ROUGH SPATIAL DESCRIPTION

Shuliang WANG [a], Deren LI [a], Wenzhong SHI [b], Xinzhou WANG [a], Deyi LI [c]

[a] National Laboratory for Information Engineering in Surveying Mapping and Remote Sensing, Wuhan University, 430079,Wuhan, China
slwang@liesmars.wtusm.edu.cn
[b] Department of Land Surveying & Geomatics, The HongKong Polytechnic University, Kowloon, Hong Kong, China
[c] China Institute of Electronic System Engineering, Beijing, China, 100039

Commission II, WG II/5

KEY WORDS: Rough set, GIS, Spatial uncertainty, Rough spatial entity, Rough topological relationship, Rough symbols

ABSTRACT:

Rough set is a new approach to uncertainties in spatial decision-making and analysis in GIS context. In this paper, rough set symbols are simplified and standardized, which are composed of rough interpretation and specialized indication. Rough spatial entities and their topological relationships are also proposed in rough space. Further, a universal intersected equation is developed, and rough membership function is extended with the grey degree of a pixel in a case study. First, rough set is simply reviewed, and a set of simplified rough symbols is advanced on the basis of different kinds of existed rough symbols. It includes both rough interpretation and specialized indication. Second, rough spatial entity is put forward in rough vector space, rough raster space and rough three-dimensional space. It is argued that GIS studies the real world as it is, without forcing uncertainties to change into crisp set. Third, rough topological relationships that are disjoint, touch, overlap, equal, cover, covered by, contain and contained by, are studied via rough matrix with their figures. They are divided into three types, CC (crisp entity and crisp entity), RC (rough entity and crisp entity) and RR (rough entity and rough entity). And a universal intersected equation is further proposed. Finally, a case is studied on river thematic map. The maximum and minimum maps of river thematic classification are generated via integrating the reduction of superfluous attributes, rough membership function and rough relationships.

1. INTRODUCTION

GIS tries to abstract, generalize and analyze a spatial entity in terms that people can understand, store and transfer via observation, relating the observation to a conceptual data model, representing the data in formal term, storage, retrieving the spatial entity, data mining in computerized information system. The terms are mainly temporal, spatial and thematic dimensions. A spatial entity may be interpreted to spatial phenomena, natural objects with geometric feature of point, line, area, volume, cases, states, processes, observations and so on. And the ideal spatial entity is defined and described crisply. However, the spatial entity is often complex and varying at many scales in time and space. And people have to select the most important spatial aspects. First the exact object model is used, and then cartographic convention enhances it (Burrough, Frank, 1996). But the procedure may lose details in one or more dimensions when the computerized GIS deals with the spatial entity. For example, a bus stop becomes a point without size or shape. Furthermore, some attribute values of the spatial entity, in many cases, are inaccessible, inexact or vague. The abovementioned make it indiscernible to associate a spatial element (e.g. pixel) to a given entity.

One of GIS fundamental functions is to determinate whether or not the spatial element belongs to the predefined entity. The classification determination is performed on the accessible attribute values that are measured by sensors. In order to improve the exact object model, continuous field model, error band, epsilon band, "S" band, fuzzy set, decision theory, cloud theory and so on have been further put forward and applied (Shi, Wang, 2001). As an extension of set theory for the study of spatial entity characterized by incomplete and inexact information (Pawlak, 1981,1982,1991), rough set is further developed and extended on spatial description in this paper.

Rough set specifies a spatial entity by giving an upper and a lower approximation. The lower approximation is the set of spatial elements that surly belong to the spatial entity, while the upper approximation is the set of spatial elements that possibly belong to it. Since the introduction, rough set has been applied in such many fields as knowledge-based medicine system, natural language processing, pattern recognition, decision systems, approximate reasoning and so on. Based on whether statistical information is used or not, existed rough set models may be grouped into such two major classes as algebraic and probabilistic models (Yao et al. 1997). Recently, rough set has also been applied in GIS for the advantages of rough set to handle spatial data with uncertainties. Spatial entities with indeterminate boundary (Burrough and Frank, 1996) may be taken as an embryonic form of rough set application in geo-informatics. The true spatial entity is the lower approximation, and the spatial entity with vague boundary is the upper approximation (Wang, Wang, Shi, 2001). In the sequence, Schneider (1997) discussed rough set in ROSE (Güting *et al*. 1995) on the formal modeling aspects, without discussing classification. Stell and Worboys (1998) used rough set to handle imprecision due to finite spatial or semantic resolution, which was affected by fuzzy set. Ahlqvist et al (2000) argued that rough set was a feasible alternative for GIS via rough classification and accuracy assessment. And a single rough classification and a relationship between two intersecting rough classification are discussed by them.

However, during the process of rough set applications and developments, various different descriptive symbols came into being, which has made it difficult to study rough set further. Simultaneously, the lower approximation is the subset of the upper approximation in rough set. That is to say, the lower approximation is always computed and stored twice, which leads to a great deal of resource waste for the certain lower approximation occupies the most part of a spatial entity.

