# A Vector Similarity Measure for Interval Type-2 Fuzzy Sets

Dongrui Wu, *Student Member, IEEE*, and Jerry M. Mendel, *Life Fellow, IEEE*

***Abstract*— Fuzzy logic is frequently used in *computing with words* (CWW). When input words to a CWW engine are modeled by interval type-2 fuzzy sets (IT2 FSs), the CWW engine's output can also be an IT2 FS, $\tilde{A}$, which needs to be mapped to a linguistic label so that it can be understood. Because each linguistic label is represented by an IT2 FS $\tilde{B}_i$, there is a need to compare the similarity of $\tilde{A}$ and $\tilde{B}_i$ to find the $\tilde{B}_i$ most similar to $\tilde{A}$. In this paper, a vector similarity measure (VSM) is proposed for IT2 FSs, whose two elements measure the similarity in shape and proximity, respectively. A comparative study shows that the VSM gives more reasonable results than all other existing similarity measures for IT2 FSs.**

## I. Introduction

Zadeh coined the phrase "*computing with words*" (CWW) [24], [25]. According to [25], CWW is *"a methodology in which the objects of computation are words and propositions drawn from a natural language"*. CWW [19] *"is fundamentally different from the traditional expert systems which are simply tools to 'realize' an intelligent system, but are not able to process natural language which is imprecise, uncertain and partially true."*

Our thesis is that *words mean different things to different people* and so there is uncertainty associated with words, which means that fuzzy logic must somehow use this uncertainty when it computes with words [9], [10]. Hence, we argue that interval type-2 fuzzy sets (IT2 FSs) should be used in CWW [12]. We will limit our discussions to IT2 FSs in this paper.

A specific architecture is proposed in [11] for making judgements by CWW. A slightly modified architecture is shown in Fig. 1. It will be called a *perceptual computer*–Per-C for short. Perceptions (i.e., granulated terms, words) activate the Per-C and are also output by the Per-C; so, it is possible for a human to interact with the Per-C just using a vocabulary of words. In Fig. 1, the *encoder*[1] transforms linguistic perceptions into IT2 FSs that activate a *CWW engine*. The *decoder*[2] maps the output of the CWW engine into a word. Usually a vocabulary (codebook) is available, in which every word is modeled as an IT2 FS. The output of the CWW engine is mapped into a word (in that vocabulary) most similar to it.

The CWW engine, e.g. rules, the linguistic weighted average (LWA) [22], etc, maps IT2 FSs into IT2 FSs. If the CWW engine is rule-based, its output may be a crisp number (e.g., after defuzzification), in which case the decoder can map this number into a word in the vocabulary, as explained in [11]. On the other hand, if the CWW engine uses the LWA, its output is an IT2 FS $\tilde{A}$, or if the CWW engine is rule-based, but its output is also an IT2 FS $\tilde{A}$, then the decoder must also map $\tilde{A}$ into a word in the vocabulary. In this paper it is assumed that the output of the CWW engine is an IT2 FS $\tilde{A}$.

How to transform linguistic perceptions into IT2 FSs, i.e. the encoding problem, has been considered in [8], [14]–[16]. This paper considers the decoding problem, i.e. how to map an IT2 FS $\tilde{A}$ into a word (linguistic label). More specifically, given a vocabulary consisting of $N$ words with their associated IT2 FSs $\tilde{B}_i$ $(i = 1, \ldots, N)$, our goal is to find the $\tilde{B}_i$ which most closely resembles $\tilde{A}$, the output of the CWW engine. The word associated with that $\tilde{B}_i$ will then be viewed as the output of the Per-C. To do this, it must be possible to compare the similarity between two IT2 FSs. A vector similarity measure (VSM) for IT2 FSs is proposed in this paper. Since a type-1 FS can be viewed as a special case of an IT2 FS, the VSM can also be used for T1 FSs [23].

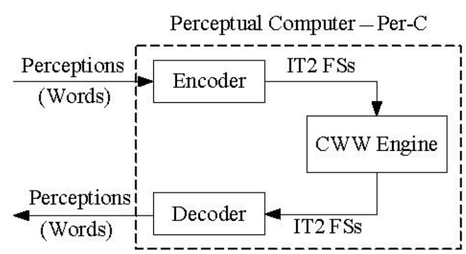


Fig. 1. Conceptual structure of CWW.

The rest of this paper is organized as follows: Section II reviews four existing similarity measures for IT2 FSs. Section III proposes the VSM and compares it with the four existing similarity measures. Section IV draws conclusions. Some background material about IT2 FSs is given in the Appendix.

## II. Existing Similarity Measures for IT2 FSs

Thought the literature on similarity measures for T1 FSs is quite extensive [3], to the best knowledge of the authors, only four similarity measures for IT2 FSs have appeared to date, and they are briefly reviewed next.

### A. Mitchell's IT2 FS Similarity Measure

Mitchell was the first to defines a similarity measure for *general* T2 FSs [18]. For the purpose of this paper, only its special case is explained, when both $\tilde{A}$ and $\tilde{B}$ are IT2 FSs:

Dongrui Wu and Jerry M. Mendel are with the Signal and Image Processing Institute, Ming Hsieh Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089-2564, USA (phone: 213-740-4445; fax: 213-740-4651; email: dongruiw@usc.edu, mendel@sipi.usc.edu).

[1] Zadeh calls this *constraint explicitation* in [24], [25]. In some of his recent talks, he calls this *precisiation.*

[2] Zadeh calls this *linguistic approximation* in [24], [25].

1-4244-1210-2/07/$25.00 

(1) Discretize the primary variable's universe of discourse, $X$, into $L$ points, that are used by both $\tilde{A}$ and $\tilde{B}$.

