International Journal of Control and Automation
Vol.9, No.1 (2016), pp.1-10
http://dx.doi.org/10.14257/ijca.2016.9.1.01

Hyperbox Granular Computing Based on Distance Measure

Hongbing Liu, Huaping Guo and Chang-an Wu

School of Computer and Information Technology, Xinyang Normal University, Xinyang 464000, China
liuhbing@126.com

Abstract

A bottle up hyperbox granular computing (HBGrC) is developed based on distance measure. Firstly, hyperbox granule is represented by the beginning point and the end point. Secondly, the distance measure between two hyperbox granules is defined by the beginning points and the end points. Thirdly, operations between two hyperbox granules are designed to the transformation between two hyperbox granule spaces with different granularities, HBGrC is developed by the join operator and the user-defined granularity threshold ρ on the basis of bottle up scheme. Experimental results shown that HBGrC achieved the better testing accuracies over the machine learning benchmark datasets.

***Keywords**: Hyperbox granule, distance, operation, granularity*

1. Introduction

Many researchers have worked in the granular computing (GrC) field. Zadeh identified three fundamental concepts of the human cognition process, namely, granulation, organization, and causation [1,2]. Granulation is a process that decomposes a universe into parts. Conversely, organization is a process that integrates parts into a universe by introducing operation between two granules. Causation involves the association of causes and effects. Pedrtcz computed information granules based on sets, fuzzy sets or relations, and fuzzy relations [3]. Karburlasos and his colleage use the fuzzy relation between two granules to realize the transformation between two granule spaces with different granularities[4-9]. These studies enable us to map the complexities of the world around us into simple theories.

In this paper, hyperbox granular computing is proposed based on distance measure. Firstly, two points, such as the beginning point and the end point, are used to represent the hyperbox granule, and each sample is regarded as the atomic hyperbox granule which cannot be divided. Secondly, the distance measure between two hyperbox granules is defined. Thirdly, two operations $\vee$ and $\wedge$ between two hyperbox granules are designed to the transformation between two hyperbox granule spaces with different granularities. Finally, HBGrC is formed on the basis of bottle up scheme.

The rest of this paper is presented as follows. The motivation and related work is described in Section 2. Section 3 designs hyperbox granular computing based on distance measure. The experiments are used to demonstrate HBGrC in Section 4. Section 5 summarizes the contribution of our work and presents future work plans.

2. Motivation and Related Work

In this section, the motivation for this proposed research work is presented, and some related works are discussed.

ISSN: 2005-4297 IJCA

2.1. Motivation

For GrC in the view of set theory, the granule is represented as the subset for the training set *S*. In general, distance between two non-empty sets is the minimum of the distances between any two of their respective points [10], *i.e*.

$$d(A,B) = \min_{x\in A, y\in B} d(x,y) \quad (1)$$

where $d(x,y)$ is Euclidean distance between two points. For aforementioned distance formula (1), it is suitable that intersection of set *A* and set *B* is empty set. In Figure 1, sets $A=\{x_1, x_2, x_3, x_4, x_5\}$ and $B=\{y_1, y_2, y_3, y_4, y_5, y_6\}$ are denoted by ball *A* and *B*. In Figure 1 (a) distance between *A* and *B* is the distance between point x_5 and y_6, obviously $d(A,B)$ is greater than 0. In Figure 1(b), distance between set *A* and *B* also is the distance between x_5 and y_6. If the distance between x_5 and y_6 in Figure 1(a) is equal to the distance between x_5 and y_6 in Figure 1(b), the distance $d(A,B)$ in Figure 1(a) is equal to the distance $d(A,B)$ in Figure 1(b). Obviously, the distance $d(A,B)$ in Figure 1(b) is less than Figure 1(a), but $d(A,B)$ in Figure 1(b) is equal to Figure 1(b) according to formula (1). Distance formula (1) does not reflect the real distance between two sets, and we define the distance between two sets, where sets are represented as the form of hyperbox, and form the hyperbox granular computing based the defined distance measure.

