PAPER • OPEN ACCESS

# Distance of Single-Valued Neutrosophic Set and Its Application in Pattern Recognition

To cite this article: Tongjuan Wang 2021 *J. Phys.: Conf. Ser.* **2025** 012019





# Distance of Single-Valued Neutrosophic Set and Its Application in Pattern Recognition

**Tongjuan Wang**

Department of Software Engineering, Shandong Polytechnic College, Jining, Shandong, China
Email: 2219185045@qq.com

**Abstract.** Single-valued neutrosophic set (SVNS) is an extension of fuzzy set, which combines the truth, indeterminacy, and falsity information. The measurement of distance of single-valued neutrosophic set will bring new ideas to pattern recognition. This paper introduces the distance definitions and properties of SVNS, and proposes the improved distance definition of single-valued neutrosophic set based on decision maker attitude towards indeterminacy information and applied them to pattern recognition. By assigning a value to decision-maker's attitude towards indeterminacy information, this attitude comes from the historical experience and risk preference etc. The distance calculated in this way can often get a matching result closer to the reality. The results show that the improved distance formula of SVNSs is more effective and more practical for pattern recognition.
**Keywords.** Single-valued neutrosophic set; pattern recognition; fuzzy pattern recognition.

## 1. Introduction

Pattern recognition is to classify a given object into several existing samples according to its characteristics, which essentially depends on mathematical methods and computer aid. It is an important research field of artificial intelligence. Pattern recognition mainly studies image processing and computer vision, language information processingetc., and studies the mechanism and effective calculation method of human pattern recognition.The main methods of pattern recognition are divided into three categories: decision theory, syntactic and statistical pattern recognition.According to the characteristics or attributes of the research object, we use certain statistical and decision-making methods to determine its category, and make the results of classification and recognition conform to the reality as much as possible.Pattern recognition technology has been widely used in the following important fields: fingerprint recognition, character recognition, speech recognition, medical diagnosis, automatic driving target detection, military target recognition, etc.

Fuzzy pattern recognition is a means and technology to simulate human processing information. It can not only deal with pattern recognition problems, but also deal with fuzzy information better. There are usually two kinds of methods to deal with common fuzzy pattern problems: one is the principle of maximum membership degree, the other is the principle of closeness degree. With the enrichment of fuzzy sets and the diversification of fuzzy pattern processing methods, the methods to deal with problems are also more changeable, but most of them are based on the above two principles. In real life, the pattern of some categories of things is clear, clear and definite. On the contrary, some categories of things are not clear, vague and uncertain. Due to the fuzziness of objective things and the fuzziness of people's cognition of things, the classical pattern recognition methods cannot effectively deal with this problem. Therefore, we try to use the fuzzy set to deal with the fuzzy information in pattern recognition. For the pattern recognition problem with "fuzzy pattern", we can use the "fuzzy recognition method" to deal with it. Fuzzy pattern recognition methods are often divided into direct methods and indirect methods. The direct method of fuzzy recognition can be divided into three steps: first, identify and extract the feature index of the object, Secondly, the membership function group of fuzzy pattern is constructed, Third, identification and judgment. The indirect steps of pattern recognition include: first, identifying and extracting the feature index of the object, Secondly, the membership function group of fuzzy pattern is constructed, Thirdly, identify the membership function with the identified object, Fourth, calculate the closeness between the object to be identified and each object in the set, Fifthly, it is classified according to the principle of proximity.

In the classification of some patterns to be recognized, it is usually based on the distance and similarity between them. Generally speaking, the greater the similarity and the smaller the distance between the two patterns, the more likely they are to belong to the same category. On the contrary, the smaller the similarity and the greater the distance between the two patterns, the less likely they are to belong to the same category. The concepts of distance and similarity are often used as criteria for statistical pattern recognition and cluster analysis.

Since 1965, Professor Zadeh [1] put forward the concept of fuzzy set (FS). Scholars began to introduce the concept of FS into the field of pattern recognition. Inheriting the basic idea of fuzzy set, until now, many variants of fuzzy sets have been developed. For example, intuitionistic fuzzy sets (IFSs)proposed by Atanassov [2], interval valued intuitionistic fuzzy sets (IVIFSs) extended on IFSs introduced by Atanassov and Gargov [3], hesitant fuzzy sets (HFSs) introduced by Torra [4], neutrosophic sets (NSs) defined by Smarandache which reflect combined the information including truth, indeterminate and false informationsimultaneously.Since is very difficult for NSs to be applied and popularized in the real-world environments, so Wang et al. gave the definition of SVNS to overcome this defect [5]. SVNS can have better flexibility and practicability. In addition, it can also be effectively applied to pattern recognition.

