2002

Fuzzy Spatial Querying with Inexact Inference

Huiqing Yang
University of Southern Mississippi

Maria Cobb
University of Southern Mississippi

Dia Ali
University of Southern Mississippi

Shahram Rahimi
University of Southern Mississippi, rahimi@cs.siu.edu

Frederick E. Petry
Tulane University of Louisiana

See next page for additional authors

Published in Yang, H., Cobb, M., Ali, D., Rahimi, S., Petry, F. E., & Shaw, K. B. (2002). Fuzzy spatial querying with inexact inference. Proceedings, 2002 Annual Meeting of the North American Fuzzy Information Processing Society, 2002, NAFIPS, 377-382. doi: 10.1109/NAFIPS.2002.1018089

Authors
Huiqing Yang, Maria Cobb, Dia Ali, Shahram Rahimi, Frederick E. Petry, and Kevin B. Shaw

Fuzzy Spatial Querying with Inexact Inference

Huiqing Yang, Maria Cobb
Dia Ali, Shahram Rahimi
Department of Computer Science & Statistics
University of Southern Mississippi
Hattiesburg, Mississippi, USA 39406-5601
hyan2@orca.st.usm.edu
Maria.Cobb@usm.edu
dia.ali@usm.edu
shahram.rahimi@usm.edu

Frederick E. Petry
Electrical Engineering & Computer Science
Tulane University, New Orleans, LA 70118
petry@eecs.tulane.edu
Kevin B. Shaw
Naval Research Laboratory Mapping,
Charting & Geodesy
Stennis Space Center, MS, USA 39529
shaw@nrlssc.navy.mil

Abstract

The issue of spatial querying accuracy has been viewed as critical to the successful implementation and long-term viability of the GIS technology. In order to improve the spatial querying accuracy and quality, the problems associated with the areas of fuzziness and uncertainty are of great concern in the spatial database community. There has been a strong demand to provide approaches that deal with inaccuracy and uncertainty in GIS. In this paper, we are dedicated to develop an approach that can perform fuzzy spatial querying under uncertainty. An inexact inferring strategy is investigated. The study shows that the fuzzy set and the certainty factor can work together to deal with spatial querying. Querying examples implemented by FuzzyClips are also provided.

Keywords: *uncertainty, inexact inferencing, fuzzy inference, spatial query, GIS, FuzzyClips*

1. Introduction

Since the spatial querying deals with some concepts expressed by verbal language, the fuzziness is frequently involved. Hence, the ability to query a spatial data under the fuzziness is one of the most important characteristics of any spatial databases. Some researchers have shown that the directional as well as topological relationships are fuzzy concepts [1-2]. To support queries of this nature, our earlier works [3-6] provided a basis for fuzzy querying capabilities based on a binary model. The Clips-based implementation [6] shows the fuzzy querying can distinguish various cases in the same relation classes. For instance, consider the example relationship *Object A* ***overlaps*** *Object B.* The fuzzy querying can answer: does *all of Object A* ***overlap*** *some of Object B, or does little of Object A* ***overlap*** *most of Object B?*

However, in these kinds of fuzzy queries, the representation of the fuzzy variables is based on classical set theory. Although classical sets are suitable for various applications and have proven to be an important tool for mathematics and computer science, they do not reflect the nature of human concepts and thoughts, which tend to be abstract and imprecise. The flaw comes from the sharp transition between inclusion and exclusion in a set. In this paper we show a way to use the fuzzy set for dealing with the vague meaning of linguistic terms, in which the smooth transition is characterized by membership function.

The queries expressed by verbal language often involve a mixture of uncertainties in the outcomes that are governed by the meaning of linguistic terms. Therefore, there is an availability-related need for skilled inexact inferring approach to handle the uncertain feature [7]. Uncertainty occurs when one is not absolutely certain about a piece of information. Although uncertainty is an inevitable problem in spatial queries, there are clear gaps in our understanding of how to incorporate uncertain reasoning into the spatial querying process. This requires performing an inexact inferencing. Recently, models of uncertainty have been proposed for spatial information that incorporate ideas from natural language processing, the value of information concept, non-monotonic logic and fuzzy set, evidential and probability theory. Each model is appropriate for a different type of inexactness in spatial data. By incorporating the fuzzy set and confirmation theory, we investigate an inexact inferencing approach for fuzzy spatial querying. The aim is to improve spatial querying accuracy and quality.