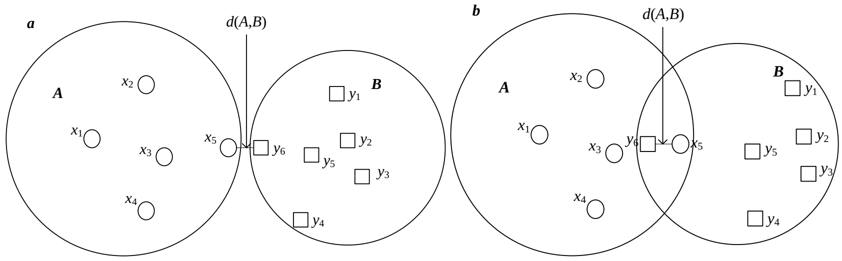

Figure 1. Distances Defined by Formula (1) between Two Sets

2.2. Related Work

GrC has been proposed and studied in many fields, including machine learning and data analysis [11-14].

Two granular structures induced by a rough set are proposed by Yao: one is a partition induced by an equivalence relationship, and the other is a covering induced by a reflexive relationship. Each equivalence class can be viewed as a granule, and each block induced by the similarity relationship is regarded as a granule. Yao also suggested the inclusion measure to form granular structures. A measure of the graded inclusion of two sets is defined as

$$\mu(A, B)=|A\cap B|/|A|$$

μ can be interpreted as the conditional probability that a randomly selected element in *A* belongs to *B*, which can be used to measure the degree to which *A* is a subset of *B*. μ can be interpreted as a fuzzy partial order relation of 2^U, and the use of a complete lattice corresponds to the lattice-based fuzzy partial order relations in the fuzzy set theory.

The difference between the granular structure proposed by Yao and GrC introduced by Kaburlasos is that the fuzzy inclusion measure in GrC is computed by the ratio of the granule to its dilation or the ratio of the erosion to the original granule.

In recent years, GrC is one of main research focus [15-17]. A notion of knowledge distance is introduced to differentiate two given knowledge structures and investigate some of its important properties [16]. This is accomplished via a near rough set framework in the approximation of a pair of disjoint sets and measurement of distances between sets using various fuzzy pseudometrics [17].

3. Hyperbox Granular Computing Based Distance Measure

For N-dimensional space, we form HBGrC in terms of the following steps. Firstly, two points called the beginning point and the end point are used to represent the hyperbox granule, and each sample is regarded as the atomic hyperbox granule which cannot be divided. Secondly, the distance measure between two hyperbox granules is defined. Thirdly, operations called join operation $\vee$ and meet operation $\wedge$ between two hyperbox granules are designed to the transformation two hyperbox granule spaces with different granularities. Finally, HBGrC is formed on the basis of bottle up scheme.

3.1. Representation and Granularity for the Hyperbox Granule

For the training set S composed of ℓ N-dimensional input vectors, two points $\boldsymbol{x}=(x_1, x_2, \ldots, x_N)$ and $\boldsymbol{y}=(y_1, y_2, \ldots, y_N)$ are used to represent the hyperbox granule. The form of the granule is HB=($\boldsymbol{x}$, $\boldsymbol{y}$, g_r), where $\boldsymbol{x}\preccurlyeq\boldsymbol{y}$. $\boldsymbol{x}\preccurlyeq\boldsymbol{y}$ is the partial order relation between two vectors and defined as follows.

$$\boldsymbol{x}\preccurlyeq\boldsymbol{y}=x_1\leq y_1\&\ x_2\leq y_2\&\ \ldots\&\ x_N\leq y_N$$

$\leq$ is the less than or equal relation between two scalars. Here, point $\boldsymbol{x}$ is called the beginning point, and $\boldsymbol{y}$ is called the end point. The granularity is the size of hyperbox granule and defined as the distance between the beginning point and the end point.

For example, in two-dimensional space, HB_1 = [0.1, 0.2, 0.4, 0.6, 0.5] represents the hyperbox granule shown in Figure 2 which has the beginning point (0.1, 0.2) and the end point (0.4, 0.6). The length of hyperbox granule equals 0.4, and its width equals 0.3. The granularity of hyperbox granule is 0.5, which is determined by the beginning point and the end point. The another example is the atomic hyperbox granule HB_2=[0.5, 0.6, 0.5, 0.6, 0] shown in Figure 2 with the granularity 0, which represents the single point (0.5, 0.6).

