

Objects with Fuzzy Spatial Extent

Tao Cheng and Martien Molenaar

Abstract

The determination of the spatial extent of geo-objects is generally approached through their boundaries or, more precisely, through the positions of their boundary points. The analysis of the geometric uncertainty of the objects is therefore often based on accuracy models for the coordinates of these points. In many survey disciplines objects are mapped, however, that are not crisp well defined. In that case, the geometric uncertainty is not only a matter of coordinate accuracy, but also a problem of object definition and thematic vagueness. The spatial uncertainty of such objects cannot be handled by a geometric approach alone, such as the epsilon band method. This paper investigates the reasons for the fuzzy spatial extent of objects and proposes an approach to map the spatial extent of objects and their uncertainties when objects are extracted from field observation data. The relationship of uncertainties between thematic aspects and geometric aspects is investigated. A practical example of a coastal geomorphology study is discussed to illustrate the approach.

Introduction

Spatial objects that are represented in a conventional GIS are generally considered to be crisp with determined boundaries. For example, the land parcels in cadastral systems are differentiated and identified by sharp boundaries. The basic assumption is that the classification of landscape units is crisp and spatial objects within these classes can be clearly determined. The second assumption is that objects are internally homogeneous and can be differentiated by crisp boundaries. Under the first assumption the threshold values or criteria for classification are sharply defined. Classes do not overlap; thus, each object will be assigned to only one class. Under the second assumption, the spatial extent of each object can be defined unambiguously and it will not contain unidentified inclusions of areas not belonging to the object. The determination of the spatial extent of geo-objects is then generally approached through their boundaries or, more precisely, through the positions of their boundary points. The analysis of the geometric uncertainty of the objects is therefore often based on accuracy models for the coordinates of these points. The epsilon band method is well known in this context (Dunn *et al.*, 1990).

These assumptions, however, are not valid when the spatial extents of objects are to be extracted from field data that change gradually and continuously over space so that no crisp boundaries can be identified. The boundary between a grassland and a woodland may be gradual through a transition zone rather than sharply being defined. In such cases, the geometric uncertainty of objects cannot be expressed through the position accuracy of boundary points. It is then not only a problem of

geometry, but it is rather a problem of object definition and thematic vagueness. For example, in interpretation of remote sensing images, uncertainty exists in the thematic aspect expressed by the likelihood of pixels belonging to thematic classes. Image segments can then be formed from adjacent pixels falling within the same class. If these segments represent spatial objects, then the uncertainty of the extracted geometry of these objects is mainly due to the fact that the value of the likelihood function varies per pixel. Therefore, existing solutions for handling the uncertainty of objects have not been found satisfactory.

The syntactic approach for handling spatial object information as presented in Molenaar (1994; 1996; 1998) makes it possible to distinguish three types of statements with respect to the existence of spatial objects:

- an *existential statement* asserting that there are spatial and thematic conditions that imply that an object exists;
- an *extensional statement* identifying the geometric elements describing the spatial extent of the object; and
- a *geometric statement* identifying the actual shape, size, and position of the object in a metric sense.

These three types of statements are intimately related. The extensional and geometric statements imply the existential statement; thus, if an object does not exist, it cannot have a spatial extent and geometry. The existential statement often relates to the uncertainty of thematic information, though that is not explicit in the other two statements. The geometric statement also implies the extensional statement, and often the actual geometry of the object is derived from the extensional description. Object detection through image interpretation is, in fact, an example of the formulation of extensional statements. These three types of statements can all have a degree of uncertainty and, although these statements are related, they give us different perspectives that may help us to understand the different aspects of uncertainty in relation to the description of spatial objects.

In this paper we will concentrate on the uncertainty related to the extensional statement. The remainder of the paper is organized as follows. The next section introduces a syntactic schema to represent the extensional uncertainty in terms of the fuzzy spatial extent of objects. The approach to extract objects from field data is proposed in the third section. It first explains the reasons for indeterminate boundaries. Then it investigates the conversion of uncertainties from thematic aspects to geometric aspects during the identification of the fuzzy spatial extent of objects. Finally, the forming of conditional boundaries and their syntactic representation is discussed. The proposed approach is illustrated by a case in the fourth section. The test area, data collection, data preparation, and modeling

International Institute for Aerospace Survey and Earth Sciences (ITC), P.O. Box 6, 7500AA Enschede, The Netherlands ([cheng, molenaar]@itc.nl).