The paper is organized as follows. Section 2 briefly overviews our previous works, and shows some basic techniques and strategies to deal with fuzzy multiple relations in spatial querying. Section 3 describes an approach that can perform fuzzy querying under the uncertainties. In section 4, FuzzyCLIPS implementation shows some improved querying results.

2. An Overview of Previous Works

Assume that the spatial objects can be approximated by their minimum bounding rectangles (MBR). Figure 1 shows two objects in two dimensions. Based on the spatial binary model [3-6], some spatial querying techniques and strategies can be briefly overviewed as follows.

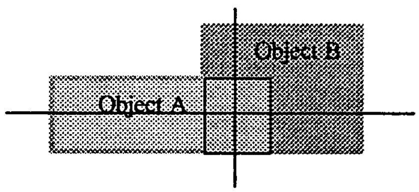

Figure 1. Two objects in 2-D

2.1 Basic Spatial Querying

Topological and directional relationships are critical components in the retrieval of information from spatial databases, including image, map and pictorial databases. Many contributions have been made. The authors in [10] define new families of fuzzy directional relations in terms of the computation of force histograms, which is based on the raster data. In this paper, we will take into account these two major spatial relations based on the vector data.

The topological relationships express the concepts of inclusion and neighborhood. A large body of related work has focused on the intersection mode that describes relations using intersections of object's interiors and boundaries. By means of geometrical similarity, we defined the topological relationships as a set:

T={disjoint, tangent, surround, overlaps ......}.

The paper [3] provides greater details on this.

The directional relationships are commonly concerned in everyday life. Most common directions are cardinal direction and their refinement. We defined the directional relations as a following set:

D={North, East, South, West, Northeast, Southeast, Southwest, Northwest}.

Such relationships provided a significant resource for the basic binary spatial queries. The examples of such queries might look like these:

Object A **overlaps** Object B.
Object A is **west** of Object B.

2.2 Fuzzy Spatial Querying

Although the above querying method can provide topological and directional information, these kinds of information do not associated with any degrees. This means it can only perform a low level query. A typical example is shown in Figure 2.

Figure 2.

For both cases that belong to the same class (or relation group), the basic spatial querying will provide the same topological and directional relationships, i.e. Object A overlaps Object B and Object A is west of Object B.

How to provide high accurate information, such as most of Object A overlaps some of Object B, or little of Object A overlaps some of Object B and so on, encourages us to make the further investigation. Some strategies and techniques can be briefly described as follows (see the details in [6]).

- Partition each object into sub-groups in eight directions based on the reference area (the common part of two objects) shown in Figure 3;
- Map each sub-group to a node, and assign two weights (area and node weights) to each node;
- Calculate two weighs to determine the special degree.

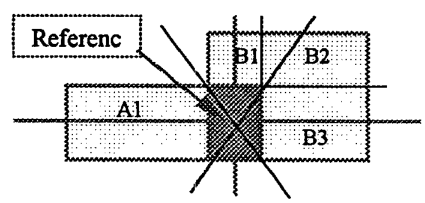

Figure 3. Partitioning two objects in 2D

Where area weight can be calculated by

AW=(area of sub-group) / (area of the entire object)

and node weight can be obtained by

NW=AW · (axis length) / (longest axis length).

In order to support fuzzy querying, the resulting quantitative figures (AW, NW) are mapped to a range that corresponds to a term known as linguistic qualifiers. There is a huge body of knowledge and techniques that deal with fuzzy spatial relations in linguistic expression. In this paper, we define the topological qualifier TQ and directional qualifier DQ as:

TQ={all, most, some, little, none}

DQ={directly, mostly, somewhat, slightly, not}.

As mentioned in [9], relative qualifiers can be represented as fuzzy subsets of the unit interval and use linguistic word. Based on the classical set, the membership function of qualifiers can be defined as a binary set, that is, complete membership has a value of 1, and no membership has a value of 0. The following tables give the quantifying description.