(2) Find $M$ embedded T1 MFs [see (27) in the Appendix] for IT2 FS $\tilde{A}$ ($m = 1, 2, \ldots, M$), i.e.

$$\mu_{A_e^m}(x_l) = r_m(x_l) \times [\overline{\mu}_{\tilde{A}}(x_l) - \underline{\mu}_{\tilde{A}}(x_l)] + \underline{\mu}_{\tilde{A}}(x_l) \quad (1)$$

where $r_m(x_l)$ is a random number chosen uniformly in $[0, 1]$, and $\underline{\mu}_{\tilde{A}}(x_l)$ and $\overline{\mu}_{\tilde{A}}(x_l)$ are the lower and upper memberships of $\tilde{A}$ at $x_l$, $l = 1, 2, \ldots, L$.

(3) Similarly, find $N$ embedded T1 MFs, $\mu_{B_e^n}$ ($n = 1, 2, \ldots, N$), for IT2 FS $\tilde{B}$, i.e.,

$$\mu_{B_e^n}(x_l) = r_n(x_l) \times [\overline{\mu}_{\tilde{B}}(x_l) - \underline{\mu}_{\tilde{B}}(x_l)] + \underline{\mu}_{\tilde{B}}(x_l) \quad (2)$$

(4) Compute the IT2 FS similarity measure $s_M(\tilde{A}, \tilde{B})$ as an average of T1 FS similarity measures $s_{mn}$ that are computed for all of the $MN$ combinations of the embedded T1 FSs for $\tilde{A}$ and $\tilde{B}$ [this uses the Representation Theorem in (29)], i.e.,

$$s_M(\tilde{A}, \tilde{B}) = \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} s_{mn}, \quad (3)$$

where

$$s_{mn} = s(A_e^m, A_e^n) \quad (4)$$

and $s_{mn}$ can be any T1 FS similarity measure.

Mitchell's IT2 FS similarity measure has the following problems:

(1) Generally $s_M(\tilde{A}, \tilde{B})$ does not equal 1 even for the special case where $\tilde{A}$ and $\tilde{B}$ are exactly the same, because the randomly generated embedded T1 FSs from $\tilde{A}$ and $\tilde{B}$ will not always be the same.

(2) Because there are random numbers involved, $s_M(\tilde{A}, \tilde{B})$ may change from experiment[3] to experiment. When both $M$ and $N$ are large, some kind of stochastic convergence can be expected to occur (e.g., convergence in probability); however, the computational cost is heavy because the computation of (3) requires direct enumeration of all $MN$ embedded T1 FSs.

### B. Gorzalczany's IT2 FS Compatibility Measure

Gorzalczany proposed a compatibility measure for interval-valued FSs (IVFSs) [5]. Because an IVFS is an IT2 FS under a different name, the terms and symbols used in [5] are changed so that they are consistent with those in this paper.

Gorzalczany defined the *degree of compatibility*, $s_G(\tilde{A}, \tilde{B})$, between two IT2 FSs $\tilde{A}$ and $\tilde{B}$ as

$$s_G(\tilde{A}, \tilde{B}) = \left[ \min \left( \frac{\max_{x \in X} \{\min(\underline{\mu}_{\tilde{A}}(x), \underline{\mu}_{\tilde{B}}(x))\}}{\max_{x \in X} \underline{\mu}_{\tilde{A}}(x)}, \frac{\max_{x \in X} \{\min(\overline{\mu}_{\tilde{A}}(x), \overline{\mu}_{\tilde{B}}(x))\}}{\max_{x \in X} \overline{\mu}_{\tilde{A}}(x)} \right), \right.$$
$$\left. \max \left( \frac{\max_{x \in X} \{\min(\underline{\mu}_{\tilde{A}}(x), \underline{\mu}_{\tilde{B}}(x))\}}{\max_{x \in X} \underline{\mu}_{\tilde{A}}(x)}, \frac{\max_{x \in X} \{\min(\overline{\mu}_{\tilde{A}}(x), \overline{\mu}_{\tilde{B}}(x))\}}{\max_{x \in X} \overline{\mu}_{\tilde{A}}(x)} \right) \right] \quad (5)$$

As pointed out by Tsiporkova and Zimmermann [20], compatibility measures do not *"perform consistently as similarity measures of FSs."* It is easy to show that Gorzalczany's compatibility measure may give counter-intuitive results when used in linguistic approximation. Consider the example shown in Fig. 2, where $\max_{x \in X} \underline{\mu}_{\tilde{A}}(x) = \max_{x \in X} \underline{\mu}_{\tilde{B}}(x) = \mu_1$. Consequently,

$$\max_{x \in X} \{\min(\underline{\mu}_{\tilde{A}}(x), \underline{\mu}_{\tilde{B}}(x))\} = \max_{x \in X} \underline{\mu}_{\tilde{A}}(x) = \mu_1 \quad (6)$$

and,

$$\frac{\max_{x \in X} \{\min(\underline{\mu}_{\tilde{A}}(x), \underline{\mu}_{\tilde{B}}(x))\}}{\max_{x \in X} \underline{\mu}_{\tilde{A}}(x)} = \frac{\mu_1}{\mu_1} = 1. \quad (7)$$

It is also easy to see that

$$\frac{\max_{x \in X} \{\min(\overline{\mu}_{\tilde{A}}(x), \overline{\mu}_{\tilde{B}}(x))\}}{\max_{x \in X} \overline{\mu}_{\tilde{A}}(x)} = \frac{1}{1} = 1. \quad (8)$$

Hence, for $\tilde{A}$ and $\tilde{B}$ shown in Fig. 2, $s_G(\tilde{A}, \tilde{B}) = [1, 1]$. Actually it can be shown that as long as $\max_{x \in X} \underline{\mu}_{\tilde{A}}(x) = \max_{x \in X} \underline{\mu}_{\tilde{B}}(x)$ and $\max_{x \in X} \overline{\mu}_{\tilde{A}}(x) = \max_{x \in X} \overline{\mu}_{\tilde{B}}(x)$, no matter how different the shapes of $\tilde{A}$ and $\tilde{B}$ are, Gorzalczany's compatibility measure always gives $s_G(\tilde{A}, \tilde{B}) = s_G(\tilde{B}, \tilde{A}) = [1, 1]$, which is counter-intuitive.

