International Journal of Uncertainty,
Fuzziness and Knowledge-Based Systems
Vol. 12, No. 6 (2004) 827–841

World Scientific
www.worldscientific.com

SIMILARITY MEASURES BETWEEN TYPE-2 FUZZY SETS

WEN-LIANG HUNG

Department of Mathematics Education, National Hsinchu Teachers College,
Hsin-Chu, Taiwan, R.O.C.
wlhung@mail.nhctc.edu.tw

MIIN-SHEN YANG

Department of Applied Mathematics, Chung Yuan Christian University,
Chung-Li, Taiwan, R.O.C.
msyang@math.cycu.edu.tw

Received 1 November 2002
Revised 14 may 2004

In this paper, we give similarity measures between type-2 fuzzy sets and provide the axiom definition and properties of these measures. For practical use, we show how to compute the similarities between Gaussian type-2 fuzzy sets. Yang and Shih's [22] algorithm, a clustering method based on fuzzy relations by beginning with a similarity matrix, is applied to these Gaussian type-2 fuzzy sets by beginning with these similarities. The clustering results are reasonable consisting of a hierarchical tree according to different levels.

Keywords: Gaussian type-2 fuzzy sets; Hausdorff distance; similarity measure; type-2 fuzzy sets.

1 Introduction

Zadeh [24] introduced the type-2 fuzzy set concept as an extension of an ordinary fuzzy set, i.e., a type-1 fuzzy set. A type-2 fuzzy set is a fuzzy set whose membership values are type-1 fuzzy sets on $[0, 1]$. Operations on type-2 fuzzy sets were studied by [5-9] and [13-15]. All of these operations are needed to implement type-2 fuzzy logic system [10] and [18]. Recently Mendel and John [11] presented a new Representation Theorem for type-2 fuzzy sets and showed that it can be used to derive formulas for the union, intersection, and complement of type-2 fuzzy sets without having to use the Extension Principle.

The similarity concept is extremely important, for it provides the degree of similarity between two fuzzy concepts. Since Zadeh [23] presented the similarity relation concept, similarity measures between fuzzy sets have been widely studied and applied in various areas. Turksen and Zhong [19] applied similarity measures between fuzzy sets for an approximate analogical reasoning. For rule matching in fuzzy control and neural networks, a similarity measure between the fuzzy sets is always used to determine whether a rule should be fired (cf. [2], [19]). In a multimedia database query, Candan *et al.* [3] applied similarity measures to develop query processing with different fuzzy semantics.

However, the similarity between type-2 fuzzy sets has received little attention. In this paper, we present similarity measures between type-2 fuzzy sets and provide the axiom definition and properties of these measures. For practical use, we show how to compute the similarities between Gaussian type-2 fuzzy sets. Furthermore, Yang and Shih's [22] algorithm, a clustering method based on fuzzy relations by beginning with a similarity matrix, is applied to these Gaussian type-2 fuzzy sets by beginning with these similarities. The clustering results are reasonable consisting of a hierarchical tree according to different levels. The remainder of this paper is organized as follows. In Section 2 type-2 fuzzy sets and the distance between fuzzy sets are reviewed. In Section 3 we extend Liu [12] axioms to type-2 fuzzy sets and create two similarity measures between type-2 fuzzy sets. Section 4 presents some examples. Conclusions are stated in Section 5.

2 Preliminaries

The following definitions and preliminaries are required in the sequel and are presented here in brief.

2.1 Type-2 fuzzy sets

Mendel and John [11] recently presented a new presentation for type-2 fuzzy sets so that it makes us easily use and effectively communicate about type-2 fuzzy sets. Using the definitions there, we introduce type-2 fuzzy sets as follows:

Definition 1 (Mendel and John [11]). A type-2 fuzzy set, denoted $\widetilde{A}$, is characterized by a type-2 membership function $\mu_{\widetilde{A}}(x,u)$, where $x \in X$ and $u \in J_x \subseteq [0,1]$, i.e.,

$$\widetilde{A} = \{((x,u), \mu_{\widetilde{A}}(x,u)) | \forall x \in X, \forall u \in J_x \subseteq [0,1]\} \tag{1}$$

in which $0 \leq \mu_{\widetilde{A}}(x,u) \leq 1$. $\widetilde{A}$ can be also expressed as

$$\widetilde{A} = \int_{x\in X} \int_{u \in J_x} \mu_{\widetilde{A}}(x,u)/(x,u), \quad J_x \subseteq [0,1] \tag{2}$$

where $\int \int$ denotes union over all admissible x and u. For discrete universes of discourse $\int$ is replaced by Σ.

Definition 2 (Mendel and John [11]). At each value of x (say $x = x'$), the 2-D plane whose axes are u and $\mu_{\widetilde{A}}(x',u)$ is called a vertical slice of $\mu_{\widetilde{A}}(x,u)$. A secondary membership function is a vertical slice of $\mu_{\widetilde{A}}(x,u)$. It is $\mu_{\widetilde{A}}(x = x',u)$ for $x \in X$ and $\forall u \in J_{x'} \subseteq [0,1]$, i.e.

$$\mu_{\widetilde{A}}(x = x', u) \equiv \mu_{\widetilde{A}}(x') = \int_{u \in J_{x'}} f_{x'}(u)/u, \quad J_{x'} \subseteq [0,1] \tag{3}$$

in which $0 \le f_{x'}(u) \le 1$. Because $\forall x' \in X$, we drop the prime notation on $\mu_{\tilde{A}}(x')$, and refer to $\mu_{\tilde{A}}(x)$ as a secondary membership function; it is a type-1 fuzzy set, which we also refer to as a secondary set.

According to the concept of secondary sets, a type-2 fuzzy set can be re-interpreted as the union of all secondary sets, i.e., $\tilde{A}$, using (3), can be re-expressed in a vertical-slice manner, as

$$\tilde{A} = \{(x, \mu_{\tilde{A}}(x)) | \forall x \in X\} \tag{4}$$

or, as

$$\tilde{A} = \int_{x \in X} \mu_{\tilde{A}}(x)/x = \int_{x \in X} \left[\int_{u \in J_x} f_x(u)/u \right] / x, \quad J_x \subseteq [0,1]. \tag{5}$$

Definition 3. (Mendel and John [11]). The domain of a secondary membership function is called the primary membership of x. In (5), J_x is the primary membership of x, where $J_x \subseteq [0,1]$ for $\forall x \in X$.