Table 1. Topological Qualifiers

Topological Qualifiers (TQ)	Area Weight (AW)
all	0.96 to 1.00
most	0.60 to 0.95
some	0.30 to 0.59
little	0.06 to 0.29
none	0.00 to 0.05

Table 2. Directional Qualifiers

Directional Qualifiers (DQ)	Node Weight (NW)
directly	0.96 to 1.00
mostly	0.60 to 0.95
somewhat	0.30 to 0.59
slightly	0.06 to 0.29
not	0.00 to 0.05

As shown in Figure 1, the based-Clips implementation can provide the following information:

Most of Object A overlaps Object B
Object A overlaps some of Object B

Most of Object A overlaps some of Object B

Most of Object A is west of Object B
Object A is mostly west of Object B

Most of Object A is mostly west of Object B

3. Fuzzy Querying Under Uncertainty

Because the spatial relationships depend on human interpretation, spatial querying should be related by fuzzy concepts. To support queries of the nature, previous works provided fuzzy queries without uncertainty that can handle the fuzziness by defining fuzzy qualifiers. However, in these kinds of fuzzy queries, the particular grades of membership have been defined as classical sets. The problem is there exist a gap between two neighboring members such as '*all*' and '*most*'. Because a jump occurs, no qualifier is defined in some intervals, for example the interval (0.95, 0.96). To improve the fuzzy querying, the fuzzy set theory is concerned in our continuous research.

3.1. Fuzziness Consideration

Fuzziness occurs when the boundary of a piece of information is not clear-cut. Hence, fuzzy querying expands query capabilities by allowing for ambiguity and partial membership. The definition of the grades of membership is subjective and depends on the human interpretation. A way to eliminate subjectivity is another interested research field. Here simple membership functions will be considered.

A fuzzy set is a set without a crisp boundary. The smooth transition is characterized by membership functions that give fuzzy sets flexibility in linguistic expressions. More formally a fuzzy set in a universe is characterized by a membership function $\mu: U \rightarrow [0,1]$. Figure 4 illustrates the primary term of fuzzy variable area weight. Each term represents a specific fuzzy set.

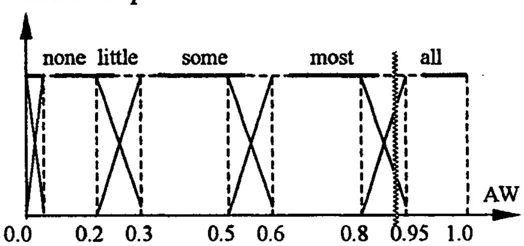

Figure 4. Membership function for TQ

The fuzzy set functions for topological qualifiers can be described as:

$$\mu_{all}(AW) = \begin{cases} 1.0 & \text{if } 0.95 \le AW \le 1.0 \\ 20\,(AW - 0.80)/3 & \text{if } 0.8 \le AW \le 0.95 \end{cases}$$

$$\mu_{most}(AW) = \begin{cases} 20\,(0.95 - AW)/3 & \text{if } 0.8 \le AW \le 0.95 \\ 1.0 & \text{if } 0.6 \le AW \le 0.80 \\ 10\,(AW - 0.5) & \text{if } 0.5 \le AW \le 0.6 \end{cases}$$

$$\mu_{some}(AW) = \begin{cases} 10\,(0.6 - AW) & \text{if } 0.5 \le AW \le 0.6 \\ 1.0 & \text{if } 0.3 \le AW \le 0.5 \\ 10\,(AW - 0.2) & \text{if } 0.2 \le AW \le 0.3 \end{cases}$$

$$\mu_{little}(AW)=\begin{cases} 10\,(0.3 - AW) & \text{if } 0.2 \le AW \le 0.3 \\ 1.0 & \text{if } 0.02 \le AW \le 0.2 \\ 100(AW - 0.01) & \text{if } 0.01 \le AW \le 0.02 \end{cases}$$

$$\mu_{none}(AW)=\begin{cases} 100\,(0.02 - AW) & \text{if } 0.01 \le AW \le 0.02 \\ 1.0 & \text{if } 0.0 \le AW \le 0.01 \end{cases}$$

In the same way, the fuzzy set functions for directional querying can be described as:

$$\mu_{directly}(NW)=\begin{cases} 1.0 & \text{if } 0.95 \le NW \le 1.0 \\ 20\,(NW - 0.80)/3 & \text{if } 0.8 \le NW \le 0.95 \end{cases}$$

$$\mu_{mostly}(NW)=\begin{cases} 20\,(0.95 - NW)/3 & \text{if } 0.8 \le NW \le 0.95 \\ 1.0 & \text{if } 0.6 \le NW \le 0.80 \\ 10\,(NW - 0.5) & \text{if } 0.5 \le NW \le 0.6 \end{cases}$$