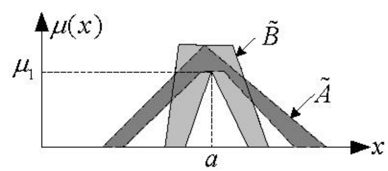


Fig. 2. Example for Gorzalczany's compatibility measure, which gives $s_G(\tilde{A}, \tilde{B}) = [1, 1]$.

### C. Bustince's IT2 FS Similarity Measure

Bustince also proposed a similarity measure for IVFSs [2]. Again, the terms and symbols used in [2] are changed so that they are consistent with those in this paper.

First, Bustince defined a *normal interval valued similarity measure* $s_B(\tilde{A}, \tilde{B})$ between two IT2 FSs $\tilde{A}$ and $\tilde{B}$, as one that satisfies five properties given in [2]. He then proposed

$$s_B(\tilde{A}, \tilde{B}) = \left[ s_L(\tilde{A}, \tilde{B}), s_U(\tilde{A}, \tilde{B}) \right] \quad (9)$$

as an *interval-valued normal similarity measure*, where

$$s_L(\tilde{A}, \tilde{B}) = \Upsilon_L(\tilde{A}, \tilde{B}) \star \Upsilon_L(\tilde{B}, \tilde{A}) \quad (10)$$

and

$$s_U(\tilde{A}, \tilde{B}) = \Upsilon_U(\tilde{A}, \tilde{B}) \star \Upsilon_U(\tilde{B}, \tilde{A}), \quad (11)$$

[3] One experiment is comprised of $M$ ($N$) randomly chosen embedded T1 FSs for $\tilde{A}$ ($\tilde{B}$).

$\star$ can be any $t$-norm (e.g. minimum), and $[\Upsilon_L(\tilde{A},\tilde{B}), \Upsilon_U(\tilde{A},\tilde{B})]$ is an *interval valued inclusion grade indicator* [2] of $\tilde{A}$ in $\tilde{B}$. $\Upsilon_L(\tilde{A},\tilde{B})$ and $\Upsilon_U(\tilde{A},\tilde{B})$ used in this paper (and taken from [2]) are computed as

$$\Upsilon_L(\tilde{A},\tilde{B}) = \inf_{x\in X}\{1, \min(1-\underline{\mu}_{\tilde{A}}(x)+\underline{\mu}_{\tilde{B}}(x), \\ 1-\overline{\mu}_{\tilde{A}}(x)+\overline{\mu}_{\tilde{B}}(x))\} \quad (12)$$

$$\Upsilon_U(\tilde{A},\tilde{B}) = \inf_{u\in U}\{1, \max(1-\underline{\mu}_{\tilde{A}}(x)+\underline{\mu}_{\tilde{B}}(x), \\ 1-\overline{\mu}_{\tilde{A}}(x)+\overline{\mu}_{\tilde{B}}(x))\} \quad (13)$$

A problem with Bustince's similarity measure is that when $\tilde{A}$ and $\tilde{B}$ are disjoint, no matter how far away they are from each other, $s_B(\tilde{A},\tilde{B})$ will always be a non-zero constant.

### D. Zeng and Li's IT2 FS Similarity Measure

Zeng and Li's similarity measure was also proposed for IVFSs [26]. Again, the terms and symbols used in [26] are changed so that they are consistent with those in this paper.

Zeng and Li defined a real function $s_Z\colon \tilde{A}\times\tilde{B}\to[0,1]$ as a similarity measure of IT2 FSs, if $s_Z$ satisfies four properties given in [26]. They then proposed the following similarity measure for IT2 FSs if the universes of discourse of $\tilde{A}$ and $\tilde{B}$ are discrete:

$$s_Z(\tilde{A},\tilde{B}) = 1-\frac{1}{2n}\sum_{i=1}^{n}[|\underline{\mu}_{\tilde{A}}(x_i)-\underline{\mu}_{\tilde{B}}(x_i)|+ \\ |\overline{\mu}_{\tilde{A}}(x_i)-\overline{\mu}_{\tilde{B}}(x_i)|], \quad (14)$$

and, if the universes of discourse of $\tilde{A}$ and $\tilde{B}$ are continuous in $[a,b]$,

$$s_Z(\tilde{A},\tilde{B}) = 1-\frac{1}{2(b-a)}\int_a^b[|\underline{\mu}_{\tilde{A}}(x)-\underline{\mu}_{\tilde{B}}(x)|+ \\ |\overline{\mu}_{\tilde{A}}(x)-\overline{\mu}_{\tilde{B}}(x)|]dx. \quad (15)$$

Zeng and Li's similarity measure has a problem similar to that of Bustince's, but may be worse depending on the choice of $a$ and $b$ [see (15)]. For example in Fig. 3, $\tilde{B}$ and $\tilde{B}'$ have the same shape but are at different distances from $\tilde{A}$; hence, $\int_a^b\left(|\underline{\mu}_{\tilde{A}}(x)-\underline{\mu}_{\tilde{B}}(x)|+|\overline{\mu}_{\tilde{A}}(x)-\overline{\mu}_{\tilde{B}}(x)|\right)dx$, $\int_a^{b'}\left(|\underline{\mu}_{\tilde{A}}(x)-\underline{\mu}_{\tilde{B}}(x)|+|\overline{\mu}_{\tilde{A}}(x)-\overline{\mu}_{\tilde{B}}(x)|\right)dx$ and $\int_a^{b'}\left(|\underline{\mu}_{\tilde{A}}(x)-\underline{\mu}_{\tilde{B}'}(x)|+|\overline{\mu}_{\tilde{A}}(x)-\overline{\mu}_{\tilde{B}'}(x)|\right)dx$ are equal, and this value is denoted as $c$. There can be two methods in computing $s_Z(\tilde{A},\tilde{B})$ and $s_Z(\tilde{A},\tilde{B}')$:

(1) If the interval $[a,b]$ is used to compute $s_Z(\tilde{A},\tilde{B})$ and $[a,b']$ is used to compute $s_Z(\tilde{A},\tilde{B}')$, then $s_Z(\tilde{A},\tilde{B}) = 1-c/[2(b-a)]$ and $s_Z(\tilde{A},\tilde{B}') = 1-c/[2(b'-a)]$. Because $b'-a > b-a$, $s_Z(\tilde{A},\tilde{B}) < s_Z(\tilde{A},\tilde{B}')$, which means $\tilde{B}'$ is more similar to $\tilde{A}$ than $\tilde{B}$ is. Additionally, as $b'-a$ increases, $s_Z(\tilde{A},\tilde{B}')$ approaches 1.