2. ROUGH SET AND ITS IMPROVEMENTS

Pawlak (1982) originally considers a rough set as a family of sets with the same lower and upper approximations. Based on it, Iwinski (1987) regards a rough set as a pair of composed sets. Then Pawlak (1994) gives another way to describe a rough set by rough membership function. Those are classifies into algebraic and probabilistic models (Yao et al. 1997). In this section, rough set will be described in these parts: rough symbols standardization, rough set concepts, rough membership, and relationships with other similar approaches.

2.1 Trial To Standardize Rough Set Symbols

There exist various rough set models to be unified. With the applications of rough set, different types of symbols on the rough set concepts are developed due to different fields and intents. Even if the rough set founder, Zdzisław Pawlak, almost gave different symbols in his different papers (Pawlak 1981, 1982, 1991, 1994, 1997, 1998, 1999,2000). In order to master a paper, readers have to compare the new symbols with their known old ones. These have made it difficult to further communicate with each other in different applicable fields of rough set. The more widely rough set is used, the worse this situation will be. In the sequel, the further development of rough set will be impeded. "In view of many generalizations and extensions of rough set theory, some kind of unification of the basic theory seems to be badly needed." (Pawlak 1999). So it becomes very necessary to standardize various symbols. As a trial to unify rough set symbols, a set of simplified genetic rough symbols are proposed on the basis of existed different rough symbols (Pawlak 1982, 1991, 1994, 1997, 1998, 1999, Komorowski et al. 1999, Skowron and Grzymala-Busse 1994, Jitender et al 1997, Yao et al. 1997, Hu et al 1997, Ahlqvist et al 2000, Wang, Wang, Shi 2001), mainly on Pawlak symbols. The new symbol is composed of two parts, one is rough interpretation, and the other is specialized indication. They are shown and interpreted in Table 1. For example, "*Lr*", "*X*" of "*Lr*(*X*)" respectively denote "*L*ower approximation in *r*ough set", "spatial entity *X*" in GIS.

2.2 Brief Rough Set

Rough set characterizes both certainties and uncertainties. In Table 1 context, $Lr(X)$ is certain "Yes", $Neg(X)$ is sure "No", while both $Ur(X)$ and $Bnd(X)$ are uncertain "Yes or no". That is to say, with respect to an element $x \in U$, it is sure that $x \in Pos(X)$ belongs to X in terms of its features, but $x \in Neg(X)$ does not belong to X; while $x \in Bnd(X)$ cannot be ensured by means of available information whether it belongs to X or not. It can be seen that $Lr(X) \subseteq X \subseteq Ur(X) \subseteq U$, $U = Pos(X) \cup Bnd(X) \cup Neg(X)$, and $Ur(X) = Pos(X) \cup Bnd(X)$. X is defined iif $Lr(X) = Ur(X)$, while X is rough with respect to $Bnd(X)$ iif $Lr(X) \neq Ur(X)$. A subset $X \subseteq U$ defined with the lower approximation and upper approximation is called rough set. Rough degree is $rd(X) = Card(Ur(X) - Lr(X)) \,/\, Card(X) \times 100\%$. Where, $Card(X)$ denotes the cardinality of set X. X is crisp when $rd(X) = 0$. For instance, regard U as an image, the rectangle becomes a pixel.

2.3 Rough Membership Function

Probabilistic rough set is with respect to rough membership function. Rough set can also be defined with a rough membership function $\mu_X(x)$, $\mu_X(x) \in [0, 1]$, (Pawlak 1994, 1997, 1998, Yao et al. 1997). (See Equation 1)

The rough membership value may be regarded as the probability of $x \in X$ given that x belongs to an equivalence class. That is to say, it is taken for a conditional probability to illustrate a certain degree of x belonging to X, $\mu_X(x) + \mu_{\sim X}(x) = 1$. Let $P(X \mid [x]_R) = \mu_X(x)$ and $\alpha \in [0,1]$, a probabilistic rough set in α context is defined as Equation 2 (Yao et al. 1997). In this sense, $\mu_X(x)$ gives a probabilistic rough space of X via a pair of upper approximation and lower approximation.

Proposed symbols	Existed symbols	Presentations
U	U	Discourse universe that is a finite and non-empty set.
R	R	Equivalence relation on U, $R \subseteq U \times U$. (U, R) formalizes an approximate space.
X	X	Arbitrary set $X \subseteq U$
$\sim X$	$-X$, $\sim X$	The complement set of X, $X \cup (\sim X) = U$
$\frac{U}{R}$	$\frac{U}{R}$	Equivalence class set composed of disjoint subsets of U partitioned by R.
$[x]_R$	$[x]_{IND(R)}$, $[x]_R$, $R(x)$, $r(x)$	Equivalence class of R including element x. It may also interpreted as all rough-related elements of x, or a neighborhood of x.
$Lr(X)$	$\underline{APR}(X)$, $\bar{A}(X)$, $\bar{A}X$, $\overline{X}$, $Int(X)$, $R\_(X)$, R^*	Lower approximation (interior set) of X on U. $Lr(X) = \{x \in U \mid [x]_R \subseteq X\}$
$Ur(X)$	$\overline{A}\overline{P}\overline{R}\,(X)$, $\underline{A}(X)$, $\underline{A}X$, $\underline{X}$, $Cl(X)$, $R^{-}(X)$, R_*	Upper approximation (closure set) of X on U. $Ur(X) = \{x \in U \mid [x]_R \cap X \neq \Phi\}$
$Pos(X)$	$POS(X)$	Positive region. $Pos(X) = Lr(X)$
$Neg(X)$	$NEG(X)$	Negative region. $Neg(X) = U - Ur(X)$
$Bnd(X)$	$BND(X)$, $Bn(X)$, $Bd(X)$, $Boundary(X)$	Boundary region. $Bnd(X) = Ur(X) - Lr(X)$

