A New Distance for Intuitionistic Fuzzy Sets Based on Similarity Matrix

CUIPING CHENG¹, FUYUAN XIAO², and ZEHONG CAO²

¹School of Computer and Information Science, Southwest University, Chongqing 400715, China
²Discipline of ICT, School of Technology, Environments and Design, University of Tasmania, Hobart, TAS 7001, Australia

Corresponding author: Fuyuan Xiao (doctorxiaofy@hotmail.com)

This work was supported by the Chongqing Overseas Scholars Innovation Program under Grant cx2018077.

ABSTRACT Measuring the distance between two intuitionistic fuzzy sets (IFSs) is an open issue. Many types of distances for the IFSs have been proposed in previous studies. Some existing methods cannot satisfy the axioms of similarity or provide counterintuitive cases. Others ignore the relationship between three parameters characterizing the information carried by the IFS. To address these issues, a new distance is proposed by analyzing the similarity among the three parameters of the IFS. The comparison with some existing distances illustrates that the new distance has a higher sensitivity and can effectively measure the similarity between the IFSs. The results of the application of pattern recognition are also shown that the proposed method has better recognition ability.

INDEX TERMS Intuitionistic fuzzy set, distance function, similarity measure, pattern recognition, medical diagnosis.

I. INTRODUCTION

In the real world, human cognition of material objects is often vague, and the measurement of uncertainty has attracted increasing interest [1]–[3]. Many math tools such as probability [4], possibility theory [5], rough sets [6], [7], belief structure [8]–[10], entropy function [11]–[15], Z numbers [16], [17] and D numbers [18]–[22] are presented and are wildly used in lots of real engineering. Among these tools, fuzzy sets theory, founded by Zadeh [23], is flexible to model linguistic uncertainty and is efficient in decision making [24]–[26]. In fuzzy set theory, a single value between zero and one is used to indicate the membership of the element. But in reality, sometimes the uncertainty of information is not completely grasped by the fuzzy set, so Atanassov [27] introduced the concept of an intuitionistic fuzzy set which consists of membership function, non-membership function and hesitancy degree. As a generalization of fuzzy sets, IFSs are considered to be more effective way to deal with vagueness than fuzzy set. It is obvious that an intuitionistic fuzzy set becomes a fuzzy set if there is no hesitation. Due to the efficiency of IFSs, it has received great attention from researchers, and at present, it has been applied in various areas such as decision making [28]–[30], medical diagnosis [31], [32], pattern recognition [33], [34] and other areas [35]–[40].

It is often inevitable to measure the similarity between IFSs in practical application. The similarity measures of IFSs have received a great attention and several distances for IFSs have been proposed [41], [42]. Chen defined the first similarity measure between IFSs which claimed that IFS is similar to vague set. Szmidt and Kacprzyk [43] proposed four distances for computing the distance measure between IFSs which were based on the geometric interpretation of IFS. Grzegorzewski [45] proposed distances based on the Hausdorff metric, the proposed new distances are straightforward generalizations of the well known Hamming distance, the Euclidean distance and their normalized counterparts. But later Chen [46] implied that some limitations exist in Grzegorzewski’s two dimensional Hausdorff based distances by showing some counter-intuitive cases. Xu [47] put forward some weighted distances on the basis of this geometric distance model. Wang and Xin [48] proposed several new distances and applied them to pattern recognition, the axiom definition of distance measure between intuitionistic fuzzy sets (IFSs) is introduced, and corresponding proofs are given, the relations between similarity measure and distance...
measure of IFSs are analyzed. Hung and Yang [49] proposed similarity measures of intuitionistic fuzzy sets based on L-p metric, and they apply the proposed measures to analyze the behavior of decision making. Yang and Chiehana [50] suggested that the three dimensional interpretation of IFSs could lead to different comparison results to the ones obtained with their two dimensional counterparts, and introduced several extended 3D Hausdorff based distances. Hung and Yang [51] presented two new similarity measures between intuitionistic fuzzy sets, and then apply the new measures to evaluate students’ answer-scripts. Hatzimichailidis et al. [52] introduced a distance measure which formulates the information of each set in matrix structure, where matrix norms in conjunction with fuzzy implications can be applied to measure the distance between the IFSs. Since the Sugeno integral provides an expected-value-like operation, it can be a useful tool in defining the expected total similarity degree between two intuitionistic fuzzy sets. Hwang et al. [53] proposed a new similarity measure with its application to pattern recognition. More discussion on similarity measures of IFSs can be found in [54], [55].

Different distances have different focus and have different advantages in measuring the similarity of IFSs, as IFSs is characterized by hesitant index to describe the state of “membership degree or non-membership degree”, unlike the first two parameters of IFSs, the information carried by the hesitancy degree is uncertain, so these three parameters should not be considered separately when comparing the differences between IFSs, but most of the existing measures only consider the differences between numerical values of the IFS parameters, and ignore the characteristics of intuitionistic fuzzy information. Therefore, in order to make the results more reliable, this paper proposes a new distance measure by analyzing the interaction between the three parameters.

The rest of this paper is organized as follows. In Section II, the basic concepts of IFSs are reviewed, some existing distances of are introduced. In Section III, the new distance measure for IFSs is proposed, and the geometric meaning of the new distance measurement is explained. In Section IV, the effectiveness of the new distance is illustrated through numerical comparisons and application of pattern recognition. The conclusion is given in Section V.

II. PRELIMINARIES

In this section, some basic concepts on IFS are introduced.

A. INTUITIONISTIC FUZZY SET [27]

Definition 1: Let A be an IFS in the finite universe of discourse X can be written as:

\[ A = \{(x, \mu_A(x), v_A(x)) | x \in X\} \]

here \( \mu_A : X \rightarrow [0, 1] \), \( v_A : \rightarrow [0, 1] \) with the condition \( 0 \leq \mu_A(x) + v_A(x) \leq 1 \) \( \forall x \in X \). The numbers \( \mu_A(x) \) and \( v_A(x) \) denote the degree of membership and non-membership of \( x \) to \( A \). For any \( x \in X \), \( 0 \leq \mu_A(x) + v_A(x) \leq 1 \).