Definition 4. (Mendel and John [11]). The amplitude of a secondary membership function is called a secondary grade. In (5), $f_x(u)$ is a secondary grade; in (1), $\mu_{\tilde{A}}(x', u')$ $(x' \in X, u' \in J_{x'})$ is a secondary grade.

Assume that each of the secondary membership functions of a type-2 fuzzy set has only one secondary grade that equals 1. A principal membership function is the union of all such points at which this occurs, i.e., $\mu_{\text{principal}}(x) = \int_{x \in X} u/x$, where $f_x(u) = 1$, and is associated with a type-1 fuzzy set.

Definition 5. (Mendel and John [11]). A type-1 fuzzy set can also be expressed as a type-2 fuzzy set. Its type-2 representation is $(1/\mu_F(x))/x$ or $1/\mu_F(x)$, $\forall x \in X$, for short. The notation $1/\mu_F(x)$ means that the secondary membership function has only one value in its domain, namely the primary membership $\mu_F(x)$, at which the secondary grade equals 1.

When the secondary membership functions (MFs) of a type-2 fuzzy set are type-1 Gaussian MFs, we call the type-2 fuzzy set a Gaussian type-2 set. When the secondary MFs are type-1 interval sets, we call the type-2 fuzzy set an interval type-2 set.

2.2 Distances between fuzzy sets

The distance concept has been extended to subsets of a metric space. One popular set distance is the Hausdorff distance, which is a metric. Let W^λ denote the operation of dilating the set W by radius λ (i.e., W^λ is the set of all points within distance λ of W). For any two non-empty compact sets U, V, let

$$L(U,V) = \inf\{\lambda \in [0,\infty) | U^\lambda \supset V\}.$$

The Hausdorff distance between U and V is then defined by

$$H(U,V) = \max\{L(U,V), L(V,U)\}.$$

Consider two type-1 fuzzy sets A and B on a metric space X. The maximum membership of A is

$$\mu^* = \max\{\mu_A(x) | x \in X\}.$$

Define the non-fuzzy set

$$A_{\max} = \{x | \mu_A(x) = \mu^*\}.$$

Let A_a be a non-fuzzy subset of X such that

$$A_a \supset A_{\max}.$$

For any two type-1 fuzzy sets A and B, we assume that

$$A_a = B_a \text{ if and only if } A_{\max} = B_{\max}.$$

Define the family of non-fuzzy sets A_t, where $t \in [0, 1]$, using

$$A_t = \begin{cases} \{x | \mu_A(x) \in [t, \mu^*]\} & \text{if } t \le \mu^*, \\ A_a & \text{if } t > \mu^*. \end{cases}$$

Note that $A_t = A_{\max}$ if $t = \mu^*$, and that the second case does not arise if $\mu^* = 1$. This non-fuzzy set A_t is usually called a level-t (or t-cut) set of A.

Suppose that the type-1 fuzzy sets can take on only a discrete set of membership values $t_1, t_2, \ldots, t_n$. Let $H(A_{t_i}, B_{t_i})$ be the crisp Hausdorff distance between two level-t_i sets A_{t_i} and B_{t_i}. Then Chaudhuri and Rosenfeld [4] defined

$$H_f(A, B) = \frac{\sum_{i=1}^n t_i H(A_{t_i}, B_{t_i})}{\sum_{i=1}^n t_i} \tag{6}$$

as the fuzzy Hausdorff distance between two type-1 fuzzy sets A and B. If A and B are continuous-valued, then (6) has the integral expression

$$H_f(A, B) = \frac{\int_0^1 tH(A_t, B_t)dt}{\int_0^1 tdt} = 2\int_0^1 tH(A_t, B_t)dt. \tag{7}$$

Thus, H_f is a generalization of the 'crisp' Hausdorff distance which is a metric defined for type-1 fuzzy sets.

Another commonly used distance between two type-1 fuzzy sets A and B is the Hamming distance d_H [21]. Let $X = \{x_1, \cdots, x_n\}$ be a discrete finite set, then d_H is defined as

$$d_H(A, B) = \sum_{i=1}^n |\mu_A(x_i) - \mu_B(x_i)|. \tag{8}$$

When X is continuous, say $X = [a, b]$, the Hamming distance between A and B is defined as

$$d_H(A, B) = \int_a^b |\mu_A(x) - \mu_B(x)|dx. \tag{9}$$

Comparing the fuzzy Hausdorff distance $H_f(A, B)$ and the Hamming distance $d_H(A, B)$ between two type-1 fuzzy sets A and B, they use the different viewpoint from the different metric space to define their distances. The fuzzy Hausdorff distance $H_f(A, B)$ uses the weighted mean based on the differences between two level-t_i sets A_{t_i} and B_{t_i} on the metric space X. However, the Hamming distance $d_H(A, B)$ uses the sum based on the differences between two membership values $\mu_A(x_i)$ and $\mu_B(x_i)$ on the interval [0,1]. In the next section, we will combine these two different viewpoints of the distance definition to define our new distance between two type-2 fuzzy sets and then define a new similarity measure.