Tao Cheng is currently with the Joint Laboratory for Geoinformation Science, The Chinese University of Hong Kong, Shatin, NT, Hong Kong.

Photogrammetric Engineering & Remote Sensing
Vol. 65, No. 7, July 1999, pp. 797-801.

0099-1112/99/6507-797\$3.00/0

© 1999 American Society for Photogrammetry
and Remote Sensing

results are presented. The last section of the paper summarizes the major findings and further research.

Representation of Fuzzy Spatial Extent

Molenaar (1994; 1998) proposed a syntactic representation of objects, which can represent the vector data and raster data in a unified way.

Let's assume that the geometry of a spatial database (shortly, a map M) has the topologic structure of a planar graph. Each edge will always have one face at its left-hand side and one on its right-hand side. The relationship can be expressed by following functions:

$$\begin{aligned} \text{Edge } e_i \text{ has a face } f_a \text{ at its left-hand side} &\Rightarrow Le[e_i, f_a] \\ &= 1, \text{ otherwise } = 0. \end{aligned}$$

$$\begin{aligned} \text{Edge } e_i \text{ has a face } f_a \text{ at its right-hand side} &\Rightarrow Ri[e_i, f_a] \\ &= 1, \text{ otherwise } = 0. \end{aligned}$$

With this function, we can define $B[e_i, f_a] = Le[e_i, f_a] + Ri[e_i, f_a]$.

If the value of this function equals 0, e_i is not related to f_a ; if the value equals 2, then the edge has a face on both sides and is thus inside the two faces; if the value equals 1, then the edge has a face on only one side, so that it must be part of the boundary. The boundary of a face f_a is then defined as

$$B\partial f_a = \{N\partial f_a, E\partial f_a\}$$

where $E\partial f_a = \{e_i | B[e_i, f_a] = 1\}$ represents all the edges that have the face f_a on only one side and $N\partial f_a = \{n_i | n_i \in e_i, e_i \in E\partial f_a\}$ represents the nodes of the edges of $E\partial f_a$.

The relationship between a face and an area object can be represented as $Part[f_a, O_a]$. If it takes a value of 1, it implies that the face belongs to the object. If it takes a value of 0, it implies that the face doesn't belong to the object.

Therefore, the relationship between edge of an object can be expressed as follows:

$$Le[e_i, O_a] = Le[e_i, f_a] * Part[f_a, O_a],$$

$$Ri[e_i, O_a] = Ri[e_i, f_a] * Part[f_a, O_a].$$

The boundary of an object is then defined as

$$B[e_i, O_a] = Le[e_i, O_a] + Ri[e_i, O_a]. \quad (1)$$

If the value is equal to 0, the edge is not related to the object; if the value is equal to 2, then the edge is inside the object; if it has a value of 1, the edge is the boundary of the objects, i.e., $B\partial O_a = \{e | B[e, O_a] = 1\}$.

If the object is fuzzy in the sense that its spatial extent is uncertain, the relationship of the fact and the object is uncertain, which can be expressed as

$$Face(O_a) = \{f_a | Part[f_a, O_a] > 0\}. \quad (2)$$

The boundary of a fuzzy object can be defined in the same manner as in Equation 1. But the value of the function will vary from 0 to 2. If the value is equal to 0, the edge is not related to the object; if the value is bigger than 0 and less than 2, the edge is an indeterminate boundary of the object; if the value is equal to 2, the edge is inside the object. Therefore, the indeterminate boundary of an object is defined as

$$B\partial O_a = \{e | 0 < B[e, O_a] < 2\} \quad (3)^*$$

Based upon certain criteria, the conditional spatial extent of objects can be identified, so the indeterminate boundary becomes a conditional boundary. This syntactic representation of the relationship between edges, boundaries, faces, objects are can be applied for both vector and raster structures. The cells of a raster are then considered as faces with a rectangular shape (Molenaar, 1994; Molenaar, 1998).