The hesitance index \( \pi_A(x) = 1 - \mu_A(x) - v_A(x) \) is used to measure hesitancy degree of \( x \) to \( A \), and for any \( x \in X \), \( 0 \leq \pi_A(x) \leq 1 \).

Definition 2 [56]: Let A and B be two IFSs in the universe of discourse X, then

1. \( A \subseteq B \) if and only if \( \forall x \in X, \mu_A(x) \leq \mu_B(x) \), and \( v_A(x) \geq v_B(x) \);
2. \( A = B \) if and only if \( \forall x \in X, \mu_A(x) = \mu_B(x) \), and \( v_A(x) = v_B(x) \);
3. \( A^C = \{(x, \mu_A(x), v_A(x)) | x \in X\} \) where \( A^C \) is the complement of \( A \).

B. EXISTING DISTANCE MEASURES BETWEEN IFSS

Distance measure plays a very important role in many applications such as pattern recognition [57]–[60] and decision making [61]–[63]. In this section, several widely used distance measures are reviewed. Let X be the universe of discourse, \( A = \{(x, \mu_A(x), v_A(x)) | x \in X\} \) and \( B = \{(x, \mu_B(x), v_B(x)) | x \in X\} \) are two IFSs in \( X = \{x_1, x_2, \ldots, x_n\} \). Szmidt and Kacprzyk [28] proposed distances between IFSs using the Hamming distance, Euclidean distance as follows:

The Hamming distance:

\[ d_H(A, B) = \frac{1}{2} \sum_{i=1}^{n} |\mu_A(x_i) - \mu_B(x_i)| + |v_A(x_i) - v_B(x_i)| \]  (2)

The normalized Hamming distance:

\[ d_{Hn}(A, B) = \frac{1}{2n} \sum_{i=1}^{n} |\mu_A(x_i) - \mu_B(x_i)| + |v_A(x_i) - v_B(x_i)| \]  (3)

The Euclidean distance:

\[ d_E(A, B) = \left( \frac{1}{2} \sum_{i=1}^{n} (\mu_A(x_i) - \mu_B(x_i))^2 + (v_A(x_i) - v_B(x_i))^2 \right)^{\frac{1}{2}} \]  (4)

The normalized Euclidean distance:

\[ d_{En}(A, B) = \left( \frac{1}{2n} \sum_{i=1}^{n} (\mu_A(x_i) - \mu_B(x_i))^2 + (v_A(x_i) - v_B(x_i))^2 \right)^{\frac{1}{2}} \]  (5)

The Hausdorff’s distance measure [44]:

\[ d_{H}(A, B) = \frac{1}{n} \sum_{i=1}^{n} \max\{|\mu_A(x_i) - \mu_B(x_i)|, |v_A(x_i) - v_B(x_i)|\} \]  (6)

Yang and Chiehana introduced the hesitancy degree into the expression based on the Hausdorff’s distance [50]:

\[ d_{HC}(A, B) = \frac{1}{n} \sum_{i=1}^{n} \max\{|\mu_A(x_i) - \mu_B(x_i)|, |v_A(x_i) - v_B(x_i)|, \] \( |v_A(x_i) - v_B(x_i)|\} \]  (7)
The greater the distance between two intuitionistic fuzzy sets, the smaller the similarity between them.

Definition 3 [48]: Let d be a mapping d: IFSs \((X) \times IFSs (X) \to [0, 1]\). If \(d(A, B)\) satisfies the following properties, \(d(A, B)\) is a distance measure between IFSs A and B.

- (P1) \(0 \leq d(A, B) \leq 1\);
- (P2) \(d(A, B) = 0\) if and only if \(A = B\);
- (P3) \(d(A, B) = d(B, A)\);
- (P4) If \(A \subseteq B \subseteq C\) then \(d(A, C) \geq d(A, B)\) and \(d(A, C) \geq d(B, C)\).

Because the hesitance index could be expressed in terms of the membership and non-membership degrees, it is argued that the hesitance index does not have to be taken into account by the distance measure. Therefore, according to whether the hesitance index is used, the existing distance can be divided into two dimensional distance and three dimensional distance. The three dimensional distances proposed by Szmidt and Kacprzyk have excellent geometric properties, whether using hamming distance or Euclidean distance to measure similarity, however, the difference caused by different parameters contributes equally to the distance. As the hesitance index indicates the unknown membership and non-membership degree, the hesitation index is somewhat similar to the first two parameters, so the three parameters cannot be measured separately, otherwise the counter-intuitive results would be obtained. For example, suppose four IFSs A, B, C, and D on \(X = \{x\}, A = \{\langle x, 0, 1 \rangle \}, B = \{\langle x, 1, 0 \rangle \}, C = \{\langle x, 0, 0 \rangle \}, D = \{\langle x, 0, 0.5 \rangle \}\). the information obtained from IFSs A and B is certain, while IFSs C and D are highly uncertain, so \(d(C, D)\) should be less than \(d(A, D)\). However, we have \(d_H(A, B) = 1, d_H(A, C) = 1, d_H(A, D) = 1, d_E(A, B) = 1, d_E(A, C) = 1, d_E(A, D) = 0.866\), it is obvious that the results are counter-intuitive, this is actually a typical example, because many existing distances cannot solve this similarity problem well, so useful conclusions can be obtained by analyzing this set of classical intuitionistic fuzzy sets. Since the uncertainty expressed by the hesitance index is related to the first two parameters, it is unreasonable to separately calculate the numerical difference between the three parameters. The geometric meaning of this example is shown below in Fig. 1. Because the three axes in the Cartesian coordinate system of the three-dimensional space are equivalent, the distance is independent of the parameters corresponding to the axis, so even if the coordinate axis corresponding to the parameter changes, the distance of the IFSs do not change. Therefore, the sensitivity of these two methods is relatively low.

