EXPRESSING NATURAL LANGUAGE INTO THE *FUZZY SET* FORM AND APPLYING THIS METHOD INTO IMAGE - CLASSIFICATION PROBLEMS FOR ELIMINATING NOISES

Lê Hoài Bắc – Trần Thái Sơn
Trường Đại Học Khoa Học Tự Nhiên
(Bài nhận ngày 21/09/1998)

ABSTRACT: This article contains two parts: In the first part, we specifically introduce our methodology of expressing natural language into the form of fuzzy sets. Furthermore, in its second part, we completely present here two topics: how to apply the fuzzy contrast intensification to image - classification problems to eliminate noises efficiently; and some of our elementary results in the image - classification of bio-cells whose photographs are taken through a microscope.

I. INTRODUCTION

Natural language is the means of human communication. In fact, it is abundant, ambiguous and vague. As a result of these characteristics, a mathematical theory dealing with fuzziness and ambiguity could be suitable for expressing and interpreting the linguistic characters of our language. The structure of a natural language usually includes a limited set of base terms - atoms in the literature, and rules in order to connect terms together into a meaningful group of words. Therefore, in our efforts, we use fuzzy sets as our mathematical foundation of expressing natural language. Fuzzy sets are used to descript terms and prescript the connection of terms into strings of intelligible expressions. Suppose that X is defined as an universal set including atomic terms and sets of atomic terms. Then, call Y as an universe of cognitive interpretations. It is difficult to define Y exactly because of the multi-meaning of a word in natural language. Fortunately, in natural sciences and technology, there is no space for the multi-meaning word in a view angle at an exact time. This means that in a problem, a word just has a meaning in the interval of its observation time. Therefore, we can define Y in a specific space, and an atom in Y is a fuzzy set to express a meaning of an atom in X. In order to solve the multi-meaning's problem of a word, we can build some mappings from X to Y. In theory, it is not reasonable, however in practical technology, since an observation time of a problem is limited, it is acceptable. Suppose that we define a specific atomic term in the universe of natural language X, as an element x, and we define a fuzzy set $\widetilde{A}$ in the universe of meanings Y, as a specific meaning for the term x. Natural language can be expressed as a mapping $\tilde{M}_f$ from X to Y, and each atomic term x in X corresponds to a fuzzy set $\tilde{A}$ in Y.

$$\tilde{M}_f : X \rightarrow Y$$
$$x \rightarrow \tilde{A}$$

This mapping can be denoted $\tilde{M}_f(x, \tilde{A})$ and the membership of $\tilde{M}_f$ is $\mu_{\tilde{M}_f}$. We can evaluate $\mu_{\tilde{M}_f}$ by the following equation

$$\mu_{\tilde{M}_f}(x,y) = \mu_{\tilde{A}}(y)$$

For instances, we have the atomic term x = "*young*" and we want to explain this linguistic atom in term of age y, by a membership function of fuzzy set $\tilde{A}$ that expresses the term "*young*". The membership function given here in the notation of Zadeh might be one interpretation of the term "*young*"

$$\tilde{A} = \int_0^{25} \frac{1}{y} + \int_{25}^{100} \frac{1}{y}\left(1+\left(\frac{y-25}{5}\right)^2\right)^{-1} \text{ or}$$

$$\mu_{\tilde{M}_f}(\text{"young"},y) = \begin{cases} \left(1+\left(\frac{y-25}{5}\right)^2\right)^{-1} & y > 25\,years \\ 1 & y \le 25\,years. \end{cases}$$

In the contrary, we can express the linguistic atom of the term "old" based on this function

$$\mu_{M_f}(\text{"old"},y) = 1-\left(1+\left(\frac{y-50}{5}\right)^2\right)^{-1} \quad for\, 50 \le y \le 100$$

The interpretation of the composite definite on universe Y can be defined by the following set-theoretic operations

x *or* z : $\mu_{x\,or\,z}(y) = max(\mu_x(y), \mu_z(y))$

x *and* z : $\mu_{x\,and\,z}(y) = min(\mu_x(y), \mu_z(y))$

not x : $\mu_{\sim x}(y) = 1 - \mu_x(y)$

II . LINGUISTIC HEDGES

In linguistics, fundamental atomic terms are often modified with adjectives or adverbs such as "very", "low", "slight", "fairly", "barely", "mostly", "approximately", etc. We will call these modifiers "linguistic hedges". These linguistic hedges have the effect of modifying the membership function for a basic term. As an example, a base linguistic atom a is defined by a function