Extraction of Fuzzy Objects

Procedure of Extraction of Objects from Field Observation Data

When natural phenomena have a field character, they can often only be sampled sparsely at a limited number of points, which can then be interpolated to generate a full raster. In the case described in the fourth section, this will be an elevation raster. Three types of objects will be extracted from these height data through a segmentation of the elevation raster: the foreshore, the beach, and the foredune areas. Each type of object will be related to a height interval. A six-step procedure will be followed to identify objects from sampled field data:

- (1) **Sampling** data values at specific sample points.
- (2) **Interpolation**, also called "regionalization," of the observed data to generate a complete elevation raster covering the observed area.
- (3) **Classification** of all grid cells into pre-defined classes. Each grid cell is assigned to a class interval that can be related to one of the natural units.
- (4) **Segmentation** of the classified raster into areas. Each contiguous set of grid cells belonging to one class will form an area that represents the spatial extent of a particular natural unit.
- (5) **Merging** areas that are smaller than a pre-defined lower threshold for mapping units with an adjacent area. Traditional merging methods, such as "window filtering," "nibbling," "dropping the longest shared boundary," and "maximum area merging" (Ma and Zhao, 1995) can be used to remove these small areas.
- (6) **Identification** of objects represented by the areas, i.e., identification of the actual objects whose spatial extents are represented by the final segments (after merging).

Figure 1 illustrates the procedure by means of a crisp example. The steps represented in this figure start with the interpolated grid cells after Step 2. In Step 3 the grid cells are classified into three elevation classes: "H" (high), ranging from 15 through 20; "M" (medium), ranging from 10 through 14; and "L" (low), ranging from 5 through 9. The segmentation of Step 4 identifies three areas. In Step 5, Area 3 has been merged into Area 2. Finally, two objects—"A" and "B"—are identified in Step 6.

Identification of Spatial Extent and Boundary of Fuzzy Objects

As presented in the previous subsection, classification and segmentation are essential for extracting objects from field observation data. The classification of grid cells is uncertain for two reasons:

- The height values of the grid cells do have a limited accuracy due to the measurement and interpolation process; and
- The height classes related to the three object types are fuzzy, as will be explained in the fourth section.

The combination of these two kinds of uncertainties has been discussed in Cheng *et al.* (1997). The uncertain classification can be expressed in terms of the membership function

*The concept of a crisp boundary of an object is only valid when the face belongs to the object ($Part[f, O] = 1$) and the edge has the face on either the left or right side. Therefore, Equation 3 is valid only for fuzzy objects.

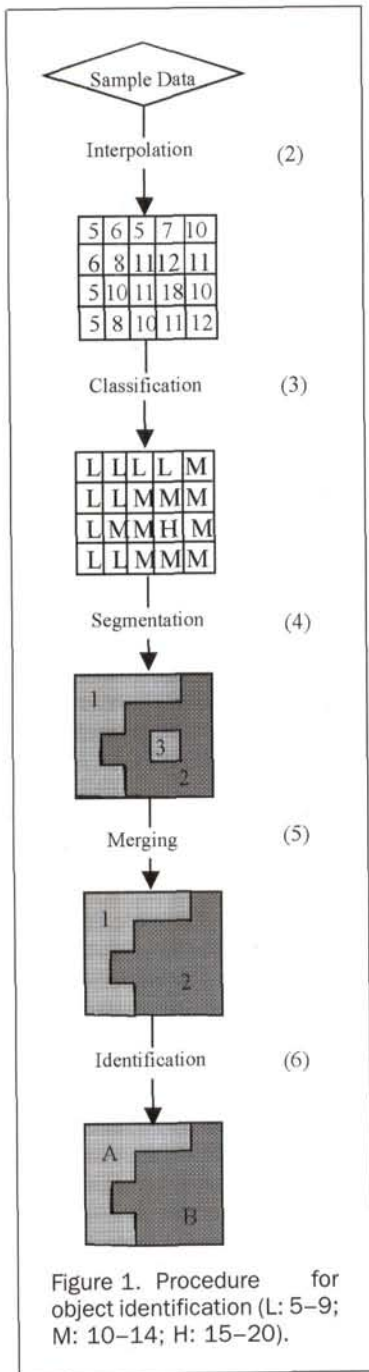


Figure 1. Procedure for object identification (L: 5-9; M: 10-14; H: 15-20).