Another measure is the two-dimensional distance without considering the hesitance index, equations 5 and 6 are classical two-dimensional distances. Although the hesitance index can be expressed in terms of two other parameters, but the information represented by the hesitance index is lost, and sometimes the results are counterintuitive. For example, Suppose A, B, and C are intuitionistic fuzzy sets in X, \(A = \{\langle x, 0.5, 0.5 \rangle \}, B = \{\langle x, 0.5, 0 \rangle \}, C = \{\langle x, 0, 0 \rangle \}\). we have \(d_{Hd}(A, B) = 0.5, d_{Hd}(A, C) = 0.5, d_{Hd}(B, C) = 0.5, d_{WX}(A, B) = 0.375, d_{WX}(A, C) = 0.5, d_{WX}(B, C) = 0.375\). It can be seen that these two distances cannot effectively distinguish the difference between the IFSs A, B and C. When measuring the similarity between intuitionistic fuzzy sets, all information contained in the intuitionistic fuzzy sets should be considered as much as possible, but uncertainty of IFSs is not well expressed in two dimensions distances because the hesitance index is ignored. Through the analysis of the above four distances, it is shown that the hesitance index cannot be ignored when measuring the similarity of IFSs, these three parameters have a certain contribution to the description of the uncertainty of information, and the hesitance index is not completely different from the first two parameters.

III. NEW METHOD FOR MEASURING THE SIMILARITY OF INTUITIONISTIC FUZZY SETS

In this section, a new method to measure the similarity between intuitionistic fuzzy sets is introduced. First, the definition of the score function is given, which is used to prove that the new distance in measuring the similarity between IFSs is related to the uncertainty of IFSs.

A. SCORE FUNCTION

The concept of score functions was initiated by Chen and Tan [64], and it is used to solve multi-attribute decision making problems of intuitionistic fuzzy set.

Definition 4: Suppose A is intuitionistic fuzzy set on the universe \(X = \{x_1, x_2, \ldots, x_n\}, A = \{\langle x, \mu(x), \nu(x) \rangle \}, \) define
$S_A$ as score function, the expression of $S_A$ is

$$S_A(x_i) = \mu(x_i) - \nu(x_i)$$

with $-1 \leq S_A(x_i) \leq 1$.

The degree that an alternative meets the decision maker’s expectations can be measured by score function $S(A)$, the larger the value of $S(A)$ is, the more suitable the alternative $A$ is for the decision-maker. The larger the value of $\mu(x_i)$, the smaller the value of $\nu(x_i)$, the greater the probability of $x \in A$, and the larger the value of the score function $S_A(x_i)$. So $S_A(x)$ can be applied to describes the degree of support of about element $x \in A$. If $S_A(x) > 0$, then the higher the degree of $x$ belongs to $A$, if $S_A(x) < 0$, then the lower the degree of $x$ belongs to $A$. In addition, the absolute value of $S$ can characterizes the degree of certainty of $A$. The larger $|S|$, then the ambiguity of the IFS is lower. For example, IFSs $A = \{(x, 0.25, 0.25)\}$ and $B = \{(x, 0.1, 0.4)\}$, although the hesitance index of these two fuzzy sets is equal, set $B$ provides more information, which means that $B$ is more certain, it is reasonable for the result that $|S_A(x)| = 0$, $|S_B(x)| = 0.3$. Therefore, the absolute value of the $S(A)$ can be used to represent the certainty of the fuzzy set.

**B. SIMILARITY MATRIX**

Let’s first discuss a classic intuitionistic fuzzy sets, for two IFSs $A$ and $B$ on $X = \{x\}$, where $A = \{(x, 0, 0)\}$, $B = \{(x, 0.5, 0.5)\}$. The hesitation index in $A$ is equal to 1, which means that $A$ is completely uncertain whether $x$ belongs to $X$, while the hesitation index in $B$ is equal to 0, but the membership degree in $B$ is equal to the non-membership degree, so the ambiguity of $B$ is also very high. This example is sufficient to show that there is a certain degree of similarity between the hesitation index and the other two parameters, as the hesitation index is not independent of the other two parameters in the new distance, considering the similarity between the three parameters, the similarity matrix is introduced to express the “similarity” between the parameters.

**Definition 5:** $M$ is defined as the similarity matrix

$$M = \begin{pmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
\frac{1}{2} & \frac{1}{2} & \frac{\sqrt{3}}{2}
\end{pmatrix}$$

The elements in the similarity matrix are determined based on the similarity between the three parameters, because uncertainty is also expressed when the membership is equal to the non-membership, so the hesitation index is divided into two parts by $\frac{1}{2}$ and then assigned to membership and nonmembership. After the similarity between the three parameters is described, the uncertainty of the intuitionistic fuzzy set should be decreased, so change the value of the hesitation index by $\frac{\sqrt{3}}{2}$. As shown in Fig. 2, projecting the hesitation index $\pi$ in a space Cartesian coordinate system. The projection length of the $\pi$-axis in the plane $\mu-\nu$ is equal to $\frac{1}{2}\pi$, and the corresponding new hesitation index in the space Cartesian coordinate system is equal to $\frac{\sqrt{3}}{2}\pi$. Therefore, the elements in the similarity matrix are determined according to the projection relationship of the Cartesian coordinate system. The numerical values of the three parameters are changed by the similarity matrix, and in fact the metric space of the distance is also changed.