Table 1. List of rough symbols and their definitions

$$\mu_X(x) = \frac{Card\left(X \cap [x]_R\right)}{Card\left([x]_R\right)} = \begin{cases} 1 & x \in Pos(x) \\ (0,\ 1) & x \in Bnd(x) \\ 0 & x \in Neg(x) \\ 1 - \mu_{\sim X}(x) & x \in \sim X \end{cases} \tag{1}$$

$$Lr_\alpha(X) = \{x \mid P(X \mid [x]_R) \geq 1\text{-}\ \alpha\},\ Ur_\alpha(X) = \{x \mid P(X \mid [x]_R) > \alpha\} \tag{2}$$

2.4 Reduction of Superfluous Attributes

It is proper to reduce superfluous attributes when making decisions. A spatial entity is characterized by spatial attributes that are divided into decisive attribute D and conditional attribute C. Suppose $C = (C_1, C_2, \ldots, C_k, \ldots, C_m)$ with values $V = (V_1, V_2, \ldots, V_k, \ldots, V_m)$, and $D = V_d$. A spatial decision on the spatial entity is often represented in the form of rules that indicate the degree of association between C and D. And the rule is always denoted as an implication:

Rule 1: $(C_1 = V_1) \wedge (C_2 = V_2) \wedge \ldots \wedge (C_k = V_k) \wedge \ldots \wedge (C_m = V_m) \rightarrow (D = V_d)$

When C_k is removed, Rule 2 appears.

Rule 2: $(C_1 = V_1) \wedge (C_2 = V_2) \wedge \ldots \wedge (C_m = V_m) \rightarrow (D = V_d)$

If *Rule* 1 = *Rule* 2, C_k is superfluous in C with respect to D. Thus any superfluous attributes in spatial decision may be removed without changing the dependency relationship of the original attributes, so as to save storage and speed up decision.

2.5 Differences between Rough Set and Other Methods

There are relationships between rough set and fuzzy set, cloud theory, evidence theory (Shi, Wang 2001). All of them can deal with uncertainties, such as transition between qualitative concept and quantitative data, characterizing indeterminate phenomena via mathematical syntax and semantics. However, rough set can be told from the others in some aspects. In the following, please be noted that x is a spatial parameter, and $\mu(x)$ is its corresponding membership value to a class X.

[1] Rough set gives an interval of $[\mu_{min}(x), \mu_{max}(x)]$ with respect to x. In other words, an element has many corresponding values, one to many. And the determination is that the element "is", "is not" or "is maybe" in a given class. These values formalize the interval. The data points in $Bnd(X)$ between the lower approximation and upper approximation is rough for set X. And it is not sure that they belong to the set X or not. As an extension of the classical (traditional, sharp or crisp) set, rough set focuses on the uncertainties caused by incomplete, insufficient or inaccessible information. Compared with other methods, rough set can close describe the spatial entities as they are in the real world, including both certainties and uncertainties.

[2] Fuzzy set gives a value $\mu(x)$, $\mu(x) \in [0,1]$, via a fuzzy membership function, with respected to x. The relationship is one parameter to one functional value. Fuzzy set is also an extensive set of the classical set, and may perform an uncertain classification. But fuzzy set pays more attention to the uncertainties caused by vague, dim or indistinct information, and it is either difficult or rather arbitrary to determine the fuzzy membership functions. Moreover, fuzzy set depends on human experience, and it loses uncertain properties once the fuzzy membership degree $\mu(x)$ is given.

[3] Cloud theory, which has three numerical characteristics, specifies a discrete data point with the value $\mu(x)$ in x context. The tuples $(x, \mu(x))$ are called cloud drops. The discrete degree is determined by the membership $\mu(x)$. But the range and interval of $\mu(x)$ is unsure. Cloud model is also the uncertainty transition between a linguistic term of a qualitative concept and its numerical representation

[4] Evidence theory, is also named as Dempster-Shafer theory, Dempster-Shafer theory of evidence or Dempster-Shafer theory of belief function. It has belief function and plausibility function, which are similar to the upper and lower approximations of rough set. The overall crispness measure can be interpreted as a belief value in the sense of Dempster-Shafer logic. However, the belief function depends on experience. This similarity has motivated the work on the relationships between rough set and evidence theory.