3 A proposed similarity measure for type-2 fuzzy sets

Distance measure is a term that describes the difference between fuzzy sets. Numerous researchers have used it but no axiom definition for it has been seen. Liu [12] first axiomatized the distance measure of fuzzy sets. The Liu axiom formulations are intuitive and have been widely employed in the fuzzy literature. They were formulated in the following way. Throughout this section, $R^+ = [0, \infty]$; X is the universal set; $\mathcal{F}_1(X)$ is the class of all type-1 fuzzy sets of X; $\mathcal{F}_2(X)$ is the class of all type-2 fuzzy sets of X; $\mathcal{P}(X)$ is the class of all crisp sets of X; A^c is the complement of $A \in \mathcal{F}_1(X)$. Let d be a real function $d : \mathcal{F}_1(X) \times \mathcal{F}_1(X) \to R^+$. d is a distance measure on $\mathcal{F}_1(X)$ if it satisfies the four Liu axioms:

(D1) $d(A, B) = d(B, A)$, $\forall A, B \in \mathcal{F}_1(X)$;

(D2) $d(A, A) = 0$, $\forall A \in \mathcal{F}_1(X)$;

(D3) $d(D, D^c) = \max_{A,B \in \mathcal{F}_1(X)} d(A, B)$, $\forall D \in \mathcal{P}(X)$;

(D4) $\forall A, B, C \in \mathcal{F}_1(X)$, if $A \subset B \subset C$, then $d(A, B) \leq d(A, C)$ and $d(B, C) \leq d(A, C)$.

Related to the distance measure concept, the similarity measure of two fuzzy sets is a measure that indicates the similarity between fuzzy sets. It is easy to see that the distance measure and similarity measure are two dual concepts. Liu [12] also gave an axiom definition for the similarity measure of fuzzy sets. Let s be a real function $s : \mathcal{F}_1(X) \times \mathcal{F}_1(X) \to R^+$. s is a similarity measure on $\mathcal{F}_1(X)$ if it satisfies the four Liu axioms:

(S1) $s(A, B) = s(B, A)$, $\forall A, B \in \mathcal{F}_1(X)$;

(S2) $s(D, D^c) = 0$, $\forall D \in \mathcal{P}(X)$;

(S3) $s(C, C) = \max_{A,B \in \mathcal{F}_1(X)} s(A, B)$, $\forall C \in \mathcal{F}_1(X)$;

(S4) $\forall A, B, C \in \mathcal{F}_1(X)$, if $A \subset B \subset C$, then $s(A, B) \geq s(A, C)$ and $s(B, C) \geq s(A, C)$.

If we normalize d and s, we can ensure $0 \leq d(A, B) \leq 1$, $0 \leq s(A, B) \leq 1$, for $A, B \in \mathcal{F}_1(X)$. The relation between d and s is $s = 1 - d$. Since the Liu axioms (D1)-(D4) and (S1)-(S4) were formulated for type-1 fuzzy sets, they are expressed for the type-2 fuzzy sets as follows:

(D1') $d(\widetilde{A}, \widetilde{B}) = d(\widetilde{B}, \widetilde{A})$, $\forall \widetilde{A}, \widetilde{B} \in \mathcal{F}_2(X)$;

(D2') $d(\widetilde{A}, \widetilde{A}) = 0$, $\forall \widetilde{A} \in \mathcal{F}_2(X)$;

(D3') $d(D, D^c) = \max_{\widetilde{A}, \widetilde{B} \in \mathcal{F}_2(X)} d(\widetilde{A}, \widetilde{B})$, $\forall D \in \mathcal{P}(X)$;

(D4') $\forall \widetilde{A}, \widetilde{B}, \widetilde{C} \in \mathcal{F}_2(X)$, if $\widetilde{A} \subset \widetilde{B} \subset \widetilde{C}$, then $d(\widetilde{A}, \widetilde{B}) \leq d(\widetilde{A}, \widetilde{C})$ and $d(\widetilde{B}, \widetilde{C}) \leq d(\widetilde{A}, \widetilde{C})$.

(S1') $s(\widetilde{A}, \widetilde{B}) = s(\widetilde{B}, \widetilde{A})$, $\forall \widetilde{A}, \widetilde{B} \in \mathcal{F}_2(X)$;

(S2') $s(D, D^c) = 0$, $\forall D \in \mathcal{P}(X)$;

(S3') $s(\widetilde{C}, \widetilde{C}) = \max_{\widetilde{A},\widetilde{B}\in\mathcal{F}_2(X)} s(\widetilde{A}, \widetilde{B})$, $\forall \widetilde{C} \in \mathcal{F}_2(X)$;

(S4') $\forall \widetilde{A}, \widetilde{B}, \widetilde{C} \in \mathcal{F}_2(X)$, if $\widetilde{A} \subset \widetilde{B} \subset \widetilde{C}$, then $s(\widetilde{A}, \widetilde{B}) \geq s(\widetilde{A}, \widetilde{C})$ and $s(\widetilde{B}, \widetilde{C}) \leq s(\widetilde{A}, \widetilde{C})$.

Remark 1. Let $\widetilde{A}, \widetilde{B} \in \mathcal{F}_2(X)$,

$$\widetilde{A} \subset \widetilde{B} \text{ if and only if } \mu_{\widetilde{A}}(x) \leq \mu_{\widetilde{B}}(x), \quad \forall x \in X.$$

Let $X = \{x_1, x_2, \cdots x_n\}$ be a discrete finite set. Consider a type-2 fuzzy set $\widetilde{A}$ in X where its membership grade of $x \in X$ is $\mu_{\widetilde{A}}(x)$, which is a type-1 fuzzy set in $[0, 1]$. For any type-2 fuzzy sets $\widetilde{A}$ and $\widetilde{B}$ in X, we propose a new distance between $\widetilde{A}$ and $\widetilde{B}$ defined by

$$d(\widetilde{A}, \widetilde{B}) = \sum_{i=1}^{n} H_f(\mu_{\widetilde{A}}(x_i), \mu_{\widetilde{B}}(x_i)), \tag{10}$$

where $H_f(\mu_{\widetilde{A}}(x_i), \mu_{\widetilde{B}}(x_i))$ is the fuzzy Hausdorff distance between $\mu_{\widetilde{A}}(x_i)$ and $\mu_{\widetilde{B}}(x_i)$. Since the membership grades $\mu_{\widetilde{A}}(x_i)$ and $\mu_{\widetilde{B}}(x_i)$ are type-1 fuzzy sets in $[0, 1]$,

$$0 \leq H_f(\mu_{\widetilde{A}}(x_i), \mu_{\widetilde{B}}(x_i)) \leq 1, \quad \forall i.$$

Thus,

$$0 \leq d(\widetilde{A}, \widetilde{B}) \leq n.$$

It is worth to mention that, when type-2 fuzzy sets $\widetilde{A}$ and $\widetilde{B}$ in X become type-1 fuzzy sets, our proposed distance $d(\widetilde{A}, \widetilde{B})$ becomes the Hamming distance. We combine the Hamming and fuzzy Hausdorff distances to define the distance for two type-2 fuzzy sets $\widetilde{A}$ and $\widetilde{B}$ where the fuzzy Hausdorff distance is used to define for secondary membership functions and the Hamming distance is used to define for the primary membership functions. It is easy to see that the defined distance measure (10) satisfies the above properties (D1') and (D2'). But we have to prove (D3') and (D4') for (10).