**C. METRIC SPACE**

The distance proposed in this paper is measured in the form of a vector in three-dimensional space, But the metric space of the new distance is different from the classical three-dimensional space. The membership and the nonmembership are independent of each other, so the $\mu$-axis and the $\nu$-axis are orthogonal. When the membership and the nonmembership are equal, the information given by the intuitionistic fuzzy set in this case is also uncertain, while the hesitation index is directly used to represent the uncertainty of the intuitionistic fuzzy set, so the $\pi$-axis and the other two axes are not orthogonal, but has a certain angle of inclination. The geometric meaning of the metric space is shown in the Fig. 2. Give an IFS $A = \{(0, 0, 1)\}$ in the Cartesian coordinate system, we can transform this IFS through the similarity matrix to get the coordinates in the new metric space, and the process is as follows:

$$\begin{pmatrix}
1 \\
0 \\
\frac{1}{2}
\end{pmatrix} \times \begin{pmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
\frac{1}{2} & \frac{1}{2} & \frac{\sqrt{3}}{2}
\end{pmatrix} = \begin{pmatrix}
1 & 1 & \sqrt{3} \\
2 & 2 & \frac{\sqrt{3}}{2}
\end{pmatrix}$$

**D. THE NEW DISTANCE**

**Definition 6:** Let $A$ and $B$ be two IFSs on the universe $X = \{x_1, x_2, \ldots, x_n\}$. The new distance is defined as

$$d_N(A, B) = \sqrt{n \sum_{i=1}^{n} \frac{\mu_A^2(x_i) + \nu_A^2(x_i) + \pi_A^2(x_i) + \mu_B^2(x_i) + \nu_B^2(x_i) + \pi_B^2(x_i)}{\overline{\sigma_i(M)}^T \overline{\sigma_i(M)}}}$$

where $\overline{\sigma_i(M)} = (1, 1, \sqrt{3})$.
with row vector \(\bar{m}_i = (\mu_A(x_i) - \mu_B(x_i), \nu_A(x_i) - \nu_B(x_i), \pi_A(x_i) - \pi_B(x_i))\).

Suppose A and B are two intuitionistic fuzzy sets on X = \{x\}. Only when A = \{(x, 1, 0)\} and B = A = \{(x, 0, 1)\}, the distance between the two intuitionistic fuzzy sets is equal to the maximum value 1. That is, when there is no hesitation in IFSs A and B, and the membership of X in them is completely opposite, then the similarity between the two intuitionistic fuzzy sets is equal to 0. This result is reasonable, because the membership degree and the non-membership degree are completely opposite states, and the hesitation index is an uncertain expression of the membership degree and the non-membership degree.

**Theorem 1:** \(d(A, B)\) is the distance between two IFSs A and B in X, the proposed distance measure satisfies the Definition 3.

**Proof (P1):**

\[
\bar{m}_i M = ((\mu_A(x_i) - \mu_B(x_i)) + \frac{1}{2}(\pi_A(x_i) - \pi_B(x_i)),
(u_A(x_i) - u_B(x_i)) + \frac{1}{2}(\pi_A(x_i) - \pi_B(x_i)),
\sqrt{\frac{3}{2}(\pi_A(x_i) - \pi_B(x_i))} \bar{m}_i M (\bar{m}_i M)^T
\]

\[
= ((\mu_A(x_i) - \mu_B(x_i)) + \frac{1}{2}(\pi_A(x_i) - \pi_B(x_i)))^2 + ((u_A(x_i) - u_B(x_i)) + \frac{1}{2}(\pi_A(x_i) - \pi_B(x_i)))^2 + \frac{3}{4}((\pi_A(x_i) - \pi_B(x_i)))^2 = (\mu_A(x_i) - \mu_B(x_i))^2
\]

\[
+ (u_A(x_i) - u_B(x_i))^2 + \frac{5}{2}(\pi_A(x_i) - \pi_B(x_i))^2 + (\mu_A(x_i) - \mu_B(x_i))2(\pi_A(x_i) - \pi_B(x_i)) + (u_A(x_i) - u_B(x_i))2(\pi_A(x_i) - \pi_B(x_i))
\]

\[
= \mu_A(x_i)^2 + u_A(x_i)^2 + \frac{1}{2}\pi_A(x_i)^2 + \mu_B(x_i)^2 + u_B(x_i)^2 + \frac{1}{4}\pi_B(x_i)^2 - \frac{1}{2}\mu_A(x_i)\mu_B(x_i) - 2\nu_A(x_i)\nu_B(x_i)
\]

\[
+ \frac{1}{2}\pi_A(x_i)\pi_B(x_i) \leq \mu_A^2(x_i) + u_A^2(x_i)
\]

\[
+ \pi_A^2(x_i) + \mu_B^2(x_i) + u_B^2(x_i) + \pi_B^2(x_i)
\]

So

\[
0 \leq \frac{\bar{m}_i M (\bar{m}_i M)^T}{\mu_A^2(x_i) + u_A^2(x_i) + \pi_A^2(x_i) + \mu_B^2(x_i) + u_B^2(x_i) + \pi_B^2(x_i)} \leq 1
\]

\[
0 \leq \frac{1}{n} \sum_{i=1}^{n} \bar{m}_i M (\bar{m}_i M)^T \leq 1.
\]

And therefore we have \(0 \leq d_n(A, B) \leq 1\).

**Proof (P2):** If \(d_n(A, B) = 0\), then \(\bar{m}_i M (\bar{m}_i M)^T\) must equal to 0, so we have \(\bar{m}_i = (0, 0, 0)\), and because \(\bar{m}_i = (\mu_A(x_i) - \mu_B(x_i), \nu_A(x_i) - \nu_B(x_i), \pi_A(x_i) - \pi_B(x_i))\), so \(\mu_A(x_i) = \mu_B(x_i), \nu_A(x_i) = \nu_B(x_i), \pi_A(x_i) = \pi_B(x_i)\). Therefore, the new distance satisfies the property that \(d(A, B) = 0\) if and only if \(A = B\).

**Proof (P3):** For \(d_n(B, A)\), we have

\[
\bar{m}_i M (\bar{m}_i M)^T
= \mu_B(x_i)^2 + u_B(x_i)^2 + \frac{1}{4}\pi_B(x_i)^2 + \mu_A(x_i)^2 + u_A(x_i)^2
\]

\[
+ \frac{1}{4}\pi_A(x_i)^2 - 2\mu_B(x_i)\mu_A(x_i) - 2\nu_B(x_i)\nu_A(x_i)
\]

\[
- \frac{1}{2}\pi_B(x_i)\pi_A(x_i) = \mu_A(x_i)^2 + u_A(x_i)^2
\]

\[
+ \frac{1}{4}\pi_A(x_i)^2 + \mu_B(x_i)^2 + u_B(x_i)^2 + \frac{1}{4}\pi_B(x_i)^2
\]

\[
- 2\mu_A(x_i)\mu_B(x_i) - 2\nu_A(x_i)\nu_B(x_i) - \frac{1}{4}\pi_A(x_i)\pi_B(x_i)
\]

So it is proved that the new distance satisfies the property that \(d(A, B) = d(B, A)\).