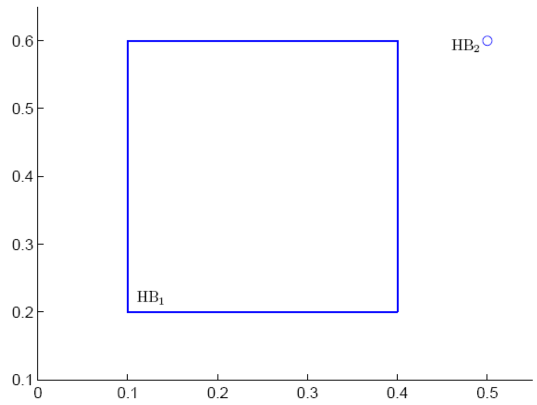

Figure 2. Hypergranules in 2-dimensional Space

3.2. Distance Measure

Distance is a numerical description of how far apart objects are. Distance between two hyperbox granules is the measure of farness between two objects, such as hyperbox granules. In analytic geometry, the distance between two points of the xy-plane can be found using the distance formula. In the Euclidean space R^N, the distance between two points is usually given by the Euclidean distance. In mathematics, in particular geometry, a distance function on a given set M is a function d: $M\times M \rightarrow R$, where R denotes the set of real numbers. Similarly, in granule space induced the hyperbox granules, we define the distance between two hyperbox granules $HB_1=(Bp_1,Ep_1,g_1)$ and $HB_2=(Bp_2,Ep_2,g_2)$ as follows.

Firstly, the distance between point P and hyperbox granule HB is defined as

$$D(\mathrm{P},\mathrm{HB})=d(\mathrm{P},\mathrm{Bp})+d(\mathrm{P},\mathrm{Ep})-d(\mathrm{Bp},\mathrm{Ep}) \tag{1}$$

where Bp is the beginning point and denoted as Bp=$(x_1,x_2,\dots,x_N)$, Ep is the end point and denoted as Ep=$(y_1,y_2,\dots,y_N)$, $d(.,.)$ is the Manhattan distance between two points.

We explain the distance between point and hyperbox granule HB in 2-dimensional space. For HB = [0.1 0.2 0.4 0.3 0.316] and the point P(0.3,0.4), d(P,Bp)=0.4, d(P,Ep)=0.2, d(Bp,Ep)=0.4, D(P,HB)=0.2. The location of P and HB is shown in Figure 3.

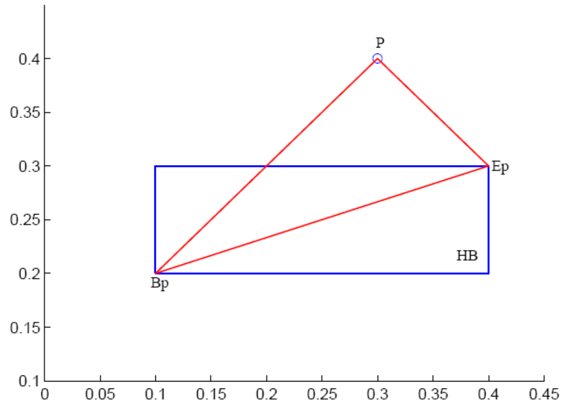

Figure 3. Distance between a Point and a Hyperbox Granule

Suppose P=$(p_1,p_2,\dots,p_N)$,

$D(\mathrm{P},\mathrm{HB})=d(\mathrm{P},\mathrm{Bp})+d(\mathrm{P},\mathrm{Ep})-d(\mathrm{Bp},\mathrm{Ep})$

$=|p_1-x_1|+|p_2-x_2|+\dots+|p_N-x_N|+|y_1-p_1|+|y_2-p_2|+\dots+|y_N-p_N|-(|y_1-x_1|+|y_2-x_2|+\dots+|y_N-x_N|)$

$=(|p_1-x_1|+|p_1-y_1|-|y_1-x_1|)+(|p_2-x_2|+|p_2-y_2|-|y_2-x_2|)+\dots+(|p_N-x_N|+|p_N-y_N|-|y_N-x_N|)$

≥ 0

Theorem 1. P is included in HB if and only if D(P,HB)=0

Proof. Suppose Bp=$(x_1,x_2,\dots,x_N)$, Ep=$(y_1,y_2,\dots,y_N)$, P=$(p_1,p_2,\dots,p_N)$.

If P is included in HB, Bp≼P and P≼Ep, $d(\mathrm{P},\mathrm{Bp})=p_1-x_1+p_2-x_2+\dots+p_N-x_N$.