Proof of (D3') for the distance measure (10).

(i) If $x_i \in D$ then

$$\mu_D(x_i, u) = \begin{cases} 1 & \text{if } u = 1\ ; \\ 0 & \text{otherwise,} \end{cases}$$

and

$$\mu_{D^c}(x_i, u) = \begin{cases} 1 & \text{if } u = 0\ ; \\ 0 & \text{otherwise.} \end{cases}$$

Using (6), we obtain

$$H_f(\mu_D(x_i), \mu_{D^c}(x_i)) = 1.$$

(ii) If $x_i \in D^c$ then

$$\mu_D(x_i, u) = \begin{cases} 1 & \text{if } u = 0\ ; \\ 0 & \text{otherwise,} \end{cases}$$

and

$$\mu_{D^c}(x_i, u) = \begin{cases} 1 & \text{if } u = 1\ ; \\ 0 & \text{otherwise.} \end{cases}$$

Using (6), we also obtain

$$H_f(\mu_D(x_i), \mu_{D^c}(x_i)) = 1.$$

Combining (i) and (ii), we have

$$d(D, D^c) = \sum_{i=1}^{n} H_f(\mu_D(x_i), \mu_{D^c}(x_i)) = n = \max_{\widetilde{A},\widetilde{B}\in\mathcal{F}_2(X)} d(\widetilde{A}, \widetilde{B}).$$

Proof of (D4') for the distance measure (10). Since $\widetilde{A} \subset \widetilde{B} \subset \widetilde{C}$ implies $\mu_{\widetilde{A}}(x_i) \le \mu_{\widetilde{B}}(x_i) \le \mu_{\widetilde{C}}(x_i)$, $\forall x_i \in X$, then

$$H_f(\mu_{\widetilde{A}}(x_i), \mu_{\widetilde{B}}(x_i)) \le H_f(\mu_{\widetilde{A}}(x_i), \mu_{\widetilde{C}}(x_i)),$$

$$H_f(\mu_{\widetilde{B}}(x_i), \mu_{\widetilde{C}}(x_i)) \le H_f(\mu_{\widetilde{A}}(x_i), \mu_{\widetilde{C}}(x_i)).$$

Therefore,

$$\sum_{i=1}^{n} H_f(\mu_{\widetilde{A}}(x_i), \mu_{\widetilde{B}}(x_i)) \le \sum_{i=1}^{n} H_f(\mu_{\widetilde{A}}(x_i), \mu_{\widetilde{C}}(x_i)),$$

$$\sum_{i=1}^{n} H_f(\mu_{\widetilde{B}}(x_i), \mu_{\widetilde{C}}(x_i)) \le \sum_{i=1}^{n} H_f(\mu_{\widetilde{A}}(x_i), \mu_{\widetilde{C}}(x_i)).$$

That is,

$$d(\widetilde{A}, \widetilde{B}) \le d(\widetilde{A}, \widetilde{C}) \text{ and } d(\widetilde{B}, \widetilde{C}) \le d(\widetilde{A}, \widetilde{C}).$$

Next, we define a similarity measure for two type-2 fuzzy sets $\widetilde{A}$ and $\widetilde{B}$. We first consider the normalized distance of $d(\widetilde{A}, \widetilde{B})$, i.e.,

$$d^N(\widetilde{A}, \widetilde{B}) = \frac{d(\widetilde{A}, \widetilde{B})}{n}. \tag{11}$$

Using the dual concept, we define the similarity measure of $\widetilde{A}$ and $\widetilde{B}$ as follows:

$$s(\widetilde{A}, \widetilde{B}) = 1 - d^N(\widetilde{A}, \widetilde{B}).$$

It is obviously that $s(\widetilde{A}, \widetilde{B})$ satisfies the above properties (S1')-(S4').

Remark 2. When uncertainties for the secondary membership function in a type-2 fuzzy set $\widetilde{A}$ disappear, it must reduce to a type-1 fuzzy set A. In such case, equation (10) will be equal to the Hamming distance between $A, B \in \mathcal{F}_1(X)$ as

$$d(\widetilde{A}, \widetilde{B}) = d_H(A, B) = \sum_{i=1}^{n} |\mu_A(x_i) - \mu_B(x_i)|.$$

Remark 3. When the universal set X is continuous, say $X = [a, b]$, we can obtain the following similar result. For any two type-2 fuzzy sets $\widetilde{A}$ and $\widetilde{B}$ in X, the distance between $\widetilde{A}$ and $\widetilde{B}$ is defined as

$$d(\widetilde{A}, \widetilde{B}) = \int_a^b H_f(\mu_{\widetilde{A}}(x), \mu_{\widetilde{B}}(x))dx. \tag{12}$$

It is well known that an exponential operation is highly useful in dealing with the classical Shannon entropy (cf. [16-17]) and cluster analysis (cf. [20]). We therefore adopted the exponential operation on $d^N(\widetilde{A}, \widetilde{B})$. We can regard

$$d_e(\widetilde{A}, \widetilde{B}) = \frac{1 - \exp(-d^N(\widetilde{A}, \widetilde{B}))}{1 - \exp(-1)},$$

as the normalized exponential-type distance measure. Hence we propose a similarity measure s_e between the type-2 fuzzy sets $\widetilde{A}$ and $\widetilde{B}$ as

$$s_e(\widetilde{A}, \widetilde{B}) = 1 - d_e(\widetilde{A}, \widetilde{B}).$$

Lemma 1. Let $f(t) = \frac{1-\exp(-t)}{1-\exp(-1)}$, then $\max_{t\in[0,1]} f(t) = f(1) = 1$, $\min_{t\in[0,1]} f(t) = f(0) = 0$.