**Proof (P4):** Let A, B, and C be three IFSs on X = \{x_1, x_2, \ldots, x_n\}, as A \subseteq B \subseteq C, so \(\mu_A(x_i) \leq \mu_B(x_i) \leq \mu_C(x_i)\) and \(\nu_A(x_i) \geq \nu_B(x_i) \geq \nu_C(x_i)\). Because the value of \(\pi(x_i)\) is uncertain, so discuss the distance in different situations. The paper will give the proof that the new distance satisfies P4 when the IFSs have equal hesitation index. The proof of other cases is similar to this proof, so omitted. Let the constant k satisfies the condition \(0 \leq k \leq 1, \) and \(\pi_A(x_i) = \pi_B(x_i) = \pi_C(x_i) = 1 - k\).

Then

\[
d_n(A, B)
= \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\mu_A(x_i) - \mu_B(x_i))^2 + (\nu_A(x_i) - \nu_B(x_i))^2 + \frac{1}{4}(1 - K)^2}
\]

and

\[
d_n(A, C)
= \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\mu_A(x_i) - \mu_C(x_i))^2 + (\nu_A(x_i) - \nu_C(x_i))^2 + \frac{1}{4}(1 - K)^2}
\]

First let the hesitation index equal zero, so the geometric relationship of the distance between these three intuitionistic fuzzy sets is shown in the Fig. 3.

\[
d_n(A, B)
= \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\mu_A(x_i) - \mu_B(x_i))^2 + (\nu_A(x_i) - \nu_B(x_i))^2}
\]

\[
d_n(A, C)
= \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\mu_A(x_i) - \mu_C(x_i))^2 + (\nu_A(x_i) - \nu_C(x_i))^2}
\]
According to the side length relationship of the triangle,
\[ 0 \leq d_N(A, B) \leq 1 \text{ and } 0 \leq d_N(A, C) \leq 1 \]
and
\[ \sqrt{(\mu_A^2(x_i) + u_A^2(x_i)}) + \sqrt{(\mu_B^2(x_i) + u_B^2(x_i)}) \]
\[ + \sqrt{(\mu_C^2(x_i) - \mu_B(x_i))^2 + (u_C(x_i) - u_B(x_i))^2} \]
\[ > \sqrt{(\mu_A^2(x_i) + u_A^2(x_i))} \]
\[ \sqrt{(\mu_B^2(x_i) + u_B^2(x_i))} \]
(16)

So
\[ \sqrt{(\mu_C(x_i) - \mu_B(x_i))^2 + (u_C(x_i) - u_B(x_i))^2} \]
\[ > \sqrt{(\mu_A^2(x_i) + u_A^2(x_i))} - \sqrt{(\mu_B^2(x_i) + u_B^2(x_i))} \]
(17)

so
\[ \frac{(\mu_A(x_i) - \mu_B(x_i))^2 + (u_A(x_i) - u_B(x_i))^2}{\mu_A^2(x_i) + u_A^2(x_i) + \mu_B^2(x_i) + u_B^2(x_i)} \]
\[ < \frac{(\mu_A(x_i) - \mu_C(x_i))^2 + (u_A(x_i) - u_C(x_i))^2}{\mu_A^2(x_i) + u_A^2(x_i) + \mu_C^2(x_i) + u_C^2(x_i)} \]
(18)

According to the properties of inequality
\[ \frac{(\mu_A(x_i) - \mu_B(x_i))^2 + (u_A(x_i) - u_B(x_i))^2}{\mu_A^2(x_i) + u_A^2(x_i) + \mu_B^2(x_i) + u_B^2(x_i) + 2 \times (1 - K)^2} \]
\[ < \frac{(\mu_A(x_i) - \mu_C(x_i))^2 + (u_A(x_i) - u_C(x_i))^2}{\mu_A^2(x_i) + u_A^2(x_i) + \mu_C^2(x_i) + u_C^2(x_i) + 2 \times (1 - K)^2} \]
(19)

So that proves \( d_N(A, B) < d_N(A, B) \).

It can be proved by the same way that the new distance still satisfies the P4 when these three hesitation index are not equal. So the distance proposed in this paper is an intuitive distance between IFSs A and B since it satisfies the Definition 3.

E. NORMALIZATION OF DISTANCE

Distance between IFSs are used to indicate the difference degree between the information carried by IFSs, improper normalization of distance will lead to opposite decision results. Most of the existing distance functions are normalized by the maximum distance between two fuzzy sets, although this can guarantee that the maximum distance obtained will not be greater than 1, sometimes the measurement result is unreasonable because the scaling factor is too large. Excessive normalization will make the result smaller, which will reduce the sensitivity of distance in measuring the difference between intuitionistic fuzzy sets. The new method is normalized by the sum of the squares of all the parameters in the two fuzzy sets, because \( \mu_A^2(x_i) + u_A^2(x_i) + \mu_B^2(x_i) + u_B^2(x_i) + \mu_C^2(x_i) + u_C^2(x_i) + \pi_A^2(x_i) + \pi_B^2(x_i) + \pi_C^2(x_i) \) can be less than 1, this means that the normalization factor plays the role of the amplification distance in this case, but this amplification is reasonable.