$$\mu_{somewhat}(NW)=\begin{cases} 10\,(0.6 - NW) & \text{if } 0.5 \le NW \le 0.6 \\ 1.0 & \text{if } 0.3 \le NW \le 0.5 \\ 10\,(NW - 0.2) & \text{if } 0.2 \le NW \le 0.3 \end{cases}$$

$$\mu_{slightly}(NW)=\begin{cases} 10\,(0.3 - NW) & \text{if } 0.2 \le NW \le 0.3 \\ 1.0 & \text{if } 0.02 \le NW \le 0.2 \\ 100(NW - 0.01) & \text{if } 0.01 \le NW \le 0.02 \end{cases}$$

$$\mu_{not}(NW)=\begin{cases} 100\,(0.02 - NW) & \text{if } 0.01 \le NW \le 0.02 \\ 1.0 & \text{if } 0.0 \le NW \le 0.01 \end{cases}$$

Unlike classical set theory that can describe membership to a set clearly, in fuzzy set theory membership of a term to a set is partial, i.e., a term belongs to a set with a certain grade of membership. Although it solves the gap problem in classical set expression, a new problem is coming. Because a common feature of the fuzzy sets is overlapping, the qualifiers may be associated with two different terms at the intersect intervals. For instance, the topological qualifier TQ may take '*all*' and '*most*' at the same time. This reveals uncertainty – the lack of adequate and correct information to make a decision.

3.2. Uncertainty Consideration

Uncertainty is an inevitable problem in GIS. In this paper, we devote ourselves to explore an approach that can perform the fuzzy querying under uncertainties. The study exemplifies whether the fuzzy set and certainty factor can incorporate in spatial querying.

Uncertainty occurs when one is not absolutely certain about a piece of information. Given AW=0.90, the fuzzy querying may give the following querying phrase:

__All__ of Object A overlaps Object B
__Most__ of Object A overlaps Object B.

How do we make the decision according to the information? Which querying information is reliable?

This reveals important deficiencies in areas such as the reliability of queries and the ability to defect inconsistencies in the knowledge. Because we cannot be completely certain that some qualifiers are true or others are false, we construct a certainty factor (CF) to evaluate the degree of certainty. The degree of certainty is usually represented by a crisp numerical value one, a scale from 0 to 1. A certainty factor of 1 indicates that it is very certain that a fact is true, and a certainty factor of 0 indicates that it is very uncertain that a fact is true. Some key ideas relevant to the determination the CF are discussed as following.

__Case 1.__ Considered a single qualifier for each querying

This is a case in which only on qualifier associated with a single object is involved in each querying result such as:

__All__ of Object A overlaps Object B
Object A is __directly__ west of Object B.

Where the fuzzy topological qualifier TQ_A ='all' which is associated with the object A; the fuzzy directional qualifier DQ_A ='directly' which is associated with the object A.

- If the qualifier only takes one term at given interval, the grade of membership μ( ) can be used as a CF that represents the degree of belief. The results will look like:

  __All__ of Object A overlaps Object B
  with $CF=\mu_{all}(AW_{ai}=0.99)=1.0$

  Object A is __directly__ west of Object B
  with $CF=\mu_{directly}(NW_{ai}=0.99)=1.0$

Where AW_{ai} is the area weight of a sub-group associated with object A; NW_{ai} is the node weight of a subgroup associated with object A; and $i, j\in I[1, 8]$, I represents an integer set.

- If the given weight is in the overlapping area, two qualifiers will be related. For example, the fuzzy topological qualifier of the object A takes both 'all' and 'most'. The querying results will be:

  __All__ of Object A overlaps Object B
  __Most__ of Object A overlaps Object B

It is acceptable if we take the qualifier that has a larger grade of membership. The certainty factor can be determined by the maximum value, that is,

$$CF = \max\{\mu_{all}(AW_{ai}=0.90), \mu_{most}(AW_{ai}=0.90)\} = \mu_{all}(AW_{ai}=0.90).$$

The final querying results should be

All of Object A overlaps Object B
with CF= μ_{all} (AW_{ai} =0.90).