$d(\mathrm{P},\mathrm{Ep})=y_1-p_1+y_2-p_2+\dots+y_N-p_N$, $d(\mathrm{P},\mathrm{Bp})+d(\mathrm{P},\mathrm{Ep})$

$= p_1-x_1+p_2-x_2+\dots+p_N-x_N+y_1-p_1+y_2-p_2+\dots+y_N-p_N$

$=y_1-x_1+y_2-x_2+\dots+y_N-x_N$

$=d(\mathrm{Bp},\mathrm{Ep})$

namely $D(\mathrm{P},\mathrm{HB})=d(\mathrm{P},\mathrm{Bp})+d(\mathrm{P},\mathrm{Ep})-d(\mathrm{Bp},\mathrm{Ep})=0$.

$D(\mathrm{P},\mathrm{HB})=d(\mathrm{P},\mathrm{Bp})+d(\mathrm{P},\mathrm{Ep})-d(\mathrm{Bp},\mathrm{Ep})$

$= |p_1-x_1|+|p_2-x_2|+\dots+|p_N-x_N|+|y_1-p_1|+|y_2-p_2|+\dots+|y_N-p_N|-(|y_1-x_1|+|y_2-x_2|+\dots+|y_N-x_N|)$

$=\Sigma^{N}_{i=1}(|y_i-p_i|+|x_i-p_i|-|y_i-x_i|)=0$

because $|y_i-p_i|+|x_i-p_i|-|y_i-x_i|\geq 0$ and $x_i\leq y_i$, $|y_i-p_i|+|x_i-p_i|-|y_i-x_i|=0$, namely $x_i\leq p_i$ and $p_i\leq y_i$. P is included in HB.

Secondly, the distance between two hyperbox granules $\mathrm{HB}_1=(\mathrm{Bp}_1, \mathrm{Ep}_1, g_1)$ and $\mathrm{HB}_2=(\mathrm{Bp}_2, \mathrm{Ep}_2,g_2)$ is defined as

$$D(\mathrm{HB}_1,\mathrm{HB}_2)=(D(\mathrm{Bp}_1,\mathrm{HB}_2)+D(\mathrm{Ep}_1,\mathrm{HB}_2))/2 \tag{2}$$

The distance between two hyperbox granule has the follow properties.

Property 1. $D(\mathrm{HB}_1,\mathrm{HB}_2)\geq 0$, $D(\mathrm{HB}_1,\mathrm{HB}_2)=0 \Leftrightarrow \mathrm{HB}_1\subseteq\mathrm{HB}_2$

Proof. Because $D(\mathrm{Bp}_1,\mathrm{HB}_2)\geq 0$ and $D(\mathrm{Ep}_1,\mathrm{HB}_2)\geq 0$,

$D(\mathrm{HB}_1,\mathrm{HB}_2)= (D(\mathrm{Bp}_1,\mathrm{HB}_2)+D(\mathrm{Ep}_1,\mathrm{HB}_2))/2\geq 0$.

If $D(\mathrm{HB}_1,\mathrm{HB}_2)=0$, $D(\mathrm{Bp}_1,\mathrm{HB}_2)=0$ and $D(\mathrm{Ep}_1,\mathrm{HB}_2)=0$. Both Bp_1 and Ep_1 are included in hyperbox granule HB_2, namely $\mathrm{HB}_1\subseteq\mathrm{HB}_2$.

If $HB_1 \subseteq HB_2$, both Bp_1 and Ep_1 are included in hyperbox granule HB_2. According to theorem 1, $D(Bp_1,HB_2)=0$ and $D(Ep_1,HB_2)=0$, namely

$D(HB_1,HB_2)= (D(Bp_1,HB_2)+D(Ep_1,HB_2))/2=0$

Property 2. $D(HB_1,HB_2) \neq D(HB_2,HB_1)$

Proof. $D(HB_1,HB_2)= (D(Bp_1,HB_2)+D(Ep_1,HB_2))/2$

$=(d(Bp_1,Bp_2)+d(Bp_1,Ep_2)-d(Bp_2,Ep_2)+d(Ep_1,Bp_2)+d(Ep_1,Ep_2)-d(Bp_2,Ep_2))/2$

$=(d(Bp_1,Bp_2)+d(Bp_1,Ep_2)+d(Ep_1,Bp_2)+d(Ep_1,Ep_2))/2-d(Bp_2,Ep_2)$

Similarly, $D(HB_1,HB_2)= (d(Bp_1,Bp_2)+d(Bp_1,Ep_2)+d(Ep_1,Bp_2)+d(Ep_1,Ep_2))/2-d(Bp_1,Ep_1)$.