3. ROUGH SPATIAL ENTITY

Both spatial entities and spatial relationships formalize an approximate space. As an alternative, rough set is proposed to characterize spatial entities in GIS. U is composed of spatial entities with attributes (interpreted as features, variables etc.), and R is the spatial relationship among the spatial entities. Both of them formalize an approximate space (U, R). Point, line and area in vector space, pixel and grid in raster space, unit cube in a multi-dimensional space are considered equivalence class of rough spatial entity. In rough set context, point, line, area and volume have size and shape. Attributes and a pair of approximations describe a point, and a series of such points linked together are lines. The lines called boundaries bound areas, and volumes are bounded by smooth area.

A pair of upper approximation and lower approximation specifies a rough spatial entity. Given a spatial entity $X \subseteq U$, X may be impossible to be represented precisely for the available information is insufficient. The observed value of an attribute is usually unequal to its true value. When an attribute has been observed for many times, the observed values may formalize an uncertain observed zone around the true value, namely a pair of approximations. As to a spatial element $x \in U$, lower approximation $Lr(X)$ is the set of x that surly belongs to the true X, while upper approximation $Ur(X)$ is the set of x that possibly belongs to X. And uncertain region of X is $Bnd(X)$ (Figure 1). Thus, during the spatial analysis based on GIS, rough set can more totally propagate the spatial entity properties (both certain and uncertain) for most spatial true values are unable to know exactly. As an alternative mathematical interpretation in the sense of rough set, object model is $Lr(X) = Ur(X)$, field model , error band, epsilon band, and "S" band are $Lr(X) \neq Ur(X)$. And for rough degree $rd(X)$, field model > error band > epsilon band > "S" band. Each of them may be taken as the special condition of rough space. Since vector data and raster data are main original data in GIS, rough vector space and rough raster space will be mainly studied in this section.

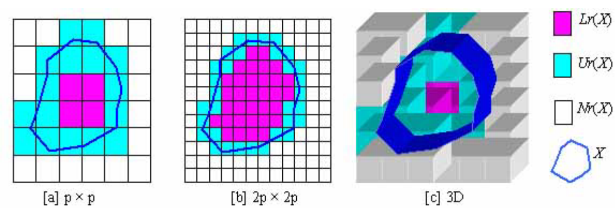

Figure 1. Rough spatial entity and its illustrations of low resolution, high resolution and 3D

3.1 Rough Vector Space

The object model represents spatial entities via crisply delineated point, line, area and volume in a defined absolute reference system. Their attributes that characterize the space at the points, along the lines, within the area or volumes are assumed to be constant over the whole object extent. It is implemented by GIS vector structure. For example, lines are linked by a defined topology to form networks, which, if open, can represent rivers, or if closed, the abstract or defined boundaries of polygons that in turn represent land parcels, soil units or administrative areas. The object model is assumed $Lr(X) = Ur(X)$ without roughness. In fact, $Lr(X) \neq Ur(X)$ when reality is described by object model in computerized GIS. And spatial vector objects often have an extension around them for errors and uncertainties made by unavailable information (Figure 2 [a[, [b]). Given uncertain positive parameters δ_1, δ_2 in rough set context, X can be represented $X = Lr(X) + \delta_1$ or $X = Ur(X) - \delta_2$. In the sense of δ_1 and δ_2, $Bnd(X) = \delta_1 + \delta_2$, $\sim X = U - X = U - Lr(X) + \delta_1 = U - Ur(X) + \delta_2$. Error ellipse may be used as their depicted mathematical model. Burrough (1996) argued that object model was suitable for a spatial entity that could be mapped on external features of the landscape, while field model adapted to a spatial entity when its single quantitative attributes were measured and mapped.

3.2 Rough Raster Space

Rough raster space brings approximations into the shapes and forms of a spatial entity. Raster data is for field model opposed to object model. Rough spatial point, line and area in the raster space are essential when the real world is put into a computerized GIS. They are illustrated in Figure 2 [a], [c]. As Figure 2 revealed, $Lr(X)$ of the point and line are both empty. $Lr(X)$ of the area has only two equivalence class. All $Ur(X)$are relatively bigger. So spatial uncertainties (positional and attribute uncertainties) in GIS really exists. Cartographic generalization is a changeable processing of the lower approximation of spatial objects and their upper approximation. However, the pair of approximations of various spatial entities changes in different directions. One becomes bigger, while the other gets smaller.