Take IFSs \( A = \{x, 0.3, 0.3\} \) and \( B = \{x, 1, 0\} \) as an example, the maximum Euclidean distance between A and other intuitionistic fuzzy sets is 0.86 while B’s is 2. The distribution range of the distance that IFS A can obtain is much smaller than IFS B, It is not appropriate to use the same criteria to define the threshold of similarity. Based on the above analysis, two reasons for using the new normalization method are summarized. First, the values of the three parameters can reflect the ambiguity of the set, and the greater the ambiguity of an intuitionistic fuzzy set, the easier it is for the set to be similar to other intuitionistic fuzzy sets, so the difference between the two intuitionistic fuzzy sets is not only related to the difference of the parameter values, but also depends on the ambiguity of each fuzzy set. In addition, when the three parameters are relatively small, the numerical difference of the parameters is relatively small. If the distance is normalized by the classical method, the distance measure is easier to get similar results, so the distribution of the distance obtained is not reasonable enough.

IV. EXAMPLES AND APPLICATIONS

The numerical examples and applications are used to illustrate that the distance proposed in this paper can effectively measure the similarity between IFSs.

A. EXAMPLES

Example 1: Suppose A and B are two intuitionistic fuzzy sets on \( X = \{x\} \), \( A = \{x, \alpha, 1 - \alpha\} \), \( B = \{x, \alpha - 0.1, 1.1 - \alpha\} \), and \( \alpha \) gradually increase from 0.1 to 1.

When the hesitant index in the fuzzy set is equal to 0, the only difference between the new distance and the Euclidean distance is the normalized method they used. Use these two
distance functions to measure the difference between the IFSs $A$ and $B$, it is proved by comparing the results that the new normalization method can make the distance value have a more reasonable distribution range. The results obtained by different methods are shown in the Tab. 1. With the increase of $\alpha$, the score function of the two IFS also change correspondingly, the score function of IFSs $A$ and $B$ are shown in the Tab. 2.

It is obvious that the distance is a fixed value while the IFSs have changed, this phenomenon is the collision of similarity which is referred above, and similarity collision has negative influence on decision making. The fact that similar collisions are easy to occur shows that Euclidean distance is not sensitive to measure differences. On the other hand, the result obtained by the new distance changes with the increase of $\alpha$ when $\alpha$ is equal to 0.5, the ambiguities of the two IFSs are relatively high, and the distance of $A$ and $B$ reaches the maximum point. This is because the new normalization method allows the change of the parameter value to affect the distance result, then the sensitivity of the distance measure is improved. The relationship between these distances and the score function of IFSs is analyzed below. When $A$ is less than 0.5, with the increase of $A$, the distance and score function between two IFSs are gradually increasing, while $|S_A(x)|$ and $|S_B(x)|$ are gradually decreasing, which means the ambiguity of the two IFSs also becomes higher, this shows that the variation trend of the results obtained by the new distance is similar to the variation trend of the ambiguity of the IFSs. Because the more uncertain the intuitionistic fuzzy set is, the more similar it is to other intuitionistic fuzzy sets. Since the new distance would be affected by the ambiguity of the intuitionistic fuzzy set, so the measurement result is more reasonable. It has been proved that the new normalization method is reasonable and effective without considering the hesitance index, the next section would be focused on comparing the differences in the handling hesitance index between different distances.

**Example 2:** Suppose $A$, $B$ and $C$ are three intuitionistic fuzzy sets on $X = \{x\}$, $A = \{(x, 0.5, 0.5)\}$, $B = \{(x, 0.1, 0.2)\}$, $C = \{(x, \alpha, 0)\}$, and $\alpha$ change gradually from 0 to 1. the results of $d(A, C)$ and $d(B, C)$ obtained by different methods are shown in the Tab. 3 and Tab. 4.

It can be seen from Fig. 5 that the results obtained by these five measurement methods are quite different. Although the

---

**TABLE 1.** The results of example 1.

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>Euclidean distance</th>
<th>New distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.1</td>
<td>0.1048</td>
</tr>
<tr>
<td>0.2</td>
<td>0.1</td>
<td>0.1155</td>
</tr>
<tr>
<td>0.3</td>
<td>0.1</td>
<td>0.1260</td>
</tr>
<tr>
<td>0.4</td>
<td>0.1</td>
<td>0.1313</td>
</tr>
<tr>
<td>0.5</td>
<td>0.1</td>
<td>0.1400</td>
</tr>
<tr>
<td>0.6</td>
<td>0.1</td>
<td>0.1400</td>
</tr>
<tr>
<td>0.7</td>
<td>0.1</td>
<td>0.1313</td>
</tr>
<tr>
<td>0.8</td>
<td>0.1</td>
<td>0.1260</td>
</tr>
<tr>
<td>0.9</td>
<td>0.1</td>
<td>0.1155</td>
</tr>
<tr>
<td>1.0</td>
<td>0.1</td>
<td>0.1048</td>
</tr>
</tbody>
</table>

**TABLE 2.** The support factor of intuitionistic fuzzy sets.

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S(A)$</td>
<td>0.8</td>
<td>0.8</td>
<td>0.6</td>
<td>0.4</td>
<td>0.2</td>
<td>0</td>
<td>0</td>
<td>0.2</td>
<td>0.4</td>
<td>0.6</td>
</tr>
<tr>
<td>$S(B)$</td>
<td>1</td>
<td>0.8</td>
<td>0.6</td>
<td>0.4</td>
<td>0.2</td>
<td>0</td>
<td>0</td>
<td>0.2</td>
<td>0.4</td>
<td>0.6</td>
</tr>
</tbody>
</table>