As important numerical concepts in fuzzy set theory, distance degree [6, 7] and similarity degree [8, 9] is widely used in pattern recognition, decision-making [10-15], machine learning, market forecasting and other fields. Because of the importance of distance in fuzzy sets, many scholars extend their concepts to SVNSs. However, it is unreasonable and inconsistent with the reality to treat the three components of SVNS equality and directly copy all kinds of distance calculation formulas and similarity calculation formulas into the calculation of distance and similarity of SVNS. Since the three components of SVNS can be roughly divided into definite information part and uncertain information part, if we integrate uncertain information into the distance calculation formula, we must consider its particularity and cannot equate it with the position of definite information. According to historical experience and risk preference, the uncertain information is selectively incorporated into the calculation formula of distance. Because many subjective factors are inevitably integrated into the date of pattern recognition, we should deal with this kind of information and treat the distance components obtained from the determined and uncertain parts differently, which will be more helpful to obtain the pattern recognition results in line with the real world.

## 2. Preliminaries

### *2.1. Definition of SVNSs*

**Definition 1.** [5] Let $U$ be a nonempty fixed set. A single valued neutrosophic set $A$ in $U$ is defined as an object of the following form.

$$A = \{< x, T_A(x), I_A(x), F_A(x) > | \, x \in U\}$$

where $T_A : U \to [0,1]$ is called the degree of truth-membership of the element $x \in U$ to $A$, $I_A : U \to [0,1]$ is the degree of indeterminacy-membership of the element $x \in U$ to $A$, $F_A : U \to [0,1]$ is called the degree of falsity-membership. They satisfy $0 \le T_A(x) + I_A(x) + F_A(x) \le 3$ for all $x \in U$. The family of all single valued neutrosophic sets in $U$ is denoted by $SNN(U)$. For convenience, a single valued neutrosophic number (SVNN) is represented by $\alpha$=<a,b,c>, where $a,b,c\in[0,1]$ and $a,b,c \leq 3$.

**Definition 2.** [5] Let $X$ be a space of points (objects), with a generic element in $X$ denoted by $x$. A NS $A$ in $X$ is characterized by $T_A(x)$, $I_A(x)$ and $F_A(x)$, which are singleton subintervals/subsets in the real standard [0,1], that is $T_A(x): X \to [0,1]$, $I_A(x): X \to [0,1]$, and $F_A(x): X \to [0,1]$.

Then, a simplification of $A$ is denoted by

$$A = \{< x, T_A(x), I_A(x), F_A(x) >| x \in X\}$$

which is called a SVNS. It is a subclass of NSs.

*2.2. Operational Rules of SVNSs*

The operational rules of SVNSs are defined as follows.

**Definition 3.** [5] Let $x = (T_1^x, I_1^x, F_1^x)$, $y = (T_2^y, I_2^y, F_2^y)$ be two SVNNs, then operational rules of single-valued neutrosophic numbers can be defined as:

(1) $\overline{x} = (F_1^x, 1 - I_1^x, T_1^x)$

(2) $x \oplus y = (T_1^x + T_2^y - T_1^x T_2^y, I_1^x I_2^y, F_1^x F_2^y)$

(3) $x \otimes y = (T_1^x T_2^y, I_1^x + I_2^y - I_1^x I_2^y, F_1^x + F_2^y - F_1^x F_2^y)$

(4) $\lambda x = (1-(1-T_1^x)^\lambda, (I_1^x)^\lambda, (F_1^x)^\lambda), \lambda > 0$

(5) $x^\lambda = ((T_1^x)^\lambda, 1-(1-I_1^x)^\lambda, 1-(1-F_1^x)^\lambda)$

**Definition 4.** [5] Let $x = (T_1^x, I_1^x, F_1^x)$, $y = (T_2^y, I_2^y, F_2^y)$ be two SVNNs, and its operation rules satisfy the following operation relations:

(1) $x \oplus y = y \oplus x$

(2) $x \otimes y = y \otimes x$

(3) $\lambda(x \oplus y) = \lambda x \oplus \lambda y, \lambda \geq 0$

(4) $\lambda_1 x \oplus \lambda_2 x = (\lambda_1 + \lambda_2)x, \lambda_1, \lambda_2 \geq 0$

(5) $x^\lambda \otimes y^\lambda = (x \otimes y)^\lambda, \lambda \geq 0$

(6) $x^{\lambda_1} \otimes x^{\lambda_2} = x^{\lambda_1 + \lambda_2}, \lambda_1, \lambda_2 \geq 0$

