International Journal of Fuzzy Logic and Intelligent Systems, vol. 7, no. 3, September 2007 pp. 205-208

Operations on the Similarity Measures of Fuzzy Sets

Saleh Omran and M. Hassaballah

Mathematics Department, Faculty of Science, South Valley University, Qena, Egypt,

Abstract

Measuring the similarity between fuzzy sets plays a vital role in several fields. However, none of all well-known similarity measure methods is all-powerful, and all have the localization of its usage. This paper defines some operations on the similarity measures of fuzzy sets such as summation and multiplication of two similarity measures. Also, these operations will be generalized to any number of similarity measures. These operations will be very useful especially in the field of computer vision, and data retrieval because these fields need to combine and find some relations between similarity measures.

Key Words *: Fuzzy sets, Similarity measures, Operations on similarity measures*

1. Introduction

Fuzzy techniques can be applied in many domains of computer vision community. The definition of an adequate similarity measure for measuring the similarity between fuzzy sets is of great importance in the field of image processing, image retrieval and pattern recognition, etc. Objective measures or measures of comparison are required to test the performance of applying algorithms to an image, to compare the output image. Visual tasks are often based on the evaluation of similarities between image-objects represented in an appropriate feature space. The performance of content-based query systems depends on the definition of a suitable similarity measure [1,2].

Since Zadeh's work [3], a lot of attentions have been paid for the development of new similarity measures between fuzzy sets and their applications [4-9]. There is no generic method for selecting a suitable similarity measure or a distance measure. However, a prior information and statistics of features can be used in selection or to establish a new measure. Dietrich *et al.* [10] gave an overview of similarity measures, originally introduced to express the degree of comparison between fuzzy sets, which can be applied to images. In [11] the authors proposed similarity measures based on neighbourhoods, so that the relevant structures of the images are observed better. In this way 13 similarity measures were found to be appropriate for the comparison of images.

Similarity measures of another type between intuitionistic fuzzy sets (IFSs) were proposed by Wang *et al.* [12], some distance measures and the corresponding proofs are given, and the relations between similarity measure and distance measure of IFSs are analyzed. Measuring the degree of similarity between three fuzzy sets under unifying form and between IFSs is presented in [13]. The authors reviewed some existing similarity measures, showed that these measures are not always effective in some cases and illustrated the problem in the context of colorectal cancer diagnosis by similarity measure between fuzzy rough sets.

Unfortunately, all these similarity measures are not always satisfactory results and have the localization of its usage. In the mean time many applications such as information retrieval systems [14] require many similarity measures to be applied together in comparison of data objects, which in turn requires some relations and operations among these similarity measures. Until now there is not any paper covering this issue. To cope with this drawback, the present work defines some of these operations and relations among similarity measures. These operation may be very useful in the fields which require a combination of similarity measures such as content-based image retrieval, data classification, and database searching [15].

The rest of this paper is organized as follows. In Section 2, basic notions and definition of similarity measure are reviewed. Sections 3 describes the suggested operations on the similarity measures and investigates their proofs. Conclusions are given in section 4.

2. Preliminaries

2.1 Fuzzy sets

The theory of fuzzy sets F(X) was proposed by Zadeh [3]. A fuzzy set A in a universe X= $\{x_1, x_2, ... x_n\}$ is characterized by a mapping $\chi_A : X \rightarrow [0,1]$, which associates with every element x in X a degree of membership $\chi_A(x)$ of x in the fuzzy set A. In the following, let a= $\{a_1, a_2, ... a_n\}$ and b= $\{b_1, b_2, ... b_n\}$ be the vector representation of the fuzzy sets A and B respectively, where a_i and b_i are membership values

Manuscript received Oct. 4, 2006; revised Feb. 28, 2007.

$\chi_A(x_i)$ and $\chi_B(x_j)$ with respect to x_i and x_j ($i, j = 0,1,2...n$) respectively. Furthermore, suppose F(X) be the class of all fuzzy sets of X, $A^c \in F(X)$ is the complement of $A \in F(X)$.