(2) If the interval $[a,b']$ is used to compute both $s_Z(\tilde{A},\tilde{B})$ and $s_Z(\tilde{A},\tilde{B}')$, then $s_Z(\tilde{A},\tilde{B}) = 1-c/[2(b'-a)]$ and $s_Z(\tilde{A},\tilde{B}') = 1-c/[2(b'-a)]$; hence, $s_Z(\tilde{A},\tilde{B}) = s_Z(\tilde{A},\tilde{B}')$.

Both results are counter-intuitive, because (Fig. 3) $\tilde{B}$ should be more similar to $\tilde{A}$ than should $\tilde{B}'$.

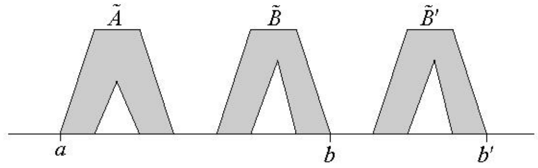


Fig. 3. Example of Zeng and Li's similarity measure for disjoint IT2 FSs.

### E. Summary

In summary, each of the four existing similarity measures for IT2 FSs has its problems:

- Mitchell's similarity measure involves randomness which can lead to different answers, and the computational cost is high.
- Gorzalczany's, Bustince's and Zeng and Li's similarity measures give counter-intuitive results for some special cases.

### F. Proposed Properties for a Similarity Measure

To avoid such problems, the following four properties are proposed for a similarity measure for IT2 FSs.

P.1) The similarity between two IT2 FSs is 1 if and only if they are exactly the same.

P.2) If two IT2 FSs overlap, there should be some similarity between them.

P.3) If two IT2 FSs become more distant from each other, similarity between them should decrease.

P.4) The similarity between two IT2 FSs should be a constant regardless of the order they are compared, i.e. $s(\tilde{A},\tilde{B}) = s(\tilde{B},\tilde{A})$.

In the next section a vector similarity measure is proposed which possesses these properties.

## III. A Vector Similarity Measure for IT2 FSs

When the similarity of two IT2 FSs $\tilde{A}$ and $\tilde{B}$ are compared, it is necessary to compare their shapes as well as proximity. In this section, a vector similarity measure (VSM) for IT2 FSs, $\mathbf{s}_v(\tilde{A},\tilde{B})$, is proposed, one that has two components, i.e.,

$$\mathbf{s}_v(\tilde{A},\tilde{B}) = \left(s_1(\tilde{A},\tilde{B}), s_2(\tilde{A},\tilde{B})\right)^T \quad (16)$$

where $s_1(\tilde{A},\tilde{B})\in[0,1]$ is a similarity measure on the shapes of $\tilde{A}$ and $\tilde{B}$, and $s_2(\tilde{A},\tilde{B})\in[0,1]$ is a similarity measure on the proximity of $\tilde{A}$ and $\tilde{B}$.

### A. Definition of $s_1(\tilde{A},\tilde{B})$

Because the proximity of $\tilde{A}$ and $\tilde{B}$ is considered in $s_2(\tilde{A},\tilde{B})$, in computing $s_1(\tilde{A},\tilde{B})$ $\tilde{A}$ and $\tilde{B}$ are "aligned" so that their shapes can be compared. Denote the centroids (see Section B in the Appendix) of $\tilde{A}$ and $\tilde{B}$ as $C_{\tilde{A}} = [c_l(\tilde{A}), c_r(\tilde{A})]$ and $C_{\tilde{B}} = [c_l(\tilde{B}), c_r(\tilde{B})]$, respectively, and the centers of $C_{\tilde{A}}$ and $C_{\tilde{B}}$ as

$$c(\tilde{A}) = [c_l(\tilde{A})+c_r(\tilde{A})]/2 \quad (17)$$

$$c(\tilde{B}) = [c_l(\tilde{B})+c_r(\tilde{B})]/2 \quad (18)$$

A reasonable alignment method is to move one or both of $\tilde{A}$ and $\tilde{B}$ so that $c(\tilde{A})$ and $c(\tilde{B})$ coincide [see Fig. 4]. The two IT2 FSs can be moved to any location as long as $c(\tilde{A})$ and $c(\tilde{B})$ coincide; this will not affect the value of $s_1(\tilde{A}, \tilde{B})$. In this paper $\tilde{B}$ is moved to $\tilde{A}$ and called $\tilde{B}'$, as shown in Fig. 4(b).

When $\tilde{A}$ and $\tilde{B}$ are aligned, $s_1(\tilde{A}, \tilde{B})$ is defined as a crisp number equal to the ratio of the *average cardinalities* [see (33)] of $\tilde{A} \cap \tilde{B}'$ and $\tilde{A} \cup \tilde{B}'$, i.e.