Generally, $D(HB_1,HB_2) \neq D(HB_2,HB_1)$, especially, $D(HB_1,HB_2)=D(HB_2,HB_1)$ when $d(Bp_1, Ep_1)=d(Bp_2,Ep_2)$.

For 2-dimensional space, two hyperbox granules HB_1=[0.2 0.1 0.3 0.4 0.316] and HB_2=[0.25 0.15 0.4 0.5 0.381], the distance between HB_1 and HB_2 are shown in Figure 3. In the figure, $d(Bp_1,Bp_2)$ =0.1, $d(Bp_1,Ep_2)$=0.6, $d(Ep_1,Bp_2)$=0.4, $d(Ep_1,Ep_2)$=0.2, $d(Bp_1,Ep_1)$=0.3, $d(Bp_2,Ep_2)$= 0.5, $D(HB_1,HB_2)$=0.15, $D(HB_2,HB_1)$=0.35.

3.3. Operations between Two Hyperbox Granules

In N-dimensional space, any two points $x=(x_1,x_2,...,x_N)$ and $y=(y_1,y_2,...,y_N)$ can be formed a hyperbox granule HB=(Bp,Ep), where $Bp=x \wedge y=(\min\{x_1,y_1\},\min\{x_2,y_2\},...,\min\{x_N,y_N\})$ and $Ep = x \vee y=(\max\{x_1,y_1\}, \max\{x_2,y_2\},...,\max\{x_N,y_N\})$.

The join operator $\vee$ between two hyperbox granules is designed to achieve the hyperbox granule with larger granularity compared with the original hyperbox granules. For two hyperbox granules $HB_1=(Bp_1,Ep_1)$ and $HB_2=(Bp_2,Ep_2)$, the join operation $\vee$ is designed as follows.

$$HB_1 \vee HB_2=(Bp_1 \wedge Bp_2, Ep_1 \vee Ep_2) \quad (3)$$

Conversely, the meet operation $\wedge$ between two hyperbox granules is designed to obtain the hyperbox granule with the smaller granularity compared with the original hyperbox granules. The meet operation $\wedge$ is designed as follows.

$$HB_1 \wedge HB_2 = \begin{cases} (Bp_1 \vee Bp_2, Ep_1 \wedge Ep_2) & Bp_1 \vee Bp_2 \preccurlyeq Ep_1 \wedge Ep_2 \\ \emptyset & \text{Otherwise} \end{cases} \quad (4)$$

From formula (3), we can see $Bp_1 \wedge Bp_2 \preccurlyeq Bp_1$, $Bp_1 \wedge Bp_2 \preccurlyeq Bp_2$, $Bp_1 \preccurlyeq Ep_1 \vee Ep_2$, $Bp_2 \preccurlyeq Ep_1 \vee Ep_2$, $\|Bp_1 \wedge Bp_2 - Ep_1 \vee Ep_2\|_2 \geq \|Bp_1 - Ep_1\|_2$, $\|Bp_1 \wedge Bp_2 - Ep_1 \vee Ep_2\|_2 \geq \|Bp_2 - Ep_2\|_2$, namely the granularity of $HB_1 \vee HB_2$ is greater than or equal to the granularities of HB_1 and HB_2, and the operation $\vee$ induces the hyperbox granule with larger granularity compared with original granules. From formula (4), we draw the opposite conclusion that the meet operation induces the hyperbox granule with the smaller granularity compared with original granules.

We explain the operation $\vee$ and operation $\wedge$ between two hyperbox granules in 2-dimensional space. For hyperbox granules HB_1=[0.2, 0.1, 0.3, 0.4, 0.316] and HB_2=[0.25, 0.15, 0.4, 0.5, 0.381], the join hyperbox granule is HB=[0.2, 0.1, 0.4, 0.5, 0.5] shown in Figure 4, the meet hyperbox granule is HB=[0.25,0.15,0.3,0.4, 0.255] shown in Figure 5.