In order to model the intersection or union between two fuzzy sets the $\wedge$ and $\vee$ operators will be used to refer to the minimum and maximum respectively. The cardinality of a finite crisp set is given by the number of elements in that set. This concept can be extend to fuzzy sets: the sigma count of a fuzzy set A (with finite support) in a universe X is defined as

$$|A| = \sum_{x_i \in X} \chi_A(x_i)$$

2.2 Similarity measures

There is no unique definition for the similarity measure, but the most common used definition is the following [5,12].

Definition 2.1 A similarity measure is a function assigning a similarity value to the pair of fuzzy sets (A,B) that indicates the degree to which A and B are equal or how similar they are. This function must be reflexive, symmetric and min-transitive. On other word, A mapping $S: F(X) x F(X) \rightarrow [0,1]$ is said to be a similarity measure between fuzzy sets $A \in F(X)$ and $B \in F(X)$, if $S(A,B)$ satisfies the following properties:

(SP1) $S(A,B) = S(B,A)$, $A,B \in F(X)$;

(SP2) $S(D,D^c) = 0$, if D is a crisp set;

(SP3) $S(E,E) = \max_{A,B \in F(X)} S(A,B)$, for all $E \in F(X)$;

(SP4) If $A \subseteq B \subseteq C$ for all $A,B,C \in F(X)$ then $S(A,B) \geq S(A,C)$ and $S(B,C) \geq S(A,C)$.

Based on this definition several similarity measures have been proposed [4,10]. But, with all this number of existed similarity measures, there is not any relation between them. In other words, there are not defined operations that control the behavior of these measures when they are applied concurrently. The next section suggests some of these operations and investigates their proof with respect to the above definition.

3. Operations on the similarity measures

Using the definition of similarity measure given in the previous section, two operations, the summation and multiplication of two similarity measures will be defined. Further these operations will be defined for any number of the similarity measures. According to the definition 2.1, the resulting of these operations is also a similarity measure in all cases and we will prove that.

Theorem 3.1 Let S_1 and S_2 be two similarity measures between two fuzzy sets A,B $\in$ F(X). The summation of S_1 and S_2 can be defined as follows

$$S_1 \oplus S_2 = \frac{S_1(A,B) + S_2(A,B)}{2}$$

Which is also a similarity measure. It is easy to see that all properties from SP1 to SP4 are satisfied.

Theorem 3.2 Let S_1 and S_2 be two similarity measures between two fuzzy sets A,B $\in$ F(X). The multiplication of S_1 and S_2 can be defined as follows

$$S_1 \otimes S_2 = \max \{ S_1(A,B), S_2(A,B) \}$$

Which is also a similarity measure. It is easy to see that all properties from SP1 to SP4 are satisfied.

Theorem 3.3 Let S_i be a set of similarity measures between two fuzzy sets A,B $\in$ F(X) and ($i = 1,2,3,...,n$). The summation operation on these similarity measures S_i can be defined as follows

$$\overset{n}{\underset{i=1}{\oplus}} S_i = \frac{1}{n}(\sum_{i=1}^{n} S_i(A,B))$$

Which is also a similarity measure.

Proof.

SP1 Since, $S_i(A,B)$ are similarity measures for each i, *i.e.*, $S_i(A,B) = S_i(B,A)$ for each i, therefore,

$$\frac{1}{n}(\sum_{i=1}^{n} S_i(A,B)) = \frac{1}{n}(\sum_{i=1}^{n} S_i(B,A)), \text{ for each } i.$$

SP2 $\overset{n}{\underset{i=1}{\oplus}} S_i(D,D^c) = 0$, if D is a crisp set;

The proof is obvious.

SP3 $\overset{n}{\underset{i=1}{\oplus}} S_i(E,E) = \frac{1}{n}(\sum_{i=1}^{n} S_i(E,E)) = \frac{1}{n}.n = 1$

SP4 If $A \subseteq B \subseteq C$ for all $A,B,C \in F(X)$ then $S_i(A,B) \geq S_i(A,C)$ and $S_i(B,C) \geq S_i(A,C)$ for each ***i***.