$$\begin{aligned} s_1(\tilde{A}, \tilde{B}) &\equiv \frac{AC(\tilde{A} \cap \tilde{B}')}{AC(\tilde{A} \cup \tilde{B}')} \\ &= \frac{card(\overline{\mu}_{\tilde{A}}(x) \cap \overline{\mu}_{\tilde{B}'}(x)) + card(\underline{\mu}_{\tilde{A}}(x) \cap \underline{\mu}_{\tilde{B}'}(x))}{card(\overline{\mu}_{\tilde{A}}(x) \cup \overline{\mu}_{\tilde{B}'}(x)) + card(\underline{\mu}_{\tilde{A}}(x) \cup \underline{\mu}_{\tilde{B}'}(x))} \\ &= \frac{\int_X \min(\overline{\mu}_{\tilde{A}}(x), \overline{\mu}_{\tilde{B}'}(x))dx + \int_X \min(\underline{\mu}_{\tilde{A}}(x), \underline{\mu}_{\tilde{B}'}(x))dx}{\int_X \max(\overline{\mu}_{\tilde{A}}(x), \overline{\mu}_{\tilde{B}'}(x))dx + \int_X \max(\underline{\mu}_{\tilde{A}}(x), \underline{\mu}_{\tilde{B}'}(x))dx}, \end{aligned} \tag{19}$$

where $\overline{\mu}_{\tilde{B}'}(x)$ and $\underline{\mu}_{\tilde{B}'}(x)$ are illustrated in Fig. 4(b). When all uncertainty disappears, $\tilde{A}$ and $\tilde{B}$ become T1 FSs $A$ and $B$, and (19) reduces to Jaccard's similarity measure [6].

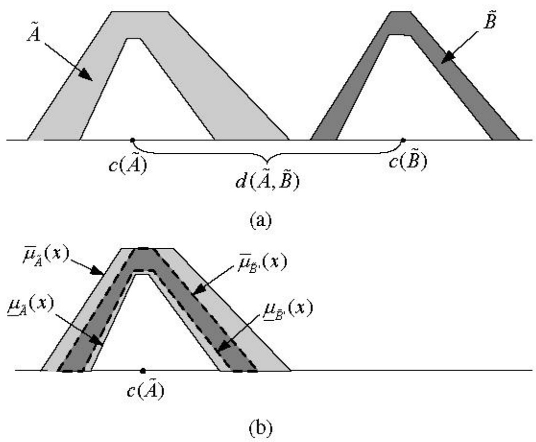


Fig. 4. An example for the proposed VSM. (a) $c(\tilde{A})$ and $c(\tilde{B})$ denote the center of the centroids of $\tilde{A}$ and $\tilde{B}$, respectively; (b) $\tilde{B}'$ is obtained by moving $\tilde{B}$ so that $c(\tilde{B})$ coincides with $c(\tilde{A})$. The solid curves are for $\tilde{A}$ and the dashed curves are for $\tilde{B}'$.

Properties of $s_1(\tilde{A}, \tilde{B})$ include: (a) $0 \leq s_1(\tilde{A}, \tilde{B}) \leq 1$; (b) $s_1(\tilde{A}, \tilde{B}) = 1 \Leftrightarrow \tilde{A} = \tilde{B}'$, i.e. $\tilde{A}$ and $\tilde{B}$ have the same shape; and, (c) $s_1(\tilde{A}, \tilde{B}) = s_1(\tilde{B}, \tilde{A})$. The proofs can be found in [21].

### B. Definition of $s_2(\tilde{A}, \tilde{B})$

$s_2(\tilde{A}, \tilde{B})$ measures the proximity of $\tilde{A}$ and $\tilde{B}$, and is defined as

$$s_2(\tilde{A}, \tilde{B}) \equiv h(d(\tilde{A}, \tilde{B})) \tag{20}$$

where

$$d(\tilde{A}, \tilde{B}) = |c(\tilde{A}) - c(\tilde{B})| \tag{21}$$

is the Euclidean distance between the centers of the centroids of $\tilde{A}$ and $\tilde{B}$ [see Fig. 4(a)], and $h$ can be any function satisfying: (1) $\lim_{x \to \infty} h(x) = 0$; (2) $h(x) = 1$ if and only if $x = 0$; and, (3) $h(x)$ decreases monotonically as $x$ increases.

Observe that $s_2(\tilde{A}, \tilde{B}) \in (0, 1]$, and $s_2(\tilde{A}, \tilde{B}) = 1$ if and only if $c(\tilde{A}) = c(\tilde{B})$.

An example of $s_2(\tilde{A}, \tilde{B})$ is

$$s_2(\tilde{A}, \tilde{B}) = e^{-rd(\tilde{A}, \tilde{B})}, \tag{22}$$

where $r$ is a positive constant. $s_2(\tilde{A}, \tilde{B})$ is chosen as an exponential function because we believe the similarity between two FSs should decrease rapidly as the distance between them increases.

### C. On Converting $\mathrm{s}_v(\tilde{A}, \tilde{B})$ to the scalar similarity measure $s_s(\tilde{A}, \tilde{B})$

$\mathrm{s}_v(\tilde{A}, \tilde{B})$ enables us to separately quantify the similarity of two features, shape and proximity. As mentioned in the Introduction, in CWW $\mathrm{s}_v(\tilde{A}, \tilde{B}_i)$ $(i = 1, 2, \ldots, N)$ need to be ranked to find the $\tilde{B}_i$ most similar to $A$. This can be achieved by first converting the vector $\mathrm{s}_v(\tilde{A}, \tilde{B}_i)$ to a scalar similarity measure $s_s(\tilde{A}, \tilde{B}_i)$ and then ranking $s_s(\tilde{A}, \tilde{B}_i)$.

In this paper, the scalar similarity between two IT2 FSs $\tilde{A}$ and $\tilde{B}$ is computed as the product of their similarities in shape and proximity [4], i.e.

$$s_s(\tilde{A}, \tilde{B}) = s_1(\tilde{A}, \tilde{B}) \times s_2(\tilde{A}, \tilde{B}) \tag{23}$$

Properties of $s_s(\tilde{A}, \tilde{B})$ include: (a) $\tilde{A} = \tilde{B} \Leftrightarrow s_s(\tilde{A}, \tilde{B}) = 1$; (b) $s_s(\tilde{A}, \tilde{B}) > 0$; (c) $s_s(\tilde{A}, \tilde{B}) > s_s(\tilde{A}, \tilde{C})$ if $\tilde{B}$ and $\tilde{C}$ have the same shape and $\tilde{C}$ is further away from $\tilde{A}$ than $\tilde{B}$ is; and, (d) $s_s(\tilde{A}, \tilde{B}) = s_s(\tilde{B}, \tilde{A})$. The proofs can be found in [21].