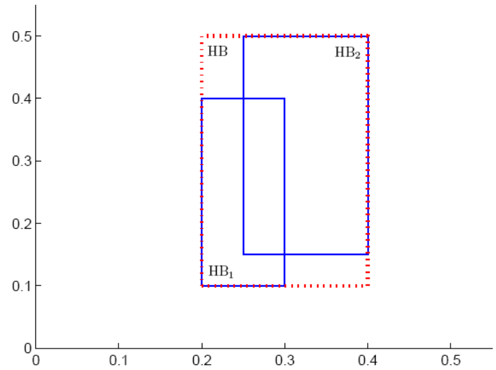

Figure 4. Two Hyperbox Granules and their join Hyperbox Granule

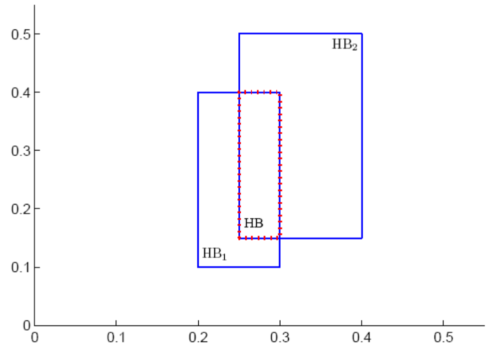

Figure 5. Two Hyperbox Granules and their Meet Hyperbox Granule

3.4. The Hyperbox Granular Computing Based on Distance Measure

For training set *S*, the granular computing classification algorithms are proposed by the following steps. Firstly, the samples are used to form the atomic granule. Secondly, the threshold of granularity is introduced to conditionally union the atomic granules by the aforementioned join operation, and the granule set is composed of all the join granules. Thirdly, if all atomic granules are included in the granules of *GS*, the join process is terminated, otherwise, the second process is continued. The algorithms include training process and testing process which are listed as follows.

Suppose the hyperbox atomic granules with the same class labels induced by *S* are g_1, g_2, g_3, g_4, g_5. The training process can be described as the following tree structure shown in Figure 6, leafs denote the atomic hyperbox granules, root denotes *GS* including its child nodes G_1, G_2, and g_3. G_1 is induced by join operation of child nodes g_1 and g_2, G_2 is the join hyperbox granule of g_4 and g_5, g_3 is the atomic hyperbox granule. The whole process of obtaining *GS* is the bottle up process.

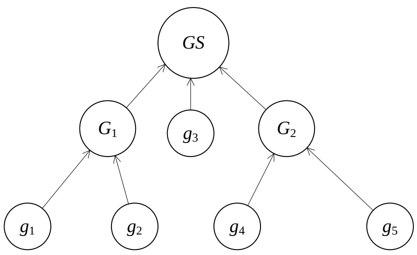

Figure 6. The Training Process of Training Set Including 5 Samples

The training algorithm and testing algorithm are described as algorithm1 and algorithm2.

Algorithm1. Training process
Input: Training set S, threshold ρ of granularity, the class number n
Output: Granule set GS, the class label lab

S1. initialize the granule set $GS=\varnothing$, $lab=\varnothing$
S2. i=1
S3. select the samples with class i, and form set X
S31. initialize the granule set $GSt=\varnothing$
S32. j=1
S33. for the jth sample x_j in X, form the corresponding atomic granule G_j
S34. k=1
S35. compute the distance d_{jk} between the atomic granule G_j and the kth granule G_k in GSt
S36. $k=k+1$
S37. find the minimal distance d_{jm}
S38. if the granularity of the join of G_j and G_m is less than or equal to ρ, the granule G_m is replace by the join, otherwise G_j is the new member of GSt.
S39. remove x_j until X is empty.
S4. $GS=GS\cup GSt$, $lab=lab\cup\{i\}$
S5. if $i=n$, output GS and class lab, otherwise $i=i+1$

Algorithm2. Testing process
Input: inputs of unknown datum x, granule set GS, the class label lab
Output: class label of x

S1. x is represented as granule g
S2. for i = 1:|GS|
S3. compute the distance d_i between g and g_i in GS
S4. find the minimal distance d_m
S5. find the corresponding class label of the g_m as the label of x

4. Experiments

We compared HBGrC with KNN by classification problems including classification in 2-dimensional space and N-dimensional space. For the selection of parameter ρ of HBGrC and parameter K of KNN, we used the stepwise refinement strategy. All the experiments are performed with an 3.2GHz Intel(R) Core(TM) i5 CPU and 8GB RAM, running Microsoft Windows7 and Matlab2008.

For the selection of parameter ρ, we used the stepwise refinement strategy. Firstly, we explored the probable optimal parameter ρ. Secondly, the optimal parameter is found near the probable optimal parameter. The maximal testing accuracy is the selection indicator of optimal parameter.

4.1. Classification in 2-dimensional Space

The spiral classification is a difficult problem to be classified and is used to evaluate the performance of classifiers. The training data are generated by the method proposed in [7]. The training set and the testing set in reference [8] are used to evaluate the performance of GrC.