assumed to be spatially exclusive, then each grid cell belongs to at most one class, and if the set of identified objects forms a spatial partition of the mapped area, then each grid cell belongs to exactly one object. In other applications, fuzzy spatial overlaps among objects are permitted, i.e., the spatial extent of objects cannot be expressed through crisp boundaries that also carry adjacency information, but the objects have fuzzy transition zones that may overlap (Burrough, 1996; Usery, 1996). In the transition zones, the pixels might belong to multiple objects. The fuzzy topologic relationships of spatial objects are discussed in Dijkmeijer and De Hoop (1996) and Zhan (1997). In our case, the landscape units form spatial partitions. So each grid cell should belong to exactly one class and therefore to one object, which can be determined by criteria such as we will define here.

Let $NM[P_{ij}, C_k] = 1 - MF[P_{ij}, C_k]$ represent no-membership, i.e., the certainty that P_{ij} does not belong to class C_k , and let $XM[P_{ij}, C_k]$ express the membership that P_{ij} belongs exclusively to C_k and not to any other classes C_l for any $l \neq k$. $XM[P_{ij}, C_k]$ can be derived by applying minimum operations as

$$XM[P_{ij}, C_k] = \text{MIN}(MF[P_{ij}, C_k], \text{MIN}_{l \neq k} (NM[P_{ij}, C_l])). \quad (4)$$

Because P_{ij} can only belong to one class, only one class is required for which the function $XM[]$ has a maximum value for P_{ij} . If there are more classes with the same maximum values, then additional evidence is required in order to arrive at a selection of a unique class. It can be represented as

$$\begin{aligned} \text{if } XM[P_{ij}, C_k] &= \text{MAX}_{C_l} (XM[P_{ij}, C_l]) \\ (l = 1, \dots, N), \text{ then let } D[P_{ij}, C_k] &= 1; \quad (5) \\ \text{otherwise, } D[P_{ij}, C_k] &= 0. \end{aligned}$$

For example, in our case a grid cell has the membership vector

$$MF[P, C] = \begin{Bmatrix} 0.2 \\ 0.7 \\ 0.1 \end{Bmatrix}$$

where C_1 is a foreshore class, C_2 is a beach class, and C_3 is a fore-dune class. Therefore,

$$NM[P, C] = \begin{Bmatrix} 1-0.2 \\ 1-0.7 \\ 1-0.1 \end{Bmatrix} = \begin{Bmatrix} 0.8 \\ 0.3 \\ 0.9 \end{Bmatrix}$$

$$\text{and } XM[P, C] = \begin{Bmatrix} \text{MIN}(0.2, \text{MIN}(0.3, 0.9)) \\ \text{MIN}(0.7, \text{MIN}(0.8, 0.9)) \\ \text{MIN}(0.1, \text{MIN}(0.8, 0.3)) \end{Bmatrix} = \begin{Bmatrix} 0.2 \\ 0.7 \\ 0.1 \end{Bmatrix}$$

Because

$$XM[P, C_2] = \text{MAX} \begin{Bmatrix} 0.2 \\ 0.7 \\ 0.1 \end{Bmatrix} = 0.7, \text{ therefore } D[P, C_2] = 1.$$

This means that this cell is assigned to class C_2 (the beach area) with a certainty of 0.7.

After assigning the cells to classes, an area S_a of class type C_k will be formed by the following two conditions (Molenaar, 1996):

$$\begin{aligned} \text{for all pixels } P_{ij} \in S_a, D[P_{ij}, C_k] &= 1, \text{ and} \\ \text{if } P_{ij} \in S_a \text{ and } \text{ADJACENT}[P_{kl}, P_{ij}] &= 1 \quad (6) \\ \text{and } D[P_{ij}, C_k] &= 1, \text{ then } P_{ij} \in S_a. \end{aligned}$$

value that varies per cell and is less than one. To identify the spatial extent, we have to first assign the grid cells into classes and then cluster the cells of the same classes into areas which represent the spatial extents of fuzzy objects. Here we will discuss the effect of fuzzy classification on identification of the spatial extent of objects. We will also investigate how existential uncertainty is converted into extensional and geometric uncertainties.