$\overset{n}{\underset{i=1}{\oplus}} S_i(A,B) \geq \overset{n}{\underset{i=1}{\oplus}} S_i(A,C)$ and similarly $\overset{n}{\underset{i=1}{\oplus}} S_i(B,C) \geq \overset{n}{\underset{i=1}{\oplus}} S_i(A,C)$.

i.e. If $A \subseteq B \subseteq C$ for all $A,B,C \in F(X)$ then,

$\overset{n}{\underset{i=1}{\oplus}} S_i(A,B) \geq \overset{n}{\underset{i=1}{\oplus}} S_i(A,C)$, and

$\overset{n}{\underset{i=1}{\oplus}} S_i(B,C) \geq \overset{n}{\underset{i=1}{\oplus}} S_i(A,C)$.

Moreover, $0 \leq S_i(A,B) \leq 1$, for each ***i***, and hence,

$$0 \leq \sum_{i=1}^{n} S_i(A,B) \leq n,$$

By dividing the inequality by n one can get

$$0 \leq \overset{n}{\underset{i=1}{\oplus}} S_i(A,B) \leq 1.$$

Theorem 3.4 Let S_i be a set of similarity measures between two fuzzy sets A,B $\in$ F(X) and ($i = 1,2,3,...,n$). The multiplication operation on these similarity measures S_i can be defined as follows

$$\overset{n}{\underset{i=1}{\otimes}} S_i = \max\{S_1(A,B), S_2(A,B), \ldots, S_n(A,B)\}$$

Based on definition 2.1, the multiplication operation is also a similarity measure. Using the same way, one can easily investigate the properties SP1-SP4.

Theorem 3.5 Let $\{A_i\}$ and $\{B_i\}$ be two collections of finite fuzzy sets with ($i = 1,2,3,\ldots,n$) such that $A = \bigcup_{i=1}^{n} A_i$ and $B = \bigcup_{i=1}^{n} B_i$. Then $S(A,B) = S(\bigcup_{i=1}^{n} A_i, \bigcup_{i=1}^{n} B_i) = \bigvee_{i=1}^{n} S(A_i,B_i)$ is a similarity measure.

Proof.

SP1 $S(A,B) = \bigvee_{i=1}^{n} S(A_i,B_i) = \bigvee_{i=1}^{n} S(B_i,A_i) = S(B,A)$

SP2 Since, $S(A_i,B_i)$ is a similarity measure for each *i*,

i.e. $S(A_i, A_i^{C}) = 0$, for each i, if A_i is crisp set.

So, $S(A, A^c) = S(A, 1-A) = \bigvee_{i=1}^{n} S(A_i, 1-A_i)$

$= \bigvee_{i=1}^{n} S(A_i, A^c{}_i) = \bigvee_{i=1}^{n} 0 = 0$.

Moreover, $S(A,A) = \bigvee_{i=1}^{n} S(A_i,A_i) = \bigvee_{i=1}^{n} 1 = 1$.

SP3 It is clear that $0 \le S(A,B) = \bigvee_{i=1}^{n} S(A_i,B_i) \le 1$, since $0 \le S(A_i,B_i) \le 1$.

SP4 If $A \subseteq B \subseteq C$ for all $A,B,C \in F(X)$, so

$A_i \subseteq B_i \subseteq C_i$ for each *i*, then

$S(A_i, B_i) \ge S(A_i,C_i)$ and $S(B_i,C_i) \ge S(A_i,C_i)$.

Therefore, $\bigvee_{i=1}^{n} S(A_i,B_i) \ge \bigvee_{i=1}^{n} S(A_i,C_i)$ and

$\bigvee_{i=1}^{n} S(B_i,C_i) \ge \bigvee_{i=1}^{n} S(A_i,C_i)$.

Theorem 3.6. Let $\{A_i\}$ and $\{B_i\}$ be two collections of finite fuzzy sets with ($i = 1,2,3,\ldots,n$) such that $A = \bigcap_{i=1}^{n} A_i$ and $B = \bigcap_{i=1}^{n} B_i$. Then $S(A,B) = S(\bigcap_{i=1}^{n} A_i, \bigcap_{i=1}^{n} B_i) = \bigwedge_{i=1}^{n} S(A_i,B_i)$ is a similarity measure.

Proof

The proof of this case is straightforward as in the proof of theorem 3.5, except using the meet $\wedge$ instead of the joint $\vee$.