Observe that $s_s(\tilde{A}, \tilde{B})$ satisfies the four properties stated in Section II-F.

### D. Examples

Comparisons of all the similarity measures for IT2 FSs introduced in this paper are given in Table I for IT2 FSs $\tilde{A} - \tilde{G}$ depicted in Fig. 5. Note that $\tilde{A} - \tilde{E}$ have the same shape. The domain of $x$ [e.g. the support of $\tilde{A} \cup \tilde{B}$ in computing $s(\tilde{A}, \tilde{B})$] was discretized into 500 equal-length intervals, $M \equiv N = 10$ and $s_{mn} = card(A_e^m \cap B_e^n)/card(A_e^m \cup B_e^n)$ (Jaccard's similarity measure [6]) in Mitchell's similarity measure, Method (1) in Section II-D was used to choose $x_i$ in Zeng and Li's similarity measure, and $r \equiv 4/|X|$ ($|X|$ is the length of the support of $\tilde{A} \cup \tilde{B}$) in the VSM [see (22)].

Observe from Table I that the outputs of the VSM are reasonable for all six cases, according to the four properties proposed in Section II-F. Observe also that:

(1) Using Mitchell's method, $s(\tilde{F}, \tilde{F}) = 0.6007$, which should be 1.
(2) Using Gorzalczany's method, $s(\tilde{F}, \tilde{G}) = [1, 1]$, which should be less than 1.
(3) Using Bustince's method, $s(\tilde{A}, \tilde{D}) = s(\tilde{A}, \tilde{E}) \neq [0, 0]$, which should be $s(\tilde{A}, \tilde{D}) > s(\tilde{A}, \tilde{E})$, or at least, $s(\tilde{A}, \tilde{D}) = s(\tilde{A}, \tilde{E}) = [0, 0]$.

[4]Recently, Bonissone, et al. [1] defined a similarity measure as a weighted minimum of several sub-similarity measures. Although similar to our idea, their objective is quite different from our objective; hence, their similarity measure is not used in this paper.

TABLE I

COMPARISON OF SIMILARITY MEASURES FOR $\tilde{A} - \tilde{G}$ SHOWN IN FIG. 5.

| Similarity Measure | $s(\tilde{A},\tilde{B})$ | $s(\tilde{A},\tilde{C})$ | $s(\tilde{A},\tilde{D})$ | $s(\tilde{A},\tilde{E})$ | $s(\tilde{F},\tilde{F})$ | $s(\tilde{F},\tilde{G})$ |
|---|---|---|---|---|---|---|
| Mitchell's Method ($s_M$) | $0.1494$ | $0.0124$ | $0$ | $0$ | $0.6007$ | $0.5762$ |
| Gorzalczany's Method ($s_G$) | $[0, 0.5980]$ | $[0, 0.1967]$ | $[0, 0]$ | $[0, 0]$ | $[1, 1]$ | $[1, 1]$ |
| Bustince's Method ($s_B$) | $[0.0017, 0.2016]$ | $[0.0010, 0.2016]$ | $[0.0007, 0.2016]$ | $[0.0007, 0.2016]$ | $[1, 1]$ | $[0.3337, 1]$ |
| Zeng and Li's Method ($s_Z$) | $0.6578$ | $0.6452$ | $0.7006$ | $0.7467$ | $1$ | $0.7782$ |
| VSM ($s_s$) | $0.2019$ | $0.0408$ | $0.0082$ | $0.0017$ | $1$ | $0.5175$ |

(4) Using Zeng and Li's method, $s(\tilde{A},\tilde{B}) < s(\tilde{A},\tilde{D}) < s(\tilde{A},\tilde{E})$, which should be $s(\tilde{A},\tilde{B}) > s(\tilde{A},\tilde{D}) > s(\tilde{A},\tilde{E})$, and, $s(\tilde{A},\tilde{C}) < s(\tilde{A},\tilde{D}) < s(\tilde{A},\tilde{E})$, which should be $s(\tilde{A},\tilde{C}) > s(\tilde{A},\tilde{D}) > s(\tilde{A},\tilde{E})$.

Finally, observe that the VSM does not have any of the short-comings of these four similarity measures.

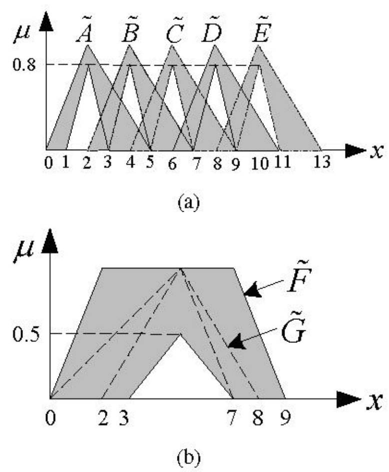


Fig. 5. Examples used in the comparative study: (a) $\tilde{A} - \tilde{E}$, which have the same shape; (b) $\tilde{F}$ (solid lines) and $\tilde{G}$ (dashed lines).

## IV. CONCLUSIONS

In this paper, four existing similarity measures for IT2 FSs are reviewed, their short-comings are pointed out, and a vector similarity measure for IT2 FSs is proposed. The VSM is the first IT2 FS similarity measure that has a vector form. It is easy to understand, and its two components enable us to consider the similarity between shapes and proximity separately and explicitly. A comparative study showed that the VSM gives reasonable similarity measures and does not have the short-comings of the four existing similarity measures. Research has also shown that the VSM can be used as a similarity measure for type-1 FSs [23].