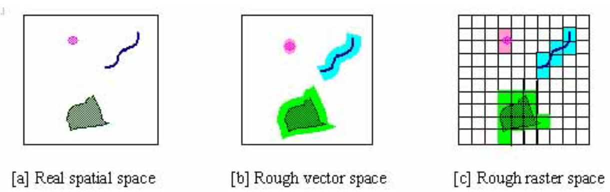

Figure 2. Rough spatial point, line and area

Rough set gives a new interpretation on image resolution. Spatial raster data become more and more important for many images are raster. A raster is regarded as a spatial equivalence class in the rough raster space. The spatial entities are defined with such raster data approximately, especially to boundaries. And a piece of spatial image is discretized to a regular grid, i.e. an image pixel at a predetermined resolution. The image resolution decides the pixel size. The higher the image resolution is, the less rough degree $rd(X)$ of the spatial raster entity X is. When the resolution is high enough, or the raster is small enough, the pair of lower and lower approximations of an entity are equal, $Lr(X) = Ur(X)$. Namely, the entity is not rough. However, bigger computerized storage is also demanded. This is another interpretation on remote sensing image changing with resolution in the sense of rough set, such remote sensing as satellite, aeroplane, and photogrammetry.

Rough multi-dimensional space is composed of a series of unit spatial cubic objects. Spatial object is composed of such blocks. This seems like building is built up with toy's blocks. Blocks belonging to the lower approximation are included in the spatial object, while the skin of the objects crosses blocks belonging to the upper approximation but not belonging to the lower approximation. That is to say, two "balls" with the same center represent a spatial entity in the multi-dimensional rough raster space. One with a smaller radius is composed of the lower approximation, while the other with a bigger radius is the upper approximation.

3.3 Study Objects as They Are

Mathematically, point has no size, line has length but no size, and area has no thickness. The attributes of a spatial entity are assumed to very continuously and smoothly, which can be described with a smooth mathematical function. However, this model is so abstract that it is not as well as the real world. Thus, uncertainties are unavoidable when abstract mathematical object is used to study the complex real object. It is ideal to study a spatial object as it is. Rough set tries its best to maintain the original characters of the real world via a pair of lower and upper approximations. True value is the lower approximation, while the observed extension is the upper approximation. When a spatial entity has been observed for several times, observed values formalize an extension around the true value because of insufficient information. The incomplete information may be from instruments, human being or mathematical functions. Rough set can keep and propagate the uncertain information until final decisions. We argue that superfluous information is better than removal information before a decision is determined.

4. ROUGH SPATIAL RELATIONSHIPS

Rough spatial relationships describe spatial relationships more completely. Rough topological relationship R is essential in a rough space (U, R). Before rough topology is advanced, it is necessary to firstly review the development of topological relationships. The meaning of standard topology is defined by Munkres (1975). Original spatial topological relationships are for simple point (0-dimensional), line (1-dimensional) and area (2-dimensional), with 4-intersection model on interior X^0 and boundary ∂X. When its limitations appear, it is extended to 9-intersection model on interior, boundary and exterior X^- (Egenhofer 1991, Egenhofer and Franzosa 1991, Egenhofer and Herring 1991, Egenhofer and Al-taha 1992, Egenhofer 1993, Egenhofer et al. 1993, Egenhofer and Sharma 1993, Egenhofer 1994, Egenhofer et al. 1994, Egenhofer and Mark 1995, Egenhofer and Ranzosa 1995, Florence and Egenhofer 1996, Clementini et al. 1993, Clementini et al. 1994, Clementini and Di Felice 1995). Then Clementini and Di Felice (1996) introduce areas with broad boundaries composed of an inner boundary and an outer boundary, and reduce the 2^9 topological matrices to the 44 matrices with 0 and 1 values. Chen et al. (2001) propose a Voronoi-based 9-intersection model via replacing the exterior X^- of an entity with its Voronoi region X^v with o (empty) and Ø (none-empty) values. However, it is difficult to ensure their interior X^0, exterior X^-, or X^v exactly because of insufficient information. In the sequel, boundary ∂X is also unsure. It is a true case that uncertainties exist, which is unavoidable in GIS. As an alternative, we propose rough topology via respectively replacing the interior, boundary and exterior with positive region, boundary region and negative region as Equation 3.

$$R_{r9}(A,B) = \begin{pmatrix} Pos(A)\cap Pos(B) & Pos(A)\cap Bnd(B) & Pos(A)\cap Neg(B) \\ Bnd(A)\cap Pos(B) & Bnd(A)\cap Bnd(B) & Bnd(A)\cap Neg(B) \\ Neg(A)\cap Pos(B) & Neg(A)\cap Bnd(B) & Neg(A)\cap Neg(B) \end{pmatrix} \quad (3)$$

Equation 3 is surely able to tell and propagate certainties ($Pos(X)$, $Neg(X)$) and uncertainties ($Bnd(X)$). 1 (none-empty) and 0 (empty) values are employed for GIS is computerized. Note that $Neg(X)$ is different from $\sim X$, the complement of X for $Neg(X) = U - Ur(X)$, while $\sim X = U - X = U - Ur(X) + \delta_2$. So rough spatial relationships give richer information that includes certain and uncertain data, which can improve the quality of image interpretation. By the way, Equation3 is universal whenever different thematic maps are overlapped. In the rough space of the same image map, it is sure that $Pos(A) \cap Pos(B) = 0$.