Theorem 3.7 Let $\{A_i\}$ and $\{B_i\}$ be two collections of finite fuzzy sets with ($i = 1,2,3,\ldots,n$) such that $A = \bigcup_{i=1}^{n} A_i$ and $B = \bigcap_{i=1}^{n} B_i$, and $S(A_i,B_j)$ is similarity measure for each *i*, *j*.

Then, $S(A,B) = S(\bigcup_{i=1}^{n} A_i, \bigcap_{i=1}^{n} B_i) = \bigvee_{i=1}^{n} \bigwedge_{j=1}^{n} S(A_i,B_j)$

is a similarity measure.

Proof.

Investigating the properties of this case can be done using the theorem 3.5 and 3.6, and applying the proof for each *i*, and *j* respectively as follows.

SP1 Using theorem 3.5, we have

$S(A,B) = \bigvee_{i=1}^{n} \bigwedge_{j=1}^{n} S(A_i,B_j) = \bigwedge_{j=1}^{n} \bigvee_{i=1}^{n} S(A_i,B_j)$

$= \bigwedge_{j=1}^{n} \bigvee_{i=1}^{n} S(B_i,A_j) = S(B,A)$.

SP2 $S(A_i, 1-A_i) = 0$,

So, $\bigvee_{i=1}^{n} S(A_i, 1-A_i) = 0$, from the theorem 3.5,

Then $\bigwedge_{j=1}^{n} \bigvee_{i=1}^{n} S(A_i, A^c{}_j) = S(A, A^c) = 0$.

SP3 $S(A,A) = \bigvee_{i=1}^{n} \bigwedge_{j=1}^{n} S(A_i,A_j) = 1$, from the theorem 3.6,

we have, $\bigwedge_{j=1}^{n} S(A_i,A_j) = 1$, So $S(A,A) = \bigvee_{i=1}^{n} 1 = 1$.

SP4 If $A \subseteq B \subseteq C$ for all $A,B,C \in F(X)$, so

$A_i \subseteq B_i \subseteq C_i$ for each *i*,

Then, from the theorem 3.5, one can conclude that

$\bigvee_{i=1}^{n} S(A_i,B_i) \ge \bigvee_{i=1}^{n} S(A_i,C_i)$ and $\bigvee_{i=1}^{n} S(B_i,C_i) \ge \bigvee_{i=1}^{n} S(A_i,C_i)$.

Similarly, from theorem 3.6, one can conclude that

$\bigwedge_{i=1}^{n} S(A_i,B_i) \ge \bigwedge_{i=1}^{n} S(A_i,C_i)$ and $\bigwedge_{i=1}^{n} S(B_i,C_i) \ge \bigwedge_{i=1}^{n} S(A_i,C_i)$

Combining both conclusions leads to the following

$\bigvee_{i=1}^{n} \bigwedge_{j=1}^{n} S(A_i,B_j) \ge \bigvee_{i=1}^{n} \bigwedge_{j=1}^{n} S(A_i,C_j)$, and

$\bigvee_{i=1}^{n} \bigwedge_{j=1}^{n} S(B_i,C_j) \ge \bigvee_{i=1}^{n} \bigwedge_{j=1}^{n} S(A_i,C_j)$.

i.e. If $A \subseteq B \subseteq C$ for all $A,B,C \in F(X)$, then

$\bigvee_{i=1}^{n} \bigwedge_{j=1}^{n} S(A_i,B_j) \ge \bigvee_{i=1}^{n} \bigwedge_{j=1}^{n} S(A_i,C_j)$, and

$\bigvee_{i=1}^{n} \bigwedge_{j=1}^{n} S(B_i,C_j) \ge \bigvee_{i=1}^{n} \bigwedge_{j=1}^{n} S(A_i,C_j)$ for each *i*, and *j*.

4. Conclusion

Fuzzy techniques can be applied in several domains of computer community. Measuring the similarity between fuzzy sets is of great importance in the field of image processing, content-based image retrieval, data classification, and database searching. The drawbacks of the current similarity measures are that, none of all well-known similarity measure methods is all-powerful, and all have the localization of its usage. Also, several applications require that a group of similarity measures to be applied together while there is not any relation between them. To cope with this drawbacks, the present paper introduced some operations among similarity measures of fuzzy sets. The proposed operations will help in combining the results of many similarity measures together improving distinguish precision and enhancing the capability of classification of some similar sets.