## APPENDIX

### A. Interval Type-2 Fuzzy Sets (IT2 FS)

An IT2 FS, $\tilde{A}$, is to-date the most widely used kind of T2 FS, and is the only kind of T2 FS that is considered in this paper. It is described as [5]

$$\tilde{A} = \int_{x\in X}\int_{u\in J_x} 1/(x,u) = \int_{x\in X}\left[\int_{u\in J_x} 1/u\right]\bigg/ x \quad (24)$$

[5]This background material is taken from [17]. See also [9].

where $x$ is the *primary variable*, $J_x \subseteq [0,1]$ is the *primary membership* of $x$, $u$ is the *secondary variable*, and $\int_{u\in J_x} 1/u$ is the *secondary membership function* (MF) at $x$. Uncertainty about $\tilde{A}$ is conveyed by the union of all of the primary memberships, called the *footprint of uncertainty* of $\tilde{A}$ [$FOU(\tilde{A})$], i.e.

$$\begin{aligned} FOU(\tilde{A}) &= \bigcup_{x\in X} J_x \\ &= \left\{(x,u) : u \in J_x = \left[\underline{\mu}_{\tilde{A}}(x), \overline{\mu}_{\tilde{A}}(x)\right] \subseteq [0,1]\right\} \end{aligned} \quad (25)$$

An IT2 FS is shown in Fig. 6. The FOU is shown as the shaded region. It is bounded by an *upper MF* (UMF) $\overline{\mu}_{\tilde{A}}(x)$ and a *lower MF* (LMF) $\underline{\mu}_{\tilde{A}}(x)$, both of which are T1 FSs; consequently, the membership grade of each element of an IT2 FS is an interval $[\underline{\mu}_{\tilde{A}}(x), \overline{\mu}_{\tilde{A}}(x)]$.

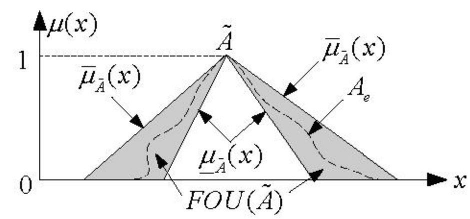


Fig. 6. An interval type-2 fuzzy set. $A_e$ is an embedded type-1 fuzzy set.

Note that an IT2 FS can also be represented as

$$\tilde{A} = 1/FOU(\tilde{A}) \quad (26)$$

with the understanding that this means putting a secondary grade of 1 at all points of $FOU(\tilde{A})$.

For discrete universes of discourse $X$ and $U$, an *embedded T1 FS* $A_e$ has $N$ elements, one each from $J_{x_1}, J_{x_2}, \ldots, J_{x_N}$, namely $u_1, u_2, \ldots, u_N$, i.e.

$$A_e = \sum_{i=1}^{N} u_i/x_i \qquad u_i \in J_{x_i} \subseteq U = [0,1] \quad (27)$$

Examples of $A_e$ are $\overline{\mu}_{\tilde{A}}(x)$ and $\underline{\mu}_{\tilde{A}}(x)$; see, also Fig. 6. Note that if each $u_i$ is discretized into $M_i$ levels, there will be a total of $n_A$ $A_e$, where

$$n_A = \prod_{i=1}^{N} M_i \quad (28)$$

Mendel and John [13] have presented a Representation Theorem for a general T2 FS, which when specialized to an IT2 FS can be expressed as:

**Representation Theorem for an IT2 FS:** Assume that primary variable $x$ of an IT2 FS $\tilde{A}$ is sampled at $N$ values, $x_1, x_2, \dots, x_N$, and at each of these values its primary memberships $u_i$ are sampled at $M_i$ values, $u_{i1}, u_{i2}, \dots, u_{iM_i}$. Let $A_e^j$ denote the $j$th embedded T1 FS for $\tilde{A}$. Then $\tilde{A}$ is represented by (26), in which[6]

$$FOU(\tilde{A}) = \bigcup_{j=1}^{n_A} A_e^j = \{\underline{\mu}_{\tilde{A}}(x), \dots, \overline{\mu}_{\tilde{A}}(x)\}$$
$$\equiv [\underline{\mu}_{\tilde{A}}(x), \overline{\mu}_{\tilde{A}}(x)]. \quad (29)$$

This representation of an IT2 FS, in terms of simple T1 FSs, the embedded T1 FSs, is very useful for deriving theoretical results; however, it is not recommended for computational purposes, because it would require the enumeration of the $n_A$ embedded T1 FSs and $n_A$ [given in (28)] can be astronomical.

### B. Centroid of an IT2 FS

The centroid of an IT2 FS has been well-defined by Karnik and Mendel [7]. Let $A_e$ be an embedded T1 FS of an IT2 FS $\tilde{A}$. The centroid of $\tilde{A}$ is defined as the union of the centroids of all $A_e$, i.e.,

$$C_{\tilde{A}} \equiv \bigcup_{\forall A_e} c(A_e) = \bigcup_{\forall A_e} \frac{\int_X x \cdot \mu_{A_e}(x)dx}{\int_X \mu_{A_e}(x)dx} = [c_l(\tilde{A}), c_r(\tilde{A})] \quad (30)$$

where $c(A_e)$ is the centroid of $A_e$, and $c_l(\tilde{A})$ and $c_r(\tilde{A})$ are the minimum and maximum centroids of all $A_e$, respectively. $c_l(\tilde{A})$ and $c_r(\tilde{A})$ can be computed by using the Karnik-Mendel (KM) algorithms [9].

### C. Cardinality of an IT2 FS

Definitions of the cardinality of T1 FSs have been proposed by many authors. De Luca and Termini's definition [4] is used in this paper:

$$card(A) = \int_X \mu_A(x)dx. \quad (31)$$

The cardinality of an IT2 FS $\tilde{A}$ is then defined as the union of all cardinalities of its embedded T1 FSs $A_e$, i.e.,

$$card(\tilde{A}) \equiv \bigcup_{\forall A_e} card(A_e) = \bigcup_{\forall A_e} \left[\int_X \mu_{A_e}(x)dx\right]$$
$$= [card(\underline{\mu}_{\tilde{A}}(x)), card(\overline{\mu}_{\tilde{A}}(x))]. \quad (32)$$

Additionally, the *average cardinality* (AC) of $\tilde{A}$ is defined as the average of its minimum and maximum cardinalities, i.e.,

$$AC(\tilde{A}) = \frac{card(\underline{\mu}_{\tilde{A}}(x)) + card(\overline{\mu}_{\tilde{A}}(x))}{2}. \quad (33)$$

Note that $AC(\tilde{A})$ is a crisp number whereas $card(\tilde{A})$ is an interval. The *average cardinality* instead of *cardinality* is used in the VSM so that $s_1(\tilde{A}, \tilde{B})$ is a crisp number.