*2.3. Definitions of Distance for SVNSs*

Suppose $A = (x_1, x_2, \cdots, x_n)$ and $B = (y_1, y_2, \cdots, y_n)$ are two single-valued neutrosophic sets, $x_i$ and $y_j$ $(i, j = 1, 2. \cdots, n)$ are two SVNNs.

$$x_i = [T_i^x, I_i^x, F_i^x], y_j = [T_j^y, I_j^y, F_j^y].$$

**Definition 5.** [5] A SVNS $A$ is contained in the other SVNS $B$, $A \subseteq B$, if and only if $T_i^x \leq T_j^y$, $I_i^x \geq I_j^y$, $F_i^x \geq F_j^y$ for any x in $X$.

(1) The Hamming distance between two SVNSs $A$ and $B$ [7].

$$d_{AB} = \frac{1}{3n} \sum_{i=1}^{n} (| T_i^x - T_i^y | + | I_i^x - I_i^y | + | F_i^x - F_i^y |) \tag{1}$$

(2) The Educlidean distance between two SVNSs $A$ and $B$ [7].

$$d_{AB} = \sqrt{\frac{1}{3n} \sum_{i=1}^{n} ((T_i^x - T_i^y)^2 + (I_i^x - I_i^y)^2 + (F_i^x - F_i^y)^2)} \tag{2}$$

(3) The Chebyshev distance between two SVNSs $A$ and $B$ .

$$d_{AB}=\frac{1}{3n}\max_{i}(|T_i^x-T_i^y|+|I_i^x-I_i^y|+|F_i^x-F_i^y|) \tag{3}$$

(4) The Minkowski distance between two SVNSs $A$ and $B$ [m].

$$d_{AB}=\sqrt[p]{\frac{1}{3\mathrm{n}}\sum_{i=1}^{n}(|T_i^x-T_i^y|^p+|I_i^x-I_i^y|^p+|F_i^x-F_i^y|^p)} \tag{4}$$

When p is 1, Minkowski degenerates to Hamming distance. When p is 2, Minkowski degenerates to Euclidean distance. In fact, when the above formula is adopted, the premise is that there are compensations between true value, hesitant value and false value. This premise is embodied in the definitions of Hamming distance, Educlidean distance and Minkowski distance.

For equation (4), calculate the distance between $O$, $R$, $T$,

$$O=(0.1,0.15,0.2),\ R=(0.35,0.25,0.25),T=(0.2,0.4,0.25)$$

According to the equation (4), when p=1, $d_{OR}=d_{OT}=0.133$,when p=2, $d_{OR}=d_{OT}=0.158$.The same can be proved: $d_{OR}=d_{OT}$ . But in fact, it's not in line with the reality. Obviously, people generally think that $d_{OT}>d_{OR}$.

## 3. Improved Definition of Distance of SVNSs

**Definition 6.** The measurement of distance mainly gathers three kinds of information: the gap of truth information, the gap of falsity information and the gap of indeterminacy information. We regard the three kinds of information as equal, but in fact, we value truth information and falsity information more than indeterminacy information. Based on the above reasons, we improve the distance formula and give the distance formula combined with the value orientation of decision makers as follows.

Suppose $R=(x_1,x_2,\cdots,x_n)$ and $T=(y_1,y_2,\cdots,y_n)$ are two SVNSs, $x_i$ and $y_j$ $(i,j=1,2.\cdots,n)$ are two SVNNs,

$$x_i=[T_i^x,I_i^x,F_i^x],\ y_j=[T_j^y,I_j^y,F_j^y]$$

The improved distance of single-valued neutrosophic weighted distance measure is defined by

$$d_{RT}=\left[\frac{1}{3}\sum_{i=1}^{n}w_i(|T_i^x-T_i^y|^p+(\varphi|I_i^x-I_i^y|)^p+|F_i^x-F_i^y|^p)\right]^{1/p} \tag{5}$$

where $p>0$, $w_i\in[0,1]$ and $\sum_{i=1}^{n}w_i=1$, $\varphi\in[0,1]$.

When the decision-maker thinks that the indeterminacy information is as important as the truth information and the falsity value information, then $\varphi=1$; when the decision-maker thinks that the indeterminacy information has no value at all, then $\varphi=0$ .The value of $\varphi$ represents the decision-maker's value orientation of indeterminacy information.