As a result, the CF in case 1 can be obtained by

CF=max{μ_{TQk}(AW_{*i} =const), k∈I[1,5], i∈I[1,8] }
CF=max{μ_{DQk}(NW_{*j} =const), k∈I[1,5], j∈I[1,8] }
Where
TQ_k is a topological qualifier such as all;
DQ_k is a directional qualifier such as directly;
AW_{*i} is an area weight associated object i-node
NW_{*j} is a node weight associated object j-node
* is used to represent object A or B

Case 2. Considered multiple qualifiers

In the querying results, many pieces of fuzzy terms are conjoined (i.e. they are joined by AND), or disjoined (i.e. joined by OR). The examples of these types of queries are as follows:

Most of Object A overlaps some of Object B
Some of Object A is slightly south of Object B.

Hence, to perform these kinds of queries, we have to handle multiple fuzzy qualifiers. It is easy to understand that the relationship between different object qualifiers is conjunction, and the relationship between the same object qualifiers is disjoined. According to the fuzzy set theory, the conjunction and disjunction of fuzzy term can be respectively defined as the minimum and maximum of the involved facts. Therefore, the certainty factor contained multiple qualifiers can be determined by the following formulas:

Consider topological relationships
CF=min{ max{μ_{TQka}(AW_{ai} =α) },
max{μ_{TQkb}(AW_{bj} =α) },
where ka, kb∈I[1,5] & i,j∈I[1,8] }
Note: a topological qualifier TQ_1=all if ka=1.

Consider topological/directional relationships
CF=min{ max{μ_{TQka}(AW_{ai} =α) },
max{μ_{DQka}(NW_{aj} =β) },
where ka∈I[1,5] & i,j∈I[1,8] }
where a_i, b_j represent object node associated with object A and B, respectively,
α and β are constant.

As seen above, an approach in which the fuzzy set and uncertainty can incorporate to perform the fuzzy queries is developed.

4. FuzzyCLIPS Implementation

FuzzyCLIPS is an enhanced version of CLIPS developed at the National Research Council of Canada to allow the implementation of fuzzy expert systems. The modifications made to CLIPS contain the capability of handling fuzzy concepts and reasoning. It allows any mix of fuzzy and normal terms, numeric-comparison logic controls, and uncertainties in the rule and facts. By using FuzzyClips, it is easy for us to deal with fuzziness in approximate reasoning, to manipulate uncertainty in the rules and facts.

In the process of our implementation, all fuzzy variables are predefined with the *deftemplate* statement. This is an extension of the standard deftemplate construct in CLIPS. For example, fuzzy variables (qualifiers) can be declared in *deftemplate* constructs as following:

```
(deftemplate TFVariable
   0 1; define the fuzzy variable area-weight
  ((all  (0.8 0.0) (0.95 1.0) (1.0 1.0) )
   (most (0.5 0.0) (0.6 1.0) (0.8 1.0) (0.95 0.0))
   (some (0.2 0.0) (0.3 1.0) (0.5 1.0) (0.6 0.0))
   (little(0.01 0.0) (0.02 1.0) (0.2 1.0) (0.3 0.0))
   (none (0.0 1.0) (0.01 1.0) (0.02 0.0) )  ) )

(deftemplate DFVariable
   0 1; define the fuzzy variable node-weight
  ((directly (0.8 0.0)(0.95 1.0)(1.00 1.0))
   (mostly  (0.5 0.0)(0.60 1.0)(0.80 1.0)(0.95 0.0))
   (somewhat (0.2 0.0)(0.30 1.0)(0.50 1.0)(0.60 0.0))
   (slightly (0.01 0.0)(0.02 1.0)(0.20 1.0)(0.30 0.0))
   (not    (0.0 1.0)(0.01 1.0)(0.02 0.0)) ) )
```

A number of commands supplied in FuuzyCLIPS are very helpful for user to access fuzzy components that they need. In our application, when the weights (fuzzy variables) are calculated, the only interested information is the value of the fuzzy set at the specified weight value. The command ***get-fs-value*** provides us a tool to access the value. The syntax of the command is:

(get-fs-value ?<fact-variable> <number>) or
(get-fs-value <integer> <number>) or
(get-fs-value <fuzzy-value> <number>),

where <number> is a value that must lie between the lower and upper bound of the universe of discourse for the fuzzy set. A simple example just look like:

```
(assert (TFVariable most) )
(defrule Get-CF
  ?f < - (TFVariable ?cf)
=> (printout t "CF for " ?cf " is = " (get-fs-value ?f AW)
          crlf) (retract ?f)
 )
```

