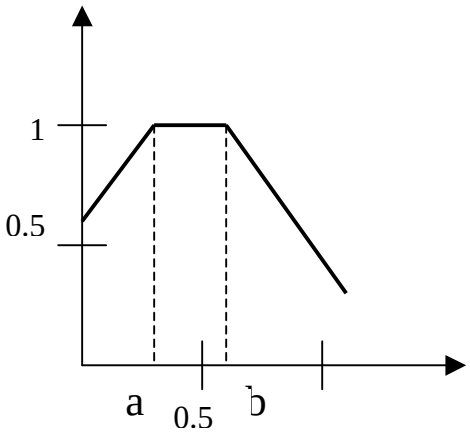

Fig.2

Thus, the label of the unknown feature is determined by the nearest prototype, which is also the most similar to it:

$$S(x,a,b) = \max(sim(x,a,b)) \quad (16)$$

Also, the feature is compared to the next nearest prototype. The dissimilarity is then determined from this reference point. The dissimilarity is considered the complementary measure of the similarity, and its relation is determined by using the negation of similarity. In the present paper, the classical negation will be used. The resulting relationship for this measure is:

$$D(x,a,b) = \min_{i=1,N} (1 - sim(y, a_i, b_i)) \quad (17)$$

From the relationship between these two measurements the measure of uncertainty can be determined, expressed as a difference between the similarity with the closest prototype and the dissimilarity from the second closest prototype:

$$U(x,a,b) = S(x,a,b) - D(x,a,b) \quad (18)$$

If the result of this measurement is positive, then small values indicate small uncertainty, because the labeled feature is similar to the designed prototype, and is weakly similar to the next closest one.
In the case of $U(x,a,b)>0$ being large, it can be concluded that uncertainty is

important, because the difference of similarity between the feature and the next nearest prototypes is small, and the possibility of it belonging to each of these two classes is high, resulting that the probability of making a mistake is considerable.
If the difference has a negative value, it can be concluded that the feature does not belong to the investigated class, because it is more similar to another prototype, or it can be labeled on the information provided by the existent prototypes.

3 Result

A set of 8 experiments with features extracted from images was performed. The goal of the experiments was to recognize capital letters, more specifically: A, E, H, I, O, and U. This was done by extracting the area and the second order moment from the images of the letters. The experiments are simple, but suggestive for the purpose of this paper. As expected, by observing the distribution of the area values, a partition of the letters in three clusters results, namely one with A, E and H, second with O and U and last containing only the letter I. Between letters from different classes, the measure of uncertainty $U(x,a,b)$ was small, but inside the classes the value was bigger, resulting in a more uncertain decision. By analyzing the feature corresponding to second order moments, it can be observed that several groups of letters exist; for example: O, U or E, H are quite similar, thus the selection of one of them has more uncertainty. Also, the uncertainty measure described before has large values, but for features corresponding to the letter I, it has a small value because the confidence in this selection is high, and the characteristic of this feature is far from any other one.

4 Conclusion

In many uncertainty measurement methods a bi-value feature determination is used, in which the degree on uncertainty is established between two limits. In this paper, a method in which the limits are not so rigid is exposed. They are based on two fuzzy relationships: the similarity and dissimilarity of the unknown feature, which is requited to be labeled, and the two most nearest prototypes. It had been proved that the method is sound, and leads to good results.
Also, the method can be generalized, by taking into consideration the k-nearest neighboring prototypes for determining the dissimilarity measure. Further investigation will be done in this direction.

References:

[1] R.Yager: A General approach to the Fusion of Imprecise Information, *Inter'l Journal of Intelligent system* Vol.12 1997

[2] F.Alkoot,J.Kittler: Modified product fusion, *Pattern Recognition Letters* Vol 23,Nr.8 2002

[3] K.Vlackos, G.Sergiadis On the Entropy of Intuitionistic Fuzzy Events, *CIMCA* 2005 Wiena pag.162-167

[4] D.Dubois,H.Prade: Possibility Theory in Information Fusion *ISIF* 2000, pag.6-19

[5] W.Pieczynski: Unsupervised Dempster-Shafer Fusion of Dependent Sensor *IEEE ICFS* 2004

[6] G.Bortolan, W.Pedrycz Reconstruction problem and information granularity *IEEE Trans. On Fuzzy Systems*, vol.2 1997 pag.234-248

[7] Z. Pawlak, Rough Sets, *International Journal of Computer and Information Sciences* vol.11 1982 pag.341-356

[8] W.Pedrycz *Knowledge-Based Clustering From Data to Information Granules* Ed.Wiley 2005

[9] A.Bargiera, W,Pedrycz, K.Hiroto Logic-based granular prototyping *26-th Annual International Computer Software COMPSAC* 2002