For the selection of parameter, if all the training data are used to form a granule, the granularity of the granule is 1.09. Firstly, the parameter is from 1.0 to 0 with step 0.1, and the probable optimal parameter is from 0.2 to 0. Secondly, the parameter is selected from 0.2 to 0 with step 0.01 in the interval [0, 1], and the optimal parameter is 0.09, which made HBGrC achieved the best testing accuracy.

HBGrC achieved the best testing accuracy and the GS included 102 hyperbox granules when ρ=0.09. The relation between ρ and training accuracy and the relation between ρ and testing accuracy are shown in Figure 6. From the figure, we saw the classification accuracy increases when the threshold of granularity decreases. Namely, the granule set including hyperbox granules with the small granularities achieved the large classification accuracy. The training data and achieved hyperbox granules were shown in Figure 7.

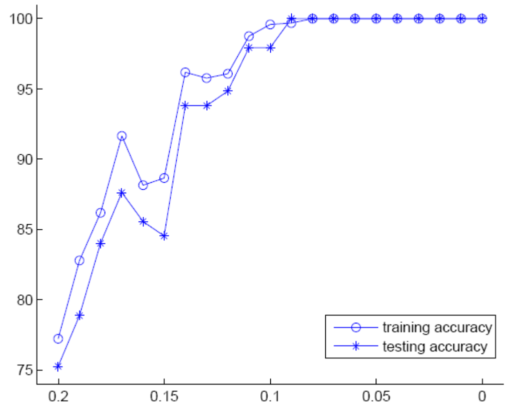

Figure 6. The Relation between ρ and the Classification Accuracies

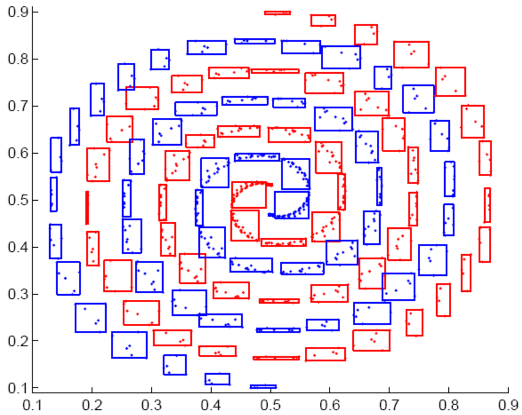

Figure 7. The Training Data and Achieved Hyperbox Granules

4.2. Classification in N-dimensional Space

Six data sets, named skin, pendigits, image, optdigits, shuttle, and madelon, are list in table 1 and selected to verify the classification performances of HBGrC in N-dimensional space.

Table 1. The Data Sets of Classification Problems in N-Dimensional Space

Datasets	Inputs	Outputs	Training size	Testing size
Skin	3	2	163371	81686
shuttle	9	7	43500	14500
pendigits	16	10	7494	3498
image	19	7	210	2100
optdigits	64	10	3823	1797
madelon	500	2	2000	600

Table 2 lists the classification performances, including the number of granules (Ng), the training accuracy, and the testing accuracy. From table 2, we can see, (1) HBGrC achieved the better testing accuracies than KNN, (2) HBGrC needed the less number of granules than KNN. Data set skin is not applicable to KNN for large number of training data. The optimal testing accuracies are 99.8839% (shuttle), 97.799% (pendigits), 87.667% (image) 97.997% (optdigits), 76.833% (madelon) by KNN algorithms. For the data set skin, owing to the large training size, the computer is out of memory. We used the selection of parameter ρ, HBGrC achieved the better or the same testing accuracies, such as 99.332% (skin), 99.91% (shuttle), 97.827% (pendigits), 92.619% (image), 97.997% (optdigits), 73.883% (madelon), and the less granule number compared with KNN.