5. Querying Examples

Given two objects A(1, 1) (7, 2) and B(2, 1)(9, 4). The previous works based on CLIPS will provide the following query information.

```
==========================================
       Query results of binary
        spatial relationships
==========================================
2D physical relations:  A |os| B.
Topological relations:  A {overlaps} B.
Directional relations:  A {South} B
                        A {South-West} B
                        A  {west } B

  Little of Object A is West of Object B
  Object A is slightly West of Object B
------------------------------------------
=> Little of A is slightly West of B.
```

Based FuzzyCLIPS, the querying results would be look like:

```
==========================================
Fuzzy Query results with certainty factor
==========================================
Topological information:
  83% of A overlaps 23.8% of B
  Most of A overlaps some of B with CF=0.778

Directional information:
  Little of A is West of B with CF = 1.0
  A is slightly West of B with CF = 1.0
------------------------------------------
    ⇨  Little of A  is slightly West of B
       with CF= 1.0
```

More details for analysis are provided as following. Table 3 shows part of quantitative information stored in nodes associated with object.

Table 3. Quantitative information

Object Name	Node	AW	NW
Object A	center	0.8333	
	west	0.1667	0.1667
Object B	center	0.2380	
	north	0.4762	0.2655
	:	:	:

From these data, we know

$AW_{a0} = 0.8333$, $TQ \rightarrow \{ \text{all, most} \}$;
$AW_{a7} = 0.1667$, $TQ \rightarrow \{ \text{little} \}$;
$AW_{b0} = 0.4762$, $TQ \rightarrow \{ \text{some} \}$;
$NW_{a7} = 0.1667$, $DQ \rightarrow \{ \text{slightly} \}$.

$$\left.\begin{array}{l} \left.\begin{array}{l} \mu_{all}(AW_{a0}=0.8333) = 0.222 \\ \mu_{most}(AW_{a0}=0.8333) = 0.778 \end{array}\right\} \max=0.778 \\ \mu_{some}(AW_{b0}=0.4762) = 1.0 \end{array}\right\} \min=0.778$$

$$\left.\begin{array}{l} \mu_{little}(AW_{a0}=0.1667) = 1.0 \\ \mu_{slightly}(NW_{b0}=0.1667) = 1.0 \end{array}\right\} \min = 1.0$$

6. Conclusion

In a real world, fuzziness and uncertainty can occur simultaneously. To improve spatial querying accuracy, our research investigates an inexact inferencing approach that can perform fuzzy querying under uncertainty. The reliability of querying information is judged by a certainty factor (CF). The improved fuzzy querying is very flexible, and it can return spatial information in a wider variety of forms.

7. Acknowledgement

This research was supported in part by the Naval Research Laboratory's Base Program, Program Element Number 0602435N.

References

[1]. E. Takahashi, N. Shima, and F. Kishino, "An Image Retrieval Method Using Inquires on Spatial Relationships", *Journal of Information Processing*, 15(3), 1992, pp441-449.

[2]. S. Winter, "Topological Relations between Discrete Regions", *SSD'95*, Portland, ME, USA, August 1995, pp310-327.

[3]. M. A. Cobb, "Modeling Spatial Relationships within a Fuzzy Framework", *Journal of the American Society for Information Science*, 49(3): 253-266, 1998.

[4]. M. A. Cobb, "An approach for the Definition, Representation and Querying of Binary Topological and Directional Relationships Between two-dimensional Objects", PhD thesis, Tulane University, 1995.

[5]. M. A. Cobb, F. E. Petry, "Fuzzy Querying of Binary Relationships in Spatial Databases", 1995 IEEE, pp3624-3629

[6]. H. Yang, M. Cobb, K. Shaw, "A Clips-Based Implementation for Querying Binary Spatial Relationships", Proceedings of Joint 9th IFSA World Congress and 20th NAFIPS International Conference, Vancouver, British Columbia, Canada. July 25-28, 2001.

[7]. M. Goodchild, S. Gopal, "The accuracy of Spatial Databases", Basingstoke, UK: Taylor and Francis, 1990.

[8]. S. D. Bruin, "Querying Probabilistic Land Cover Data Using Fuzzy Set Theory", *GIS*, Vol 14, No.4.

[9]. L.A. Zadeh, "Fuzzy Sets", *Information and Control*, 8:338-353, 1965.

[10]. P. Matsakis, J.M. Keller, L. Wendling, J. Marjamaa, and S. Sjahputera, "Linguistic Description of Relative Positions of Objects in Images", *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 31, No. 4, 2001, pp573-588.