**TABLE 3.** The results of $d(A, C)$.

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$d_{H}(A, C)$</th>
<th>$d_{G}(A, C)$</th>
<th>$d_{H}(A, C)$</th>
<th>$d_{WX}(A, C)$</th>
<th>$d_{SG}(A, C)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.300</td>
<td>0.300</td>
<td>0.300</td>
<td>0.300</td>
<td>0.300</td>
</tr>
<tr>
<td>0.1</td>
<td>0.500</td>
<td>0.781</td>
<td>0.500</td>
<td>0.475</td>
<td>0.681</td>
</tr>
<tr>
<td>0.2</td>
<td>0.300</td>
<td>0.700</td>
<td>0.500</td>
<td>0.450</td>
<td>0.651</td>
</tr>
<tr>
<td>0.3</td>
<td>0.500</td>
<td>0.900</td>
<td>0.500</td>
<td>0.300</td>
<td>0.300</td>
</tr>
<tr>
<td>0.4</td>
<td>0.500</td>
<td>0.624</td>
<td>0.624</td>
<td>0.425</td>
<td>0.618</td>
</tr>
<tr>
<td>0.5</td>
<td>0.400</td>
<td>0.557</td>
<td>0.500</td>
<td>0.400</td>
<td>0.386</td>
</tr>
<tr>
<td>0.6</td>
<td>0.500</td>
<td>0.458</td>
<td>0.500</td>
<td>0.400</td>
<td>0.542</td>
</tr>
<tr>
<td>0.7</td>
<td>0.500</td>
<td>0.436</td>
<td>0.500</td>
<td>0.425</td>
<td>0.538</td>
</tr>
<tr>
<td>0.8</td>
<td>0.500</td>
<td>0.436</td>
<td>0.500</td>
<td>0.450</td>
<td>0.555</td>
</tr>
<tr>
<td>0.9</td>
<td>0.500</td>
<td>0.458</td>
<td>0.500</td>
<td>0.475</td>
<td>0.559</td>
</tr>
<tr>
<td>1.0</td>
<td>0.500</td>
<td>0.500</td>
<td>0.500</td>
<td>0.500</td>
<td>0.577</td>
</tr>
</tbody>
</table>

**TABLE 4.** The results of $d(B, C)$.

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$d_{H}(B, C)$</th>
<th>$d_{G}(B, C)$</th>
<th>$d_{H}(B, C)$</th>
<th>$d_{WX}(B, C)$</th>
<th>$d_{SG}(B, C)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.300</td>
<td>0.265</td>
<td>0.200</td>
<td>0.175</td>
<td>0.300</td>
</tr>
<tr>
<td>0.1</td>
<td>0.200</td>
<td>0.200</td>
<td>0.200</td>
<td>0.150</td>
<td>0.242</td>
</tr>
<tr>
<td>0.2</td>
<td>0.200</td>
<td>0.173</td>
<td>0.200</td>
<td>0.174</td>
<td>0.222</td>
</tr>
<tr>
<td>0.3</td>
<td>0.200</td>
<td>0.200</td>
<td>0.200</td>
<td>0.200</td>
<td>0.267</td>
</tr>
<tr>
<td>0.4</td>
<td>0.300</td>
<td>0.285</td>
<td>0.300</td>
<td>0.275</td>
<td>0.363</td>
</tr>
<tr>
<td>0.5</td>
<td>0.400</td>
<td>0.346</td>
<td>0.400</td>
<td>0.300</td>
<td>0.480</td>
</tr>
<tr>
<td>0.6</td>
<td>0.500</td>
<td>0.436</td>
<td>0.500</td>
<td>0.425</td>
<td>0.599</td>
</tr>
<tr>
<td>0.7</td>
<td>0.600</td>
<td>0.529</td>
<td>0.600</td>
<td>0.500</td>
<td>0.707</td>
</tr>
<tr>
<td>0.8</td>
<td>0.700</td>
<td>0.624</td>
<td>0.700</td>
<td>0.575</td>
<td>0.799</td>
</tr>
<tr>
<td>0.9</td>
<td>0.800</td>
<td>0.721</td>
<td>0.800</td>
<td>0.650</td>
<td>0.891</td>
</tr>
<tr>
<td>1.0</td>
<td>0.900</td>
<td>0.818</td>
<td>0.900</td>
<td>0.725</td>
<td>0.933</td>
</tr>
</tbody>
</table>
hesitance index of the IFS A is equal to 0, but \( \mu_A(x_i) = \upsilon_A(x_i) = 0.5 \), whether A belongs to B or can not be judged based on this IFS, so the ambiguity of IFSs A is relatively high. The ambiguity of the IFS C changes with the increase of \( \alpha \), therefore the similarity between IFSs A and C should also be changed. But the result obtained by Hamming distance and Hausdorff’s distance is a fixed value, and the result obtained by Wang and Xin’s distance is symmetric with respect to \( \alpha = 1.5 \), it is obvious these results are unreasonable. The range of results obtained by Euclidean distance is relatively wide, when \( \alpha \) is relatively small, the distance between the IFSs will obviously change even if the parameters are only slightly changed. When \( \alpha \) is relatively large, the results obtained by Euclidean distance is the smallest, which is due to the normalization method it uses, the result shows that Euclidean distance is a little bit more sensitive to the difference between values. The distance obtained by the new method can not only change correspondingly with the increase of \( \alpha \), but also the distribution of distance values is relatively uniform. Therefore, the distance proposed in this paper has good sensitivity. The distances between B and C are relatively uniform. Therefore, the distance proposed in this paper has good sensitivity. The distances between B and C are almost equal. This example illustrates that the new distance can effectively measure the similarity between IFSs.

B. APPLICATIONS

Intuitionistic fuzzy sets are widely used to deal with uncertain information, in order to prove the effectiveness of the new distance in practical application, we would presents the application of IFSs in pattern recognition and medical diagnosis.