Proof. Since $f'(t) = \frac{\exp(-t)}{1-\exp(-1)} > 0$, $t \in [0, 1]$, then $f(t)$ is increasing in $[0, 1]$.

Using the properties of $d(\widetilde{A}, \widetilde{B})$, and Lemma 1, we have that $d_e(\widetilde{A}, \widetilde{B})$, also satisfies the above properties (D1')-(D4'). Thus, $s_e(\widetilde{A}, \widetilde{B})$ satisfies the above properties (S1')-(S4') using the dual concept.

4 Examples

In this section, we present two examples to illustrate the proposed similarity measures s and s_e.

Example 1. Assume that there are two patterns denoted with type-2 fuzzy sets in $X = \{x\}$. The two patterns are denoted as follows:

$$\widetilde{A} = \{(x, \mu_{\widetilde{A}}(x))\} \text{ and } \widetilde{B} = \{(x, \mu_{\widetilde{B}}(x))\},$$

where

$$\mu_{\widetilde{A}}(x) = \{(0.8, 1), (0.7, 0.5), (0.6, 0.4)\},$$
$$\mu_{\widetilde{B}}(x) = \{(0.8, 1), (0.7, 0.6), (0.6, 0.5)\}.$$

Assume that a sample $\widetilde{C} = \{(x, \mu_{\widetilde{C}}(x))\}$ is given, where

$$\mu_{\widetilde{C}}(x) = \{(0.8, 1), (0.7, 0.7), (0.6, 0.6)\}.$$

By the above definitions, we know that the membership values of type-1 fuzzy sets $\mu_{\widetilde{A}}$ and $\mu_{\widetilde{C}}$ are $1, 0.5, 0.4$ and $1, 0.7, 0.6$, respectively. According to Eq. (6), we take $t_1 = 1$, $t_2 = 0.7$, $t_3 = 0.6$, $t_4 = 0.5$ and $t_5 = 0.4$. Then the level-t_1 sets of $\mu_{\widetilde{A}}$ and $\mu_{\widetilde{C}}$ are $\{0.8\}$ and $\{0.8\}$, respectively. Hence, the crisp Hausdorff distance between these two level-t_1 sets is 0. On the other hand, the level-t_2 sets of $\mu_{\widetilde{A}}$ and $\mu_{\widetilde{C}}$ are $\{0.8\}$ and $\{0.8, 0.7\}$, respectively. Thus, the crisp Hausdorff distance between these two level-t_2 sets is 0.1. By the same argument, we also have that the crisp Hausdorff distances of level-t_3 sets, level-t_4 sets and level-t_5 sets are 0.2, 0.1 and 0, respectively. Therefore, by Eq. (6), the fuzzy Hausdorff distance between type-1 fuzzy sets $\mu_{\widetilde{A}}$ and $\mu_{\widetilde{C}}$ is given by

$$H_f(\mu_{\widetilde{A}}, \mu_{\widetilde{C}}) = \frac{0.1 \times 0.7 + 0.2 \times 0.6 + 0.1 \times 0.5}{1 + 0.7 + 0.6 + 0.5 + 0.4} = 0.075.$$

Similarly, we have $H_f(\mu_{\widetilde{B}}, \mu_{\widetilde{C}}) = 0.046$. Using (11), we have

$$d^N(\widetilde{A}, \widetilde{C}) = 0.075, \quad \text{and} \quad d^N(\widetilde{B}, \widetilde{C}) = 0.046.$$

Therefore, we get

$$s(\widetilde{A}, \widetilde{C}) = 0.925, \quad \text{and} \quad s(\widetilde{B}, \widetilde{C}) = 0.954.$$

Based on the above, it is seen that the sample $\widetilde{C}$ belongs to the pattern $\widetilde{B}$ according to the principle of the maximum degree of similarity between type-2 fuzzy sets.

Example 2. Consider Gaussian type-2 fuzzy sets $\widetilde{A}_j$, $j = 1, 2, 3, 4, 5$, with a discrete x-domain consisting of only 3 points, $x_1 = 1$, $x_2 = 3$ and $x_3 = 5$. Suppose that $m(x_1) = 0.1$, $m(x_2) = 0.8$ and $m(x_3) = 0.6$.

$$\mu_{\widetilde{A}_j}(x_i) = \exp\{-(u - m(x_i))^2/2(\sigma_j(x_i))^2\}, \quad i = 1, 2, 3, \quad j = 1, 2, 3, 4, 5,$$

where $\sigma_1^2(x_i) = m(x_i)$, $\sigma_2^2(x_i) = 0.1m(x_i)$, $\sigma_3^2(x_i) = 0.05m(x_i)$, $\sigma_4^2(x_i) = 3m(x_i)$, $\sigma_5^2(x_i) = 0.01m(x_i)$, $i = 1, 2, 3$. We want to cluster $\widetilde{A}_j$, $j = 1, 2, 3, 4, 5$ based on the similarity values s and s_e. Our initial interest is in the normalized distance between $\widetilde{A}_1$ and $\widetilde{A}_2$. By (7), we obtain

$$H_f(\mu_{\widetilde{A}_1}(x_1), \mu_{\widetilde{A}_2}(x_1)) = 2\int_0^1 t(\sqrt{-2 \times 0.1 \ln t} - \sqrt{-2 \times 0.01 \ln t})dt = 0.192,$$

$$H_f(\mu_{\widetilde{A}_1}(x_2), \mu_{\widetilde{A}_2}(x_2)) = 2\int_0^1 t(\sqrt{-2 \times 0.1 \ln t} - \sqrt{-2 \times 0.01 \ln t})dt = 0.542,$$

$$H_f(\mu_{\widetilde{A}_1}(x_3), \mu_{\widetilde{A}_2}(x_3)) = 2\int_0^1 t(\sqrt{-2 \times 0.1 \ln t} - \sqrt{-2 \times 0.01 \ln t})dt = 0.469.$$