The rough relationships may be divided into three kinds, i.e., CC (rough relationships between crisp entities and crisp entities), RC (Rough relationships between rough entities and crisp entities) and RR (rough relationships between rough entities and rough entities). Here, rough area-area topological relationships in 2-dimensional space are proposed mainly. Because area is from line, and line is from point, area is studied as a case. The topologies of point-point, point-line, point-area, line-line and line-area may be regarded as the special cases of area-area. Figure 3 illustrates the intersection relationships between two rough spatial entities. Where, $Lr(A)$, $Lr(B)$ are respectively the lower approximation of rough entities A, B; $Ur(A)$, $Ur(A)$ are respectively the upper approximation; $Bnd(AB)$ is a rough region between A and B, which is the most uncertain part. Because the indeterminate region often happens in the boundary, it is unable for an uncertainty to take place between the lower approximation A and B. So the meet relationship often exists at the indeterminate transition zone in image classification, which is composed of two neighboring upper approximations. In the rough space, the set of topological relationships are {disjoint, touch/meet, overlap, equal, covers, covered by, contains, contained by/ inside} which are studied via rough matrices with their figures (Figure 3). Excluding spatial entities that contain roughness, there are also crisp spatial entities (e.g. administrative boundary) in rough space. According to the abovementioned, a crisp spatial entity X is a special rough entity where $Lr(X) = Ur(X)$. So rough spatial relationships in the same rough space are divided into three types, CC (crisp entity and crisp entity, Figure 3 (a)), RC (rough entity and crisp entity, Figure 3 (b)) and RR (rough entity and rough entity, Figure 3 (c)).

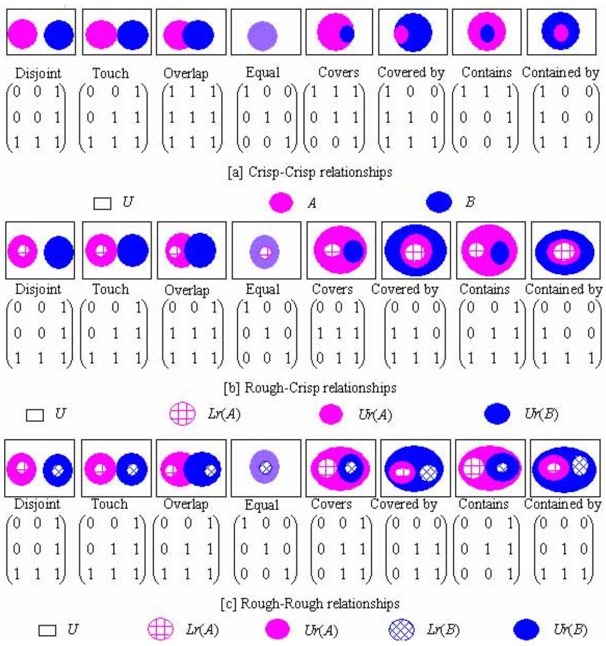

Figure 3. Rough spatial topological relationships and their corresponding rough matrices

Moreover, a universal equation can be deduced from Equation 3 to represent the intersected rough regions. When more than two rough spatial entities are intersected, rough regions among them are proposed to describe with $Bnd(A_1, \ldots, A_i, \ldots, A_n)$ (See Equation 4).

$$R_{r9}(A_1,\ldots,A_i,\ldots,A_n)=\begin{pmatrix} R_{r9}(A_1,A_1) & \ldots & R_{r9}(A_1,A_i) & \ldots & R_{r9}(A_1,A_n) \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ R_{r9}(A_i,A_1) & \ldots & R_{r9}(A_i,A_i) & \ldots & R_{r9}(A_i,A_n) \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ R_{r9}(A_n,A_1) & \ldots & R_{r9}(A_n,A_i) & \ldots & R_{r9}(A_n,A_n) \end{pmatrix} \quad (4)$$

Here, we take $n = 3$ as an example to interpret the equation. Supposed there are three rough spatial entities A, B and C, which are intersected with each other. Besides the two intersected regions, $Bnd(A,B)$, $Bnd(A,C)$ and $Bnd(B,C)$, a new rough region $Bnd(ABC)$ also appears (See Figure 4).

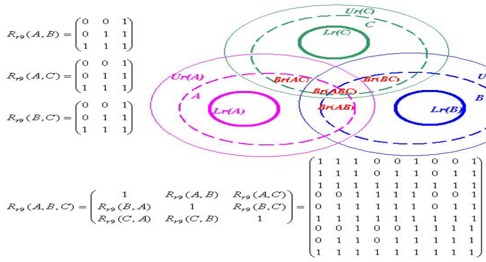

Figure 4. Rough topological relationships among A, B and C and their rough matrices