Table 2. Classification Performance of HBGrC for N-dimensional Space

Datasets	algorithms	Parameter ρ/K	Ng	Training accuracy	Testing accuracy
Skin	HBGrC	90	91	99.59	99.332
	KNN	N/A	N/A	N/A	N/A
shuttle	HBGrC	0.001	1335	99.995	99.91
	KNN	1	43500	100	99.883
pendigits	HBGrC	60	1468	100	97.827
	KNN	3	7494	100	97.799
image	HBGrC	17	170	100	92.619
	KNN	1	210	100	87.667
optidigits	HBGrC	39.2	743	100	97.997
	KNN	1	3823	100	97.997
madelon	HBGrC	1730	271	100	73.833
	KNN	16	2000	100	76.833

5. Conclusion

The hyperbox granular computing classification algorithms are proposed based on distance measures in the paper. Firstly, a training datum is represented as an atomic hyperbox granule. Secondly, the distance measure between two hyperbox granules is form based on the beginning points and the end points. Thirdly, the training process is constructed based on the join operator and the user-defined threshold of granularity jointly. Finally, the proposed granular computing classification algorithms are demonstrated by the dataset selected from machine learning benchmark datasets. HBGrC is affected by the sequence of the training data the same as the other granular computing. The distance measure defined in the paper does not satisfy the properties, such as the symmetrical characteristic. For the future work, we will focus on the novel distance measure between two hyperbox granules.

Acknowledgements

This work was supported in part by the Natural Science Foundation of China (Grant No. 61170202, 61202287) and Natural Science Foundation of Henan (No. 132300410421, 132300410422).

References

[1] L. A. Zadeh, "Fuzzy sets and information granulation", Advances in Fuzzy Set Theory and Applications, North Holland Publishing, **(1979)**.
[2] L. A. Zadeh, "Towards a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic", Fuzzy Sets and Systems, vol. 19, **(1997)**.
[3] L. Wang, X. Liu and W. Pedrycz, "Effective intervals determined by information granules to improve forecasting in fuzzy time series", Expert Systems with Applications, vol. 40, no. 14, **(2013)**.
[4] V.G. Kaburlasos and S. E. Papadakis, "A granular extension of the Fuzzy-ARTMAP (FAM) neural classifier based on Fuzzy Lattice Reasoning (FLR)", Neurocomputing, vol. 72, no. 10-12, **(2009)**.
[5] V.G. Kaburlasos, L. Moussiades and A. Vakali, "Fuzzy lattice reasoning (FLR) type neural computation for weighted graph partitioning", Neurocomputing, vol.72, no.10-12, **(2009)**.
[6] V.G. Kaburlasos, "Towards a Unified Modeling and Knowledge-Representation Based on Lattice Theory - Computational Intelligence and Soft Computing Applications", Studies in Computational Intelligence, vol. 27, **(2006)**.
[7] V.G. Kaburlasos, S.E. Papadakis and G.A. Papakostas, "Lattice computing extension of the FAM neural classifier for human facial expression recognition", IEEE Transactions on Neural Networks and Learning Systems, vol. 24, no. 10, **(2013)**.
[8] V. G. Kaburlasos and L. Moussiades, "Induction of formal concepts by lattice computing techniques for tunable classification", Journal of Engineering Science and Technology Review, vol. 7, no. 1, **(2014)**.
[9] V. G. Kaburlasos and A. Kehagias, "Fuzzy inference system (FIS) extensions based on lattice theory", IEEE Transactions on Fuzzy Systems, vol. 22, no. 3, **(2014)**.
[10] E. Deza and M. Deza, "Dictionary of Distances", Elsevier, **(2006)**.
[11] H. Hu and Z. Shi, "Machine Learning in Granular Computing", Lecture Notes in Computer Science Vol. 5589, **(2009)**.
[12] Y.Y. Yao, N. Zhang, D.Q. Miao and F.F. Xu, "Set-theoretic approaches to granular computing", Fundamenta Informaticae, vol. 115, no. 2-3, **(2012)**.
[13] Y.Y. Yao, "Interpreting concept learning in cognitive informatics and granular computing", IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, vol. 39, no. 4, **(2009)**.
[14] Y.Y. Yao, "Information granulation and rough set approximation", International Journal of Intelligent Systems, vol. 16, no. 1, **(2001)**.
[15] Y. Qian, H. Zhang, F. Li, Q. Hu, J. Liang, "Set-based granular computing: A lattice model", International Journal of Approximate Reasoning, vol. 55, no. 3, **(2014)**.
[16] Y. Qian, J. Liang and C. Dang, "Knowledge structure, knowledge granulation and knowledge distance in a knowledge base", International Journal of Approximate Reasoning, vol.50, no.1, **(2009)**.
[17] S. Ramanna, A. H. Meghdadi and J.F. Peters, "Nature-inspired framework for measuring visual image resemblance: A near rough set approach", Theoretical Computer Science, vol. 412, no. 42, **(2011)**.