Hence, by (10), we have

$$d(\widetilde{A}_1, \widetilde{A}_2) = 0.192 + 0.542 + 0.469 = 1.203.$$

Next, we can obtain the normalized distance of $d(\widetilde{A}_1, \widetilde{A}_2)$ as follows:

$$d^N(\widetilde{A}_1, \widetilde{A}_2) = \frac{d(\widetilde{A}_1, \widetilde{A}_2)}{3} = 0.401.$$

Finally, the similarity measures $s(\widetilde{A}_1, \widetilde{A}_2)$ and $s_e(\widetilde{A}_1, \widetilde{A}_2)$ are obtained as follows:

$$s(\widetilde{A}_1, \widetilde{A}_2) = 1 - d^N(\widetilde{A}_1, \widetilde{A}_2) = 0.599,$$

$$s_e(\widetilde{A}_1, \widetilde{A}_2) = 1 - d_e(\widetilde{A}_1, \widetilde{A}_2) = 1 - \frac{1 - \exp(-d^N(\widetilde{A}_1, \widetilde{A}_2))}{1 - \exp(-1)} = 0.477.$$

By the similar argument, we can obtain the others which are presented in Tables 1 and 2.

Table 1. Similarities between Gaussian type-2 fuzzy sets using the similarity measure s.

	$\widetilde{A}_1$	$\widetilde{A}_2$	$\widetilde{A}_3$	$\widetilde{A}_4$	$\widetilde{A}_5$
$\widetilde{A}_1$	1.000	0.599	0.545	0.571	0.472
$\widetilde{A}_2$	0.599	1.000	0.946	0.170	0.873
$\widetilde{A}_3$	0.545	0.946	1.000	0.115	0.928
$\widetilde{A}_4$	0.571	0.170	0.115	1.000	0.043
$\widetilde{A}_5$	0.472	0.873	0.928	0.043	1.000

Table 2. Similarities between Gaussian type-2 fuzzy sets using the similarity measure s_e.

	$\widetilde{A}_1$	$\widetilde{A}_2$	$\widetilde{A}_3$	$\widetilde{A}_4$	$\widetilde{A}_5$
$\widetilde{A}_1$	1.000	0.477	0.422	0.448	0.351
$\widetilde{A}_2$	0.477	1.000	0.917	0.108	0.811
$\widetilde{A}_3$	0.422	0.917	1.000	0.071	0.890
$\widetilde{A}_4$	0.448	0.108	0.071	1.000	0.026
$\widetilde{A}_5$	0.351	0.811	0.890	0.026	1.000

By examining the original relations in Tables 1 and 2, one finds that

$$s(\widetilde{A}_2, \widetilde{A}_3) = 0.946,\ s(\widetilde{A}_2, \widetilde{A}_5) = 0.873,\ s(\widetilde{A}_3, \widetilde{A}_5) = 0.928,$$

$$s_e(\widetilde{A}_2, \widetilde{A}_3) = 0.917,\ s_e(\widetilde{A}_2, \widetilde{A}_5) = 0.811,\ s_e(\widetilde{A}_3, \widetilde{A}_5) = 0.890,$$

respectively. Intuitively, it seems to be reasonable that $\widetilde{A}_2$, $\widetilde{A}_3$, $\widetilde{A}_5$ can be in a class, but it is not so clear that $\widetilde{A}_1$ and $\widetilde{A}_4$ are in a class.

To obtain greater perspective for clustering $\widetilde{A}_1$, $\widetilde{A}_2$, $\widetilde{A}_3$, $\widetilde{A}_4$ and $\widetilde{A}_5$, Yang and Shih's algorithm (see Appendix), a clustering method based on fuzzy relations by beginning with a similarity matrix, is applied to these Gaussian type-2 fuzzy sets by beginning with Tables 1 and 2, respectively.

From Table 1 we obtain the similarity matrix $R_1^{(0)}$ on $\{\widetilde{A}_1, \widetilde{A}_2, \widetilde{A}_3, \widetilde{A}_4, \widetilde{A}_5\}$ with

$$R_1^{(0)} = \begin{pmatrix} 1 & & & & \\ 0.599 & 1 & & & \\ 0.545 & 0.946 & 1 & & \\ 0.571 & 0.170 & 0.115 & 1 & \\ 0.472 & 0.873 & 0.928 & 0.043 & 1 \end{pmatrix}.$$

Then, by max-Δ composition, we have

$$R_1^{(2)} = R_1^{(3)} = \begin{pmatrix} 1 & & & & \\ 0.599 & 1 & & & \\ 0.545 & 0.946 & 1 & & \\ 0.571 & 0.170 & 0.116 & 1 & \\ 0.473 & 0.874 & 0.928 & 0.044 & 1 \end{pmatrix}.$$

is a max-Δ similarity relation matrix. The following clustering results are obtained using Yang and Shih's algorithm:

$$0 < \alpha \le 0.044 \Rightarrow \{\widetilde{A}_1, \widetilde{A}_2, \widetilde{A}_3, \widetilde{A}_4, \widetilde{A}_5\},$$

$$0.044 < \alpha \le 0.473 \Rightarrow \{\widetilde{A}_1, \widetilde{A}_2, \widetilde{A}_3, \widetilde{A}_5\}, \{\widetilde{A}_4\},$$

$$0.473 < \alpha \le 0.571 \Rightarrow \{\widetilde{A}_2, \widetilde{A}_3, \widetilde{A}_5\}, \{\widetilde{A}_1, \widetilde{A}_4\},$$

$$0.571 < \alpha \le 0.874 \Rightarrow \{\widetilde{A}_2, \widetilde{A}_3, \widetilde{A}_5\}, \{\widetilde{A}_1\}, \{\widetilde{A}_4\},$$

$$0.874 < \alpha \le 0.946 \Rightarrow \{\widetilde{A}_2, \widetilde{A}_3\}, \{\widetilde{A}_1\}, \{\widetilde{A}_4\}, \{\widetilde{A}_5\},$$

$$0.946 < \alpha \le 1 \Rightarrow \{\widetilde{A}_1\}, \{\widetilde{A}_2\}, \{\widetilde{A}_3\}, \{\widetilde{A}_4\}, \{\widetilde{A}_5\}.$$