5. CASE STUDY

As a case study, the method of rough spatial description -based rough classification is proposed and is used to extract river information from a remote sensing image. Based on a pair of lower and upper approximations, maximum and minimum maps of river thematic classification are generated via integrating the reduction of superfluous attributes, rough membership function and rough relationships. The original image (Figure 5 (a)) is a piece of remote sensing TM image. There are many conditional attributes affecting the decisional attributes, image classification. The conditional attributes include image grey degrees, the satellite parameters, air refraction and so on. After other conditional attributes are reduced, grey degree is selected to extract the river classification from the image. Let G_x be the grey degree of a pixel x and G_X is the grey degree of river pixel. Then the rough membership function (See Equation 5) can be extended from Equation 1. As Figure 5 (b)(c) revealed, the lower approximation $Lr(X)$ is the minimum water map with certainties, while the upper approximation $Ur(X)$ is the maximum water map with uncertainties. Here, $rd(X) = Card(Ur(X) - Lr(X)) / Card(X) \times 100\% = 10.37\%$. Compared with the crisp classification with only one result, the rough classification not only includes both certainties and uncertainties, but also tells the certainties from the uncertainties.

$$\mu_X(x)=\frac{G_x}{G_X}=\begin{cases} 1 & x\in Lr(X) \\ (0,\ 1] & x\in Ur(X) \\ 0 & x\in Neg(X) \end{cases} \quad (5)$$

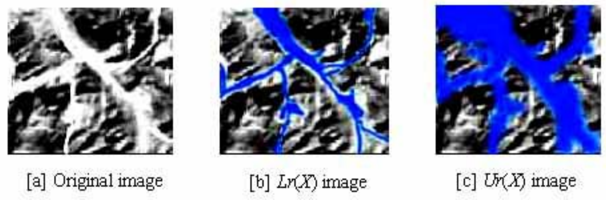

Figure 5. Rough water thematic maps continuums）

Furthermore, the result are compared with that from the maximum likelihood classification and the fuzzy classification, which indicates that the rough classification based on rough set contains more information and with high precision. We get the maximum possible river, minimum certain river, rough confidence degree, possible error, etc. The maximum possible river approaches the river in floodtime, while the minimum certain river comes near the river in low water. Moreover, the precision is improved 7% than the maximum likelihood classification or 2% than fuzzy classification.

6. CONCLUSION

Based on existed various rough symbols, this paper simplified and standardized a set of rough set symbols. A proposed symbol is composed of rough interpretation and specialized indication. Then rough set is compared with such other similar methodologies, the differences of which show rough set can close describe the spatial entity as it is in the real world. In rough set context, the paper also proposed rough spatial entities and their topological relationships in rough space. A universal intersected equation and rough membership function with grey degree are further developed. Rough topological relationships that are disjoint, touch, overlap, equal, cover, covered by, contain and contained by, were studied via rough matrix with their figures. And a universal intersected equation is further proposed. The result of the case study not only included rich information but also was confidential and practical. This indicates the method of rough spatial description is a valuable alternative to study Geomatics.

ACKNOWLEDGEMENTS

National High Technology R&D Program (863) (2001AA135081), National Natural Science Foundation of China (40023004), and Grant of Earmarked Provision for Special Projects (1.34.37.9709) of Hong Kong support this study.

REFERENCES

AHLQVIST,O. et al., 2000, Rough classification and accuracy assessment. *International Journal of Geographical Information Science*, 14(5), 475-496

BURROUGH, P. A., Frank, A. U. (editors), 1996, *Geographic Objects with Indeterminate Boundaries* (London: Taylor & Francis).

CLEMENTINI, E., DI FELICE, P., and VAN OOSTEROM, P., 1993, A small set of formal topological relationships suitable for end-user interaction. In *Advances in Spatial Databases (SSD'93)*, Lecture Notes in Computer Science, No. 692, edited by D. Abel and B. C. Ooi (New York: Springer-Verlag), pp. 277–295.

CLEMENTINI, E., SHARMA, J., and EGENHOFER, M., 1994, Modelling topological spatial relations: Strategies for query processing. *Computer and Graphics*, 18,pp.815–822

CLEMENTINI, E. and DI FELICE, P., 1995, A comparison of methods for representing topological relationships. *Information Sciences,* 3pp.149-178

CLEMENTINI, E. and DI FELICE, P., 1996, An algebraic model for spatial objects with indeterminate boundaries. In *Spatial Objects with Indeterminate Boundaries*, edited by P. A. Burrough and A. U. Frank, 1996 (London: Taylor & Francis), pp.155–169

EGENHOFER, M. J., 1991, Reasoning about binary topological relations. In *Proceedings of the 2nd Symposium on Large Spatial Databases*, Lecture Notes in Computer Science, No.523 (New York: Springer-Verlag), pp. 143–160

EGENHOFER, M. J. and FRANZOSA, R. D., 1991, Point-set topological spatial relations. *International Journal of Geographical Information Systems*, 5(2),pp.161–174.

EGENHOFER, M. J. and AL-TAHA, K., 1992, Reasoning about gradual changes of topological relations, In *Theories and Methods of Spatio-temporal Reasoning in Geographic Science*, edited by A. U. Frank, I. Campari and Formentini, Lecture Notes in Computer Science 639 (New York: Springer-Verlag), pp. 196–219

EGENHOFER, M. J., 1993, Definition of line-line relations for geographic databases. *IEEE Data Engineering Bulletin*, 7,pp.40–45

EGENHOFER, M. J., SHARMA, J., and MARK, D. M., 1993, A critical comparison of the 4-intersection and 9-intersection models for spatial relations: formal analysis. *Auto-Carto*, 11,pp.1–11.