$$a = \int_Y \mu_a(y)/y$$

We assign a some hedges like "very", " very , very", "plus", "slightly", "minus", then the membership function of a has changed

$$"Very"\, a = a^2 = \int_Y \frac{[\mu_a(y)]^2}{y} \qquad (1)$$

$$"Very,\ very"\ a = a^4 \qquad (2)$$

$$"Plus"\, a = a^{1.25} \qquad (3)$$

$$"Slightly"\, a = \sqrt{a} = \int_Y \frac{[\ _a(y)]^{0.5}}{y} \qquad (4)$$

$$"Minus"\, a = a^{0.75} \qquad (5)$$

The expressions shown in *Eqs (1) - (3)* are linguistic hedges known as concentrations. Concentrations tend to concentrate the elements of a fuzzy set by reducing the degree of membership of all elements that are only "partial" in the set. Alternatively, the expressions given in *Eqs (4) - (5)* are linguistic hedges known as dilations. Dilations stretch a fuzzy set by increasing the membership of elements that are "partial" in the set.

Especially, there is an operation which is a combination of concentration and dilation here called intensification. This operation increases the degree of membership of those elements in the set with original membership values greater than 0.5 and it decreases the degree of membership of those elements in the set with original membership values less than 0.5. The membership function of intensification proposed by Zadeh is

$$"Intensify"\ \mathrm{a} = \begin{cases} 2\mu_a^2(y) & 0 \le \mu_a(y) \le 0.5 \\ 1 - 2[1 - \mu_a(y)]^2 & 0.5 \le \mu_a(y) \le 1 \end{cases}$$

Intensification increases the contrast between the elements of the set whose values are larger than 0.5 and ones whose values are less than 0.5.

III. APPLYING THE INTENSIFICATION INTO ELIMINATING NOISES OF IMAGE

The classification of colors in an image is processed by intensifying the brightness of these image pixels. We have an image *X* as $M * N$ dimensions, and the value of the $(m,n)^{th}$

element is considered as the colors of the $(m,n)^{th}$ pixel. To fuzzify an image, we select a maximum constant called Max with condition

$\forall x \in X, \; Max \geq x,$

then the membership value of an image pixel is

$$\mu_x = \frac{x}{Max}, \quad 0 \leq \mu_x \leq 1, \qquad (7)$$

with x is the color value of x pixel .

The contrast of an image is evaluated on the brightness of each pixel, so a high contrast image is the image in which there are all black or all white, and very little gray regions. The intensification operator on a fuzzy set $\widetilde{A}$ generates the other fuzzy set $\widetilde{A}' = INT\,(\widetilde{A})$. Moreover, in this fuzzy set $\widetilde{A}'$, the fuzziness is reduced by increasing the values of $\mu_{\widetilde{A}}(x)$ which are greater than 0.5 and by decreasing the values of $\mu_{\widetilde{A}}(x)$ which are less than 0.5. If we define this intensification as T, we could also define T for the membership values of the brightness for an image as follows

$$T(\mu_{ij}) = T'(\mu_{ij}) = 2\,\mu_{ij}^2. \quad 0 \leq \mu_{ij} \leq 0.5. \quad (8)$$
$$= T''(\mu_{ij}) = 1 - 2\,(1 - \mu_{ij})^2 \quad 0.5 < \mu_{ij} \leq 1.$$

T represents the operator INT as specified above and the transformation $\widetilde{A}$ to $\widetilde{A}'$ is defined by the recursive relation below

$T_r(\mu_{ij}) = T\,\{\,T_{r-1}(\mu_{ij})\,\}. \quad r = 1, 2 \,.....$

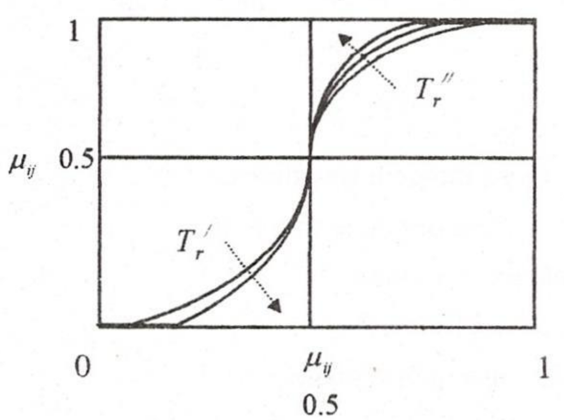


Figure 1: Graph of membership value through T transformations .

In practice, an image from a scanner usually has some noises. Hence, we have to smooth the image before using the *T* transformation. A useful smoothing algorithm is called defocusing.