**Proposition 1.** Let R, T are two SVNSs, then the distance between R and T $d(R,T)$ satisfies:

(d1) $0\le d(R,T)\le 1$;

(d2) $d(R,T)=0$, if and only if $R=T$;

(d3) $d(R,T)=d(T,R)$

(d4) If $R\subseteq T\subseteq O$, $O=(\mathrm{o}_1,\mathrm{o}_2,\cdots,\mathrm{o}_\mathrm{n})$ is an SVNS in $X$, $\mathrm{o}_k=[T_k^o,I_k^o,F_k^o]$, then

$$d(R,T)\le d(R,O)\ ,d(T,O)\le d(R,O)$$

**Proof.** It is easy to proof that $d(R,T)$ satisfies the properties (d1), (d2), (d3). So we only need to prove (d4).

Let $R \subseteq T \subseteq O$, then

$$T_i^r \le T_j^t \le T_k^o,\ I_i^r \ge I_j^t \ge I_k^o, F_i^r \ge F_j^t \ge F_k^o$$
$$|T_i^r - T_i^t|^p \le T_i^r - T_k^o|^p$$
$$|I_i^r - I_i^t|^p \le |I_i^r - I_k^o|^p$$
$$|F_i^r - F_i^t|^p \le |F_i^r - F_k^o|^p$$

Hence,

$$|T_i^r - T_i^t|^p + (\varphi|I_i^r - I_i^t|)^p + |F_i^r - F_i^t|^p \le$$
$$|T_i^r - T_k^o|^p + (\varphi|I_i^r - I_k^o|)^p + |F_i^r - F_k^o|^p$$

So we can obtain the result, $d(R,T) \le d(R,O)$.

The same can be proved, $d(T,O) \le d(R,O)$.

The minimum distance degree and the maximum similarity principle of SVNSs. The smaller the distance, the closer they are, or the more similar they are, the closer we are.

## 4. Numerical Example

The following examples illustrate the validity of the distance formula given in this paper.

**Example 1.** It is assumed that there are two modes: single-valued neutrosophic sets are $R = (0.75, 0.6, 0.2)$, $T = (0.6, 0.2, 0.1)$. The mode to be identified is $O = (0.6, 0.6, 0.25)$.

Let $w_i = 1/n$, $p = 2$, $\varphi = 1$, then according to the equation (5), $C$ belongs to $A$. To illustrate the distance proposed in this paper.

The validity of the formula can be obtained by substituting the above data into equation (5):
$d(R,O) = 0.158$, $d(T,O) = 0.427$.

Hence,

$$d(R,\mathrm{O}) < d(T,O)$$

Different values of $\varphi$ will affect the distance calculation results, and eventually lead to different distance comparison results. Table 1 shows the calculation results of distance $d(R,O)$ and $d(T,O)$ under different values of $\varphi$.

**Table 1.** The effect of $\varphi$ value on $d(R,O)$ and $d(T,O)$.

| **Order** | **The value of** $\varphi$ | $d(R,O)$ | $d(T,O)$ |
|---|---|---|---|
| 1 | 1 | 0.158 | 0.427 |
| 2 | 0.9 | 0.158 | 0.39 |
| 3 | 0.8 | 0.158 | 0.353 |
| 4 | 0.7 | 0.158 | 0.317 |
| 5 | 0.6 | 0.158 | 0.283 |
| 6 | 0.5 | 0.158 | 0.25 |
| 7 | 0.4 | 0.158 | 0.219 |
| 8 | 0.3 | 0.158 | 0.192 |
| 9 | 0.2 | 0.158 | 0.17 |
| 10 | 0.1 | 0.158 | 0.155 |
| 11 | 0 | 0.158 | 0.15 |

According to the evaluation and historical experience data of decision makers, the value of $\varphi$ can be set. In general, the value of $\varphi$ is set to [0.5,1]. Here, based on historical experience data, we set $\varphi$=0.7 .

$$d(R,\mathrm{O}) = 0.158\ ,\ d(T,O) = 0.317$$

Hence,

$$d(R,O) < d(T,O)$$

So we can judge that $O$ belongs to $R$ , which shows that the distance formula proposed in this paper is effective in the process of pattern recognition. The value of $\varphi$ can affect the result of distance comparison.

**Example 2.** It is assumed that there are two modes: single-valued neutrosophic sets are

$$R = (0.2, 0.4, 0.25)\ ,\ T = (0.35, 0.25, 0.25)$$

The mode to be identified is $O = (0.1, 0.15, 0.2)$.