From Table 2, we can also obtain the other similarity matrix $R_2^{(0)}$ on $\{\widetilde{A}_1, \widetilde{A}_2, \widetilde{A}_3, \widetilde{A}_4, \widetilde{A}_5\}$ with

$$R_2^{(0)} = \begin{pmatrix} 1 & & & & \\ 0.477 & 1 & & & \\ 0.422 & 0.917 & 1 & & \\ 0.448 & 0.108 & 0.071 & 1 & \\ 0.351 & 0.811 & 0.890 & 0.026 & 1 \end{pmatrix}.$$

Then, by max-Δ composition, we also have

$$R_2^{(0)} = R_2^{(1)} = \begin{pmatrix} 1 & & & & \\ 0.477 & 1 & & & \\ 0.422 & 0.917 & 1 & & \\ 0.448 & 0.108 & 0.071 & 1 & \\ 0.351 & 0.811 & 0.890 & 0.026 & 1 \end{pmatrix}.$$

is a max-Δ similarity relation matrix. The following clustering results are also obtained using Yang and Shih's algorithm:

$$0 < \alpha \le 0.026 \Rightarrow \{\widetilde{A}_1, \widetilde{A}_2, \widetilde{A}_3, \widetilde{A}_4, \widetilde{A}_5\},$$

$$0.026 < \alpha \le 0.351 \Rightarrow \{\widetilde{A}_1, \widetilde{A}_2, \widetilde{A}_3, \widetilde{A}_5\}, \{\widetilde{A}_4\},$$

$$0.351 < \alpha \le 0.448 \Rightarrow \{\widetilde{A}_2, \widetilde{A}_3, \widetilde{A}_5\}, \{\widetilde{A}_1, \widetilde{A}_4\},$$

$$0.448 < \alpha \le 0.811 \Rightarrow \{\widetilde{A}_2, \widetilde{A}_3, \widetilde{A}_5\}, \{\widetilde{A}_1\}, \{\widetilde{A}_4\},$$

$$0.811 < \alpha \le 0.917 \Rightarrow \{\widetilde{A}_2, \widetilde{A}_3\}, \{\widetilde{A}_1\}, \{\widetilde{A}_4\}, \{\widetilde{A}_5\},$$

$$0.917 < \alpha \le 1 \Rightarrow \{\widetilde{A}_1\}, \{\widetilde{A}_2\}, \{\widetilde{A}_3\}, \{\widetilde{A}_4\}, \{\widetilde{A}_5\}.$$

Based on the above clustering results, we find that they are reasonable consisting of a hierarchical tree according to different levels of α.

5 Conclusions

In this paper, we presented the axiom definition for a similarity measure between type-2 fuzzy sets. We proposed two simple similarity measures to measure the degree of similarity between type-2 fuzzy sets, based on the Hausdorff distance between type-1 fuzzy sets. For practical use, we showed how to compute the similarities between Gaussian type-2 fuzzy sets. Furthermore, we applied Yang and Shih's algorithm, which is a clustering method based on fuzzy relations by beginning with a similarity matrix, to these Gaussian type-2 fuzzy sets beginning with these similarities. The clustering results are reasonable consisting of a hierarchical tree according to different levels of α.

Appendix (see Yang and Shih [23])

A crisp relation R between two sets, X and Y, is defined as a subset of $X \times Y$. Denoted by $R(X,Y)$, this relation is associated with an indicator function $\mu_R(x,y) \in \{0,1\}$, $\forall(x,y) \in X \times Y$. That is, $\mu_R(x,y) = 1$ if $(x,y) \in R(X,Y)$, and $\mu_R(x,y) = 0$ if $(x,y) \notin R(X,Y)$. Zadeh [23] defined a fuzzy relation R between X and Y as a fuzzy subset of $X \times Y$ by an extension to allow $\mu_R(x,y)$ to be membership functions assuming values in the interval $[0,1]$. The value of $\mu_R(x,y)$ represents the strength of the relationship between x and y. Furthermore, he defined a similarity relation R in X if and only if R is a fuzzy relation and for all x, y, $z \in X$, $\mu_R(x,x) = 1$ (reflexivity), $\mu_R(x,y) = \mu_R(y,x)$ (symmetry), and $\mu_R(x,z) \geq \max_{y\in X}\{\min\{\mu_R(x,y), \mu_R(y,z)\}\}$ (transitivity). Later, Bezdek and Harris [1] proposed a max-Δ transitivity: $\mu_R(x,z) \geq \max_{y\in X}\{\max\{0, \mu_R(x,y) + \mu_R(y,z) - 1\}\}$.

There are many data presented in subjective relations are presented by a proximity relation which is a fuzzy relation with reflexivity and symmetry. Since proximity relations do not have transitivity, they cannot be used in clustering. Theorem A states that a max-Δ similarity relation is obtained by an n-step procedure base on max-Δ compositions. Then, this max-Δ similarity-relation matrix is used for clustering. Beginning with a proximity-relation matrix, this clustering algorithm produces a level α $(0 < \alpha \leq 1)$ partition through an n-step procedure based on max-Δ compositions. The algorithm is stated as follows.

Definition A. (max-Δ composition) Given an initial fuzzy-relation matrix, $R^{(0)} = [r_{ij}^{(0)}]$, then $R^{(n)} = [r_{ij}^{(n)}]$ with $r_{ij}^{(n)} = \max_k\{\max\{0, r_{ik}^{(n-1)} + r_{kj}^{(n-1)} - 1\}\}$, $n = 1,2,3,\cdots$ is called a max-Δ composition.