The $(m,n)^{th}$ smoothed pixel intensity is found from *[Pal and King]*

$$\mu_{mn}' = a_0 \mu_{mn} + a_1 \sum_{Q_1} \mu_{ij} + a_2 \sum_{Q_2} \mu_{ij} + \ldots + a_S \sum_{Q_S} \mu_{ij} . \quad (9)$$

where

$a_0 + N_1 a_1 + N_2 a_2 + \ldots\ldots + N_S a_S = 1.$

$1 > a_1 > a_2 > \ldots > a_S > 0$ *and* $(i,j) \neq (m,n).$

To simplify this, we use the foundation

$$\mu_{mn} = \frac{1}{4}(\mu_{-10} + \mu_{10} + \mu_{01} + \mu_{0-1}) . \quad (10)$$

where μ_{mn} is evaluated from its 4–connected pixels

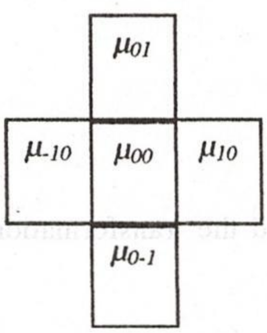


Figure 2 : 4–connected pixels of the center pixel.

Execution

- Apply the intensification operator to pixels of an image a few times.
- Repeat the following steps until the image is clear or can not be better

Step 1: Apply the *Eq(10)* to all pixels of the image for *n* times.

Step 2: Apply the *T* transformation to the image for *k* times.

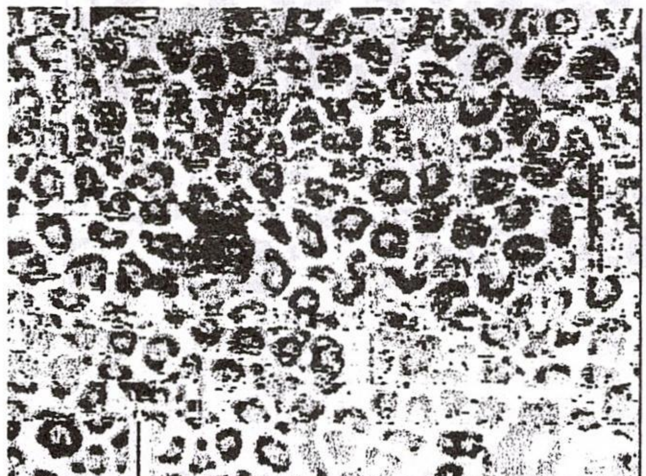
Note: Values of *k* and *n* belong to the degree of noise in the image .

Experimentation

We apply this execution to a bio-cell image whose photographs are taken through a microscope and scanned in low quality (we performed it to scatter noise in the image). The value of *n* is equal to 2 and *k* is also defined as 1.

The following part is the result of our test:

- The image before using the execution



- The image after using the execution.

After running for four times:

The last result:

MỘT PHƯƠNG PHÁP BIỂU DIỄN NGÔN NGỮ TỰ NHIÊN DƯỚI DẠNG TẬP MỜ VÀ ỨNG DỤNG TRONG BÀI TOÁN PHÂN LỚP ẢNH ĐỂ KHỬ NHIỄU.

Lê Hoài Bắc – Trần Thái Sơn

TÓM TẮT : Nội dung bài báo gồm hai phần : Phần đầu giới thiệu phương pháp biểu diễn ngôn ngữ tự nhiên dưới dạng tập mờ. Phần thứ hai, trình bày cách áp dụng phép biến đổi phân cực mờ trong bài toán phân lớp ảnh để khử nhiễu và một số kết quả bước đầu trong việc phân lớp ảnh các tế bào được chụp từ kính hiển vi

TÀI LIỆU THAM KHẢO

[1] Timothy J.Ross , " Fuzzy Logic with Engineering Applications" , McGraw-Hill ,New York ,(1995) .

[2] Toshiro Terano, Kiyoji Asai and Michio Sugeno , "Fuzzy Systems Theory and Its Applications" , Academics Press , Boston , (1992) .

[3] Dubois D. and Prade H, "Fuzzy sets and systems, theory and application" Academic NY, (1980).

[4] Zimmermann H, "Fuzzy set theory and its applications" 2nd ed, Kluwer Acad Publ, (1991).

[5] Ronald. Yager, Lofti A. Zadeh, An introduction to Fuzzy logic applications in intelligent systems", Kluwer Academic Publishers, 1992.