1) PATTERN RECOGNITION

Assume that there are four patterns \( A_1, A_2, A_3 \) and \( A_4 \) denoted by IFSs in the universe of discourse \( X = \{x_1, x_2, x_3, x_4\} \). The patterns are denoted as follows:

\[
A_1 = \{(x_1, 0.3, 0.5), (x_2, 0.4, 0.5), (x_3, 0.2, 0.6), (x_4, 0.6, 0.2)\}
\]

\[
A_2 = \{(x_1, 0.5, 0.4), (x_2, 0.2, 0.3), (x_3, 0.0, 0.8), (x_4, 0.8, 0.2)\}
\]

\[
A_3 = \{(x_1, 0.4, 0.3), (x_2, 0.6, 0.2), (x_3, 0.0, 0.7), (x_4, 0.5, 0.1)\}
\]

\[
A_4 = \{(x_1, 0.5, 0.4), (x_2, 0.3, 0.5), (x_3, 0.2, 0.8), (x_4, 0.5, 0.2)\}
\]

It is aimed to classify an unknown pattern represented by an IFSs \( B \) into one of the patterns \( A_1, A_2, A_3 \) or \( A_4 \). The IFS B is shown as follows:

\[
B = \{(x_1, 0.4, 0.3), (x_2, 0.3, 0.4), (x_3, 0.1, 0.7), (x_4, 0.7, 0.2)\}
\]

The results obtained by different distances are shown in the Tab. 5. Because the four known patterns are very similar, so the measurement method should be sensitive enough in order to correctly identify the pattern of B. Based on analysis in Tab. 5, it can be seen that that Hamming distance and Hausdorff distance can only roughly measure similarity, and it is easy to fail in pattern recognition. Although the distance values obtained by different methods are different, the results obtained from other measures are the same except Hamming distance, the unrecognized pattern B should be classified into pattern \( A_2 \). This pattern recognition example illustrates that the new distance can effectively measure the difference between IFSs, at the same time, it can measure the difference between intuitionistic fuzzy sets more reasonably and sensitively.

2) MEDICAL DIAGNOSIS

[65]–[69] Suppose four patients Jack, Bob, Joe, Bill in a hospital, let \( P = \{\text{Jack, Bob, Joe, Bill}\} \). Their symptoms are represented as \( S = \{\text{Temperature, Headache, Stomach pain, Cough, Chest pain}\} \). Let the set of Diagnosis be \( D = \{\text{Viral fever, Malaria, Typhoid, Stomach problem, Chest problem}\} \). The intuitionistic fuzzy relation \( P \rightarrow S \) is given as in Tab. 6, and the intuitionistic fuzzy relation \( S \rightarrow D \) is given as in Tab. 7. In Tab. 8, the distance degree between patients is presented. In Tab. 9, the diagnosis results for this case obtained by different methods have been presented.
According to the results shown in the Tab. 9, Jack suffers from Malaria, Bob suffers from Stomach problem, and Joe suffers from Typhoid. Four out of seven methods indicate that Bill suffers from Malaria, and other methods indicate that Bill suffers from Viral Fever. It is hard to tell whether Bill suffers from Malaria or Viral Fever, because the symptoms of the two diseases are similar. The proposed distance provides the same results obtained in [67] and [70].

V. CONCLUSION

A new distance of IFSs is proposed by this paper, which not only considers the numerical difference between IFSs, but also the characteristics of the fuzziness information carried by IFSs. The proposed distance satisfies all the property requirements of intuitive distance. In addition, the square sum of the parameters is used to normalize the distance. This work indicated that the information carried by the hesitation index cannot be ignored, and the normalization method used for the distance also affects the rationality of the results. Intuitionistic fuzzy sets can better measure the degree of ambiguity because of the hesitance index. In order to better solve the problem of intuitionistic fuzzy sets, the characteristics of the three parameters and their relations should be reasonably analyzed. In the future, we would like to generalize the new distance and apply it to areas that describe vagueness and uncertainty.

### REFERENCES


CUIPING CHENG is currently pursuing the master’s degree with the School of Computer and Information Science, Southwest University, China. Her research interests include information fusion and evidence theory.

FUYUAN XIAO received the D.E. degree from the Graduate School of Science and Technology, Kumamoto University, Japan, in 2014. Since 2014, she has been with the School of Computer and Information Science, Southwest University, China, where she is currently an Associate Professor. She has published over 30 papers in the prestigious journals and conferences, including *Information Fusion*, *Applied Soft Computing*, *IEEE Access*, *Engineering Applications of Artificial Intelligence*, *the IEICE Transactions on Information and Systems*, *Artificial Intelligence in Medicine*. Her research interests include information fusion, uncertain information modeling and processing, and complex event processing. She serves as a Reviewer of the prestigious journals, such as *Information Sciences*, *Knowledge-Based Systems*, *Engineering Applications of Artificial Intelligence*, *Future Generation Computer Systems*, *IEEE Access*, and *Artificial Intelligence in Medicine*.

ZEHONG CAO (M’17) received the B.S. degree from The Chinese University of Hong Kong, the M.S. degree from Northeastern University, and the dual Ph.D. degrees in information technology from UTS and in electrical and control engineering from National Chiao Tung University (NCTU), Taiwan. He is currently a Lecturer (Assistant Professor) with the Discipline of Information and Communication Technology (ICT), School of Technology, Environments and Design, College of Sciences and Engineering, University of Tasmania, Hobart, Australia, and an Adjunct Fellow with the School of Computer Science, Faculty of Engineering and IT, University of Technology Sydney (UTS), Australia. His research interests include fuzzy sets and systems, fuzzy neural networks, game artificial intelligence, deep reinforcement learning, brain–computer interface, and biosignal processing. He received the UTS Centre for Artificial Intelligence Best Paper Award, in 2017, the UTS Faculty of Engineering and IT Publication Award, in 2017, the UTS President Scholarship, in 2015, and the NCTU & Songshanhu Scholarship, in 2013. He has a string of successful over 30 publications among the most respected journals, including *Nature Scientific Data*, the *IEEE Transactions on Fuzzy Systems*, the *IEEE Transactions on Neural Networks and Learning Systems*, the *IEEE Transactions on Cybernetics*, the *IEEE Systems, Man, and Cybernetics Magazine*, the *IEEE Transactions on Biomedical Engineering*, *IEEE Access*, the *International Journal of Neural Systems*, *Information Science, Neurocomputing*, and *Neural Computing and Applications*. He has been serving as an Associate Editor for *IEEE Access*, since 2018, and the *Journal of Intelligent and Fuzzy Systems*, since 2019, and the Guest Editor of *Swarm and Evolutionary Computation*, in 2019, *Neurocomputing*, in 2018, and the *International Journal of Distributed Sensor Networks*, in 2018.

* * *