EGENHOFER, M. J. and SHARMA, J., 1993, Topological relations between regions in IR 2 and ZZ 2. In *Advances in Spatial Databases (SSD'93*), edited by D. Abel and B. C. Ooi, Lecture Notes in Computer Science 692 (New York: Springer-Verlag), pp. 316–226.

EGENHOFER, M. J., 1994. Deriving the composition of binary topological relations. *Journal of Visual Languages and Computing*, 5,pp.133–149

EGENHOFER, M. J., CLEMENTINI, E., and DI FELICE, P., 1994, Topological relations between regions with holes. *International Journal of Geographical Information Systems*, 8(2), pp.129–142

EGENHOFER, M. J. and MARK, D. M., 1995, Modelling conceptual neighborhoods of topological line-region relations. *International Journal of Geographical Information Systems*, 9(5),pp. 555–565

EGENHOFER, M. J. and FRANZOSA, R. D., 1995, On the equivalence of topological relations. *International Journal of Geographical Information Systems*, 9(2),pp. 133–152.

FLORENCE, J. and EGENHOFER, M. J., 1996, Distribution of topological relations in geographic datasets. In *Proceedings of 1996 ASPRS/ACSM Annual Convention and Exposition* (American Society for Photogrammetry and Remote Sensing: Bethesda), pp. 314–324.

IWINSKI, T.B., 1987, Algebraic approach to rough sets. *Bulletin of the Polish Academy of Sciences, Mathematics,* 35, pp.673-683

KOMOROWSKI,J., et al., 1999, Rough Sets: A Tutorial. In *Rough Fuzzy Hybridization,* edited by S.K.PAL and A.SKOWRON (Singapore: Springer), pp.3-98

MUNKRES, J.R., 1975, *Topology: A first course* (Englewood Cliffs, NJ: Prentice Hall)

PAWLAK, Z., 1981, Information systems – theoretical foundations. *Information Systems*, 6, pp.205-218

PAWLAK,Z., 1982, Rough Sets. *International Journal of Computer and Information Sciences*, 11(5), pp.341-356

PAWLAK,Z., 1984, Rough Classification. *International Journal of Man-Machine Studies*, 20, pp.469-483

PAWLAK,Z., 1985, Rough Sets and Fuzzy Sets. *Fuzzy Sets and System,* 17, pp.99-102

PAWLAK,Z., 1991, *Rough sets: theoretical aspects of reasoning about data* (London: Kluwer Academic Publishers)

PAWLAK,Z., 1997, Rough Sets. In *Rough Sets and Data Mining Analysis for Imprecise Data,* edited by Y.LIN and N.CERCONE (London: Kluwer Academic Publishers), pp.3-7

PAWLAK,Z., 1998, Rough set elements. In *Rough Sets in Knowledge Discovery 1: Methodologies and Applications*. Studies in fuzziness and soft computing, Vol.18, edited by POLKOWSKI, L., SKOWRON, A. (Heidelberg: Physica-Verlag), pp.10-30

PAWLAK,Z., 1999, Rough sets, rough Function and rough calculus. *Rough-Fuzzy Hybridization: A New Trend in Decision-Making*, edited by PAL S, SKOWRON A (Singapore: Spring-Verlag), pp.99-108

PAWLAK,Z., POLKOWSKI,L., SKOWRON,A., 2000, Rough sets and rough logic: a KDD perspective. Rough Methods and Applications: New Developments in Knowledge Discovery in Information Systems, edited by Polkowski,L., Tsumoto S., LIN, T. Y.（Berlin: Phisica-Verlag), pp.583~648

SCHNEIDER, M., 1997, *Spatial Data Types for Database Systems*, Lecture Notes in Computer Science, vol. 1288 (Berlin: Springer)

SHI, W.Z., WANG, S.L., 2001, State of the art of research on the attribute uncertainty in GIS data. Journal of Image and Graphics, 6[A](9), pp.918-924

SKOWRON, A., and GRZYMALA-BUSSE, J., 1993, From rough set theory to evidence theory. In *Advances in the Dempster-Shafer Theory of Evidence*, edited by R. Yager, M. Fedrizzi and J. Kasprzykpages (New York: John Wiley and Sons), pp. 193–236

STELL, J., and WORBOYS, M., 1998, Stratified map spaces: A formal basis for multi-resolution spatial databases. In *Proceedings of the 8th International Symposium on Spatial Data Handling* (Burnaby, Canada: International Spatial Union), pp. 180–189

WANG, S.L., X.Z.WANG and W.Z. SHI, 2001, Development of a Data Mining Method for Land Control. Geo-spatial information science, 4(1), pp.68-76

YAO,Y.Y., WONG,S.K.M., LIN,T.Y., 1997, A Review of Rough Set Models. In *Rough Sets and Data Mining Analysis for Imprecise Data,* edited by Y.LIN and N.CERCONE (London: Kluwer Academic Publishers), pp.47-75