Theorem A. (An n-step procedure) Suppose that $R^{(0)}$ is a proximity-relation matrix. Then, by max-Δ compositions, one has

$$I < R^{(0)} < R^{(1)} < \cdots < R^{(n)} = R^{(n+1)} = \cdots,$$

where $R^{(n)}$ is a max-Δ similarity relation. If n is not finite, then $\lim_{n\to\infty} R^{(n)} = R^{(\infty)}$ with $R^{(\infty)}$ a max-Δ similarity relation, i.e.

$$I < R^{(0)} < R^{(1)} < \cdots < R^{(n)} < R^{(n+1)} < \cdots < R^{(\infty)}.$$

Yang and Shih's Clustering Algorithm:

S0. Let $R^{(0)} = [r_{ij}^{(0)}]_{n\times n}$ be a given proximity-relation matrix and let $I = \{1, 2, \cdots, n\}$ be the index set. It is given that α with $0 < \alpha \leq 1$.

S1. Obtain a max-Δ similarity-relation matrix $R = [r_{ij}]_{n\times n}$ from the given proximity-relation matrix $R^{(0)}$ using the n-step procedure.

S2. Set $r_{ij} = 0$ for all $i = j$ and set $r_{ij} = 0$ for all $r_{ij} < \alpha$.

S3. Choose s and t in I so that $r_{st} = \max\{r_{ij} | i < j, i, j \in I\}$. Note that a tie is broken randomly.

IF $r_{st} \neq 0$ *THEN* link s and t into the same cluster $C = \{s, t\}$ and *GOTO* S4.

ELSE PRINT all indices in I into separated cluster and *STOP*.

S4. Choose u in $I \setminus C$ so that

$$\sum_{i\in C} r_{iu} = \max\left\{\sum_{i\in C} r_{ij} | j \in I \setminus C, \text{ with } r_{ij} \neq 0 \text{ for all } i \in C\right\}.$$

A tie is broken randomly.

IF there is such an u, *THEN* link u into C, i.e. $C = \{s, t, u\}$ and *GOTO* S4.

ELSE PRINT the cluster C.

S5. Let $I = I \setminus C$ and *GOTO* S3.

Acknowledgements

The authors are grateful to anonymous referees for their helpful comments and suggestions to improve the presentation of the paper.

References

[1] J.C. Bezdek and J.D. Harris, Fuzzy partitions and relations: an axiomatic basis for clustering, *Fuzzy Sets and Systems* **1** (1978) 111-127.

[2] J.J. Buckley and Y. Hayashi, Fuzzy input-output controllers are universal approximates, *Fuzzy Sets and Systems* **58** (1993) 273-278.

[3] K.S. Candan, W.S. Li and M.L. Priya, Similarity-based ranking and query processing in multimedia databases, *Data & Knowledge Engineering* **35** (2000) 259-298.

[4] B.B. Chaudhuri and A. Rosenfeld, A modified Hausdorff distance between fuzzy sets, *Information Sciences* **118** (1999) 159-171.

[5] D. Dubois and H. Prade, Operations on fuzzy numbers, *Int. J. Syst. Sci.* **9** (1978), 613-626.

[6] D. Dubois and H. Prade, Operations in a fuzzy-valued logic, *Inform. and Control* **43** (1979), 224-240.

[7] D. Dubois and H. Prade, *Fuzzy sets and Systems: Theory and Applications.* New York: Academic, 1980.

[8] N.N. Karnik and J.M. Mendel, Operations on type-2 fuzzy sets, *Fuzzy Sets and Systems* **122** (2001) 327-348.

[9] N.N. Karnik and J.M. Mendel, Centroid of a type-2 fuzzy sets, *Information Sciences* **132** (2001) 195-220.

[10] N.N. Karnik and J.M. Mendel, Applications of type-2 fuzzy logic systems to forecasting of time series, *Information Sciences* **120** (1999) 89-111.

[11] J.M, Mendel and Robert I. Bob John, Type-2 fuzzy sets made simple, *IEEE Trans. Fuzzy Syst.* **10** (2002) 117-127.

[12] Liu Xuechang, Entropy, distance measure and similarity measure of fuzzy sets and their relations, *Fuzzy Sets and Systems* **52** (1992) 305-318.

[13] M. Mizumoto and K. Tanaka, Some properties of fuzzy sets of type-2, *Inform. Control* **31** (1976) 312-340.

[14] M. Mizumoto and K. Tanaka, Fuzzy sets of type 2 under algebraic product and algebraic sum, *Fuzzy Sets and Systems* **5** (1981) 277-290.

[15] J. Nieminen, On the algebraic structure of fuzzy sets of type-2, *Kybernetica* **13** (4) (1977).

[16] N.R. Pal and S.K. Pal, Entropy, a new definition and its applications, *IEEE. Trans. System Man Cybernet* **21** (1991) 1260-1270.

[17] N.R. Pal and S.K. Pal, Some properties of the exponential entropy, *Inform. Sci.* **66** (1992) 119-137.

[18] Q. Liang and J.M, Mendel, Interval type-2 fuzzy logic systems: Theory and design, *IEEE Trans. Fuzzy Syst.* **8** (2000) 535-550.

[19] I.B. Turksen and Z. Zhong, An appoximate analogical reasoning approach based on similarity measures, *IEEE Trans. System Man Cybernet* **18** (1988) 1049-1056.

[20] K.L. Wu and M.S. Yang, Alternative c-means clustering algorithms, *Pattern Recognition* **35** (2002) 2267-2278.

[21] R.R. Yager, On the measure of fuzziness and negation, Part I: membership in the unit interval, *Internat. J. General Systems* **5** (1979) 221-229.

[22] M.S. Yang and H.M. Shih, Cluster analysis on fuzzy relations, *Fuzzy Sets and Systems* **120** (2001) 197-212.

[23] L.A. Zadeh, Similarity relations and fuzzy orderings, *Information Sciences* **3** (1971) 177-200.

[24] L.A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning-1, *Information Sciences* **8** (1975) 199-249.