Let $w_i = 1/n$ , $p = 2$ , $\varphi$=1, then according to the equation (5), the values of $d(R,O)$ and $d(T,O)$ are as follows:

$$d(R,O) = 0.273\ ,\ d(T,O) = 0.273$$

But when $\varphi$ is set different values, the result is completely different. Table 2 shows the value of $\varphi$ influence the values of $d(R,O)$ and $d(T,\mathrm{O})$ .

According to table 2, if we set $\varphi$=0.9 , the comparison result is completely different, $d(R,O) > d(T,O)$ . This comparison result is more in line with the reality, so it has more reference value.

The biggest advantage of this distance calculation formula is to consider the value orientation of indeterminacy value, and the disadvantage is to evaluate the value of $\varphi$ . Here we mainly adopt expert evaluation method, or infer from historical experience data. There are obviously some difficulties. This has some obstacles to the extension and application of the distance formula.

**Table 2.** The effect of $\varphi$ value on $d(R,O)$ and $d(T,O)$ .

| **Order** | **The value of** $\varphi$ | $d(R,O)$ | $d(T,O)$ |
|---|---|---|---|
| 1 | 1 | 0.274 | 0.274 |
| 2 | 0.9 | 0.270 | 0.251 |
| 3 | 0.8 | 0.267 | 0.229 |
| 4 | 0.7 | 0.264 | 0.208 |
| 5 | 0.6 | 0.262 | 0.187 |
| 6 | 0.5 | 0.260 | 0.168 |
| 7 | 0.4 | 0.258 | 0.150 |
| 8 | 0.3 | 0.257 | 0.135 |
| 9 | 0.2 | 0.256 | 0.122 |
| 10 | 0.1 | 0.255 | 0.115 |
| 11 | 0 | 0.255 | 0.112 |

## 5. Conclusion

In this paper, considering that, we pay more attention to certain information than uncertain information, we try to introduce the value orientation of decision-makers into the distance calculation formula of SVNSs, and propose the improved distance formula of SVNSs with decision maker's attitude towards indeterminacy information, and based on this further improve the distance calculation method of

SVNSs. And we introduce an improved fuzzy pattern recognition method based on this method. The effectiveness of this method is verified by an example.In the future, we look forward to applying the improved distance formula of SVNSs to speech recognition, image recognition, text recognition, etc.

## References

[1] Zadeh L A 1965 Fuzzy sets *Inf. Control* **8** 338.
[2] Atanassov K T 1986Intuitionistic fuzzy sets *Fuzzy Sets Syst.* **20** 87.
[3] Atanassov K and Gargov G 1989 Interval valued intuitionistic fuzzy sets *Fuzzy Sets Syst*. **31** 343.
[4] Torra V 2010 Hesitant fuzzy sets *Int. J. Intell. Syst.* **25** 529.
[5] Wang H B, Smarandache F, Zhang Y Q and Sunderraman R 2012 Single valued neutrosophic sets *Multispace Multistruct* **4** 410.
[6] Huang H L 2016 New distance measure of single-valued neutrosophic sets and its application *Int. J. Intell. Syst*. **31** 1021.
[7] Ye J 2014 Clustering methods using distance-based similarity measures of single-valued neutrosophic sets *Journal of Intelligent Systems* **23** 379.
[8] Ye J 2014 Similarity measures between interval neutrosophic sets and their applications in multicriteria decision-making *J. Intell. Fuzzy Syst.* **26** 165.
[9] Ye J 2014 Vector similarity measures of simplified neutrosophic sets and their application in multicriteria decision making *Int. J. Fuzzy Syst.* **16** 204.
[10] Karaaslan F 2018 Multicriteria decision-making method based on similarity measures under single-valued neutrosophic refined and interval neutrosophic refined environments *Int. J. Intell. Syst.* **33** 928.
[11] Liu P D and Wang Y M2014 Multiple attribute decision-making method based on single-valued neutrosophic normalized weighted Bonferroni mean *Neural Comput. Appl.* **26** 2001.
[12] Peng X D and Dai J G 2018 Approaches to single-valued neutrosophic MADM based on MABAC, TOPSIS and new similarity measure with score function *Neural Comput. Appl.* **29** 939.
[13] Ye J 2013 Multicriteria decision-making method using the correlation coefficient under single-valued neutrosophic environment *Int. J. Gen. Syst.* **42** 386.
[14] Ye J 2014 Single valued neutrosophic cross-entropy for multicriteria decision making problems *Appl. Math. Model.* **38** 1170.
[15] Ye J 2014 A multicriteria decision-making method using aggregation operators for simplified neutrosophic sets *J. Intell. Fuzzy Syst*. **26** 2459.