References

[1] J. P. Eakins, “Towards intelligent image retrieval,” *Pattern Recognition*, vol. 35, pp. 3-14, 2002.

[2] T. Chaira and A. K. Ray, “Fuzzy measures for color image retrieval,” *Fuzzy Sets and Systems*, vol. 150, pp. 545-560, 2005.

[3] L. A. Zadeh, “Fuzzy sets,” *Inform. Control*, vol. 8, pp. 338-353, 1965.

[4] S. Omran and M. Hassaballah, “A new class of similarity measures for fuzzy sets,” *Journal of Fuzzy Logic and Intelligent Systems*, vol. 6, no. 2, pp. 100-104, Korea Fuzzy Logic and Intelligent Systems Society, 2006.

[5] S. H. Kwon, “A similarity measure of fuzzy sets,” *Journal of Fuzzy Logic and Intelligent Systems*, vol. 11, No.3, pp. 270-274, Korea Fuzzy Logic and Intelligent Systems Society, 2001.

[6] X. Liu, Entropy, “Distance measure and similarity measure of fuzzy sets and their relations,” *Fuzzy Sets and Systems*, vol. 52, pp. 305-318, 1992.

[7] X. Wang, B. De Baets, E. E. Kerre, “A comparative study of similarity measures,” *Fuzzy Sets and Systems*, vol. 73, pp. 259-268, 1995.

[8] S. M. Chen, “Measures of similarity between vague sets,” *Fuzzy Sets and Systems*, vol. 74, pp. 217-223, 1995.

[9] W. J. Wang, “New similarity measure on fuzzy sets and on elements,” *Fuzzy Sets and Systems*, vol. 85, pp. 305-309, 1997.

[10] D. Van der Weken, M. Nachtegael, E. E. Kerre, “An overview of similarity measures for images,” *Proceeding of ICASSP 2002 (IEEE International Conference on Acoustics, Speech and Signal processing)*, Orland, USA, pp.3317-3320, 2002.

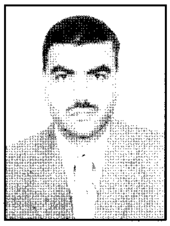
[11] D. Van der Weken, M. Nachtegael , E. E. Kerre, “Using similarity measures and homogeneity for the comparison of images,” *Image and Vision Computing*, vol. 22, pp.695-702, 2004.

[12] W. Wang, X. Xin, “Distance measure between intuitionistic fuzzy sets,” *Pattern Recognition Letters*, vol. 26, pp. 2063-2069, 2005.

[13] C. Zhang, H. Fu, “Similarity measures on three kinds of fuzzy sets,” *Pattern Recognition Letters*, vol. 27, no. 12, pp. 1307-1317, 2006.

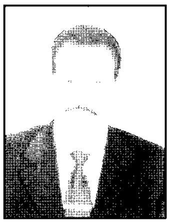
[14] H. Muller, N. Michoux, D. Bandon, A. Geissbuhler, “A review of content-based image retrieval systems in medical applications-clinical benefits and future directions,” *Medical Informatics*, vol. 73, pp. 1-23, 2004.

[15] A. Martinez and J. R. Serra, “Semantic access to database of images: an approach to object-related image retrieval,” *Proc. of IEEE Multimedia Systems (ICMCS)*, Florence, Italy, pp. 194-203, 1999.



Salah Omran received his Ph.D. degree in Mathematics in 2005 from University of Muenster, Germany. He is a lecturer in the Department of Mathematics at South Valley University, Egypt. His research interests are operator algebras, *k*-theory, and fuzzy sets and its applications.

Phone : +2 0102594601
Fax : +2 0965211279
E-mail : salahomran@yahoo.com



M. Hassaballah received his B.Sc. degree in Mathematics in 1997 and M.Sc. degree in Computer Science in 2003, all from South Valley University, Egypt. He is an assistant lecturer in the Department of Mathematics at South Valley University, Egypt. His research interests are in computer vision, fractal image compression, content-based image retrieval, similarity measures, and high performance computing.

Phone : +2 0965224219
Fax : +2 0965211279
E-mail : mahcs2010@yahoo.ca