[6] Although there are a finite number of embedded T1 FSs, it is customary to represent $FOU(\tilde{A})$ as an interval set $[\underline{\mu}_{\tilde{A}}(x), \overline{\mu}_{\tilde{A}}(x)]$ at each $x$. Doing this is equivalent to discretizing with infinitesimally many small values and letting the discretizations approach zero.

## REFERENCES


[1] P. P. Bonissone, A. Varma, K. S. Aggour, and F. Xue, "Design of local fuzzy models using evolutionary algorithms," to be published in *Journal of Computational Statistics and Decision Analysis*, 2006.

[2] H. Bustince, "Indicator of inclusion grade for interval-valued fuzzy sets. application to approximate reasoning based on interval-valued fuzzy sets," *Int'l Journal of Approximate Reasoning*, vol. 23, no. 3, pp. 137–209, March 2000.

[3] V. V. Cross and T. A. Sudkamp, *Similarity and Compatibility in Fuzzy Set Theory: Assessment and Applications*. Heidelberg, NY: Physica-Verlag, 2002.

[4] A. De Luca and S. Termini, "A definition of nonprobabilistic entropy in the setting of fuzzy sets theory," *Information and Computation*, vol. 20, pp. 301–312, 1972.

[5] M. B. Gorzalczany, "A method of inference in approximate reasoning based on interval-valued fuzzy sets," *Fuzzy Sets and Systems*, vol. 21, pp. 1–17, 1987.

[6] P. Jaccard, "Nouvelles recherches sur la distribution florale," *Bulletin de la Societe de Vaud des Sciences Naturelles*, vol. 44, p. 223, 1908.

[7] N. N. Karnik and J. M. Mendel, "Centroid of a type-2 fuzzy set," *Information Sciences*, vol. 132, pp. 195–220, 2001.

[8] F. Liu and J. M. Mendel, "An interval approach to fuzzistics for interval type-2 fuzzy sets," submitted to IEEE Trans. on Fuzzy Systems, Feb 2007.

[9] J. M. Mendel, *Rule-Based Fuzzy Logic Systems: Introduction and New Directions*. Upper Saddle River, NJ: Prentice-Hall, 2001.

[10] ——, "Computing with words, when words can mean different things to different people," in *Proc. 3rd Int'l ICSC Symposium on Fuzzy Logic and Applications*, Rochester, NY, June 1999, pp. 158–164.

[11] ——, "An architecture for making judgement using computing with words," *Int'l Journal of Applied Mathematics and Computer Science*, vol. 12, no. 3, pp. 325–335, 2002.

[12] ——, "Computing with words and its relationships with fuzzistics," *Information Sciences*, vol. 177, pp. 988–1006, 2007.

[13] J. M. Mendel and R. I. John, "Type-2 fuzzy sets made simple," *IEEE Trans. on Fuzzy Systems*, vol. 10, no. 2, pp. 117–127, April 2002.

[14] J. M. Mendel and H. Wu, "Type-2 fuzzistics for symmetric interval type-2 fuzzy sets: Part 1, forward problems," *IEEE Trans. on Fuzzy Systems*, vol. 14, no. 6, pp. 781–792, 2006.

[15] ——, "Type-2 fuzzistics for symmetric interval type-2 fuzzy sets: Part 2, inverse problems," *IEEE Trans. on Fuzzy Systems*, 2006, in press.

[16] ——, "Centroid uncertainty bounds for interval type-2 fuzzy sets: Forward and inverse problems," in *Proc. FUZZ-IEEE*, vol. 2, Budapest, Hungary, July 2004, pp. 947–952.

[17] J. M. Mendel, H. Hagras, and R. I. John, "Standard background material about interval type-2 fuzzy logic systems that can be used by all authors," http://ieee-cis.org/_files/standards.t2.win.pdf.

[18] H. B. Mitchell, "Pattern recognition using type-II fuzzy sets," *Information Sciences*, vol. 170, pp. 409–418, 2005.

[19] M. Nikravesh, "Soft computing for reservoir characterization and management," in *Proc. IEEE Int'l Conf. Granular Computing*, vol. 2, Beijing, China, July 2005, pp. 593–598.

[20] E. Tsiporkova and H.-J. Zimmermann, "Aggregation of compatibility and equality: A new class of similarity measures for fuzzy sets," in *Proc. 7th Int'l Conf. on Information Processing and Management of Uncertainty in Knowledge-Based Systems*, Paris, 1998, pp. 1769–1776.

[21] D. Wu and J. M. Mendel, "A vector similarity measure for linguistic approximation: Interval type-2 and type-1 fuzzy sets," submitted to *Information Sciences*, 2006.

[22] ——, "Aggregation using the linguistic weighted average and interval type-2 fuzzy sets," accepted by *IEEE Trans. on Fuzzy Systems*, 2007.

[23] ——, "A vector similarity measure for type-1 fuzzy sets," in *Proc. IFSA World Congress*, Cancun, Maxico, June 2007.

[24] L. A. Zadeh, "Fuzzy logic = computing with words," *IEEE Trans. on Fuzzy Systems*, vol. 4, pp. 103–111, 1996.

[25] ——, "From computing with numbers to computing with words – from manipulation of measurements to manipulation of perceptions," *IEEE Trans. on Circuits and Systems-I: Fundamental Theory and Applications*, vol. 4, pp. 105–119, 1999.

[26] W. Zeng and H. Li, "Relationship between similarity measure and entropy of interval valued fuzzy sets," *Fuzzy Sets and Systems*, vol. 157, pp. 1477–1484, 2006.