A membership vector $[MF[P_{ij}, C_1], MF[P_{ij}, C_2], \dots, MF[P_{ij}, C_N]]^T$ ($0 \leq MF[P_{ij}, C_k] \leq 1$) will be created for each grid cell P_{ij} after fuzzy classification. Here, $MF[P, C]$ represents the membership function value of grid cell P_{ij} belonging to class C , and N is the total number of the class types.

For each class C_k , areas can be identified with $MF[P_{ij}, C_k] > \text{Threshold}_k$. These areas can then be interpreted as the fuzzy extent of spatial objects belonging to C_k . If the classes are

$ADJACENT[P_{kl}, P_{ij}]$ expresses the adjacency relationship between cells P_{kl} and P_{ij} , and it has a value of either 0 or 1. P_{ij} will only be assigned to S_a if " $D[P_{ij}, C_k] = 1$." The certainty that this assignment is correct depends on the certainty that the pixel has been assigned correctly to C_k . Therefore, the relationship between P_{ij} and S_a , $Part[P_{ij}, S_a]$, can be written as

$$Part[P_{ij}, S_a] = MIN(D[P_{ij}, C_k], XM[P_{ij}, C_k]). \quad (7)$$

Because S_a represents the spatial extent of the object O_a , the relationship between P_{ij} and O_a can be defined as (Cheng, 1999)

$$Part[P_{ij}, O_a] = Part[P_{ij}, S_a] = MIN(D[P_{ij}, C_k], XM[P_{ij}, C_k]). \quad (8)$$

Equation 8 expresses the relationship between the uncertainty of a pixel belonging to the spatial extent of an object and the uncertainty of a pixel belonging to classes, i.e., the relationship between the geometric uncertainty and the thematic uncertainty. This means that the uncertainty is transferred from thematic aspects to geometric aspects of objects during spatial clustering, i.e., the *existential uncertainty is converted to extensional uncertainty*.

Because the spatial extent is fuzzy, there is no crisp boundary between objects. After identification of the spatial extent of objects, i.e., assigning pixels to areas, boundaries are formed. We call them *conditional boundaries to distinguish them from crisp boundaries*. In our case, the conditional boundary between two objects is the transit boundary between two classes.

According to the Equation 1, the boundary of an object consists of edges that have the object on one side. To check if the edge has an object on one side, the relationship between the edge and the faces belonging to the object should be checked. Therefore, to identify the conditional boundary of a fuzzy object, we have to first identify the faces that belong to the object, and then find the edges that have faces on only one side.

The face of two fuzzy objects should satisfy

$$Face(O_a) = Cell(O_a) = \{P_{ij} | Part[P_{ij}, O_a] > Part[P_{ij}, O_b]\} \quad (9)$$

and

$$Face(O_b) = Cell(O_b) = \{P_{ij} | Part[P_{ij}, O_b] > Part[P_{ij}, O_a]\}.$$

Then the transition boundary consists of edges that have simultaneously the cells of O_a on the left side and the cells of O_b on the right side, or the cells of O_a on the right side and the cells of O_b on the left side. Therefore, the edges of the boundary should satisfy (Cheng, 1999)

$$\begin{aligned} E_{a,b} &= \{e_i | B[e_i, f_a] = 1 \\ &\text{and } B[e_i, f_b] = 1 \\ &\text{and } f_a \in Face(O_a) \text{ and } f_b \in Face(O_b)\}. \end{aligned}$$

Then the transition boundary is

$$B\partial(O_a, O_b) = \{N_{a,b}, E_{a,b}\} \text{ and } N_{a,b} = \{n_i | n_i \in e_i, e_i \in E_{a,b}\}.$$

The Case

Ameland is a barrier island in the north of The Netherlands, where geomorphologic processes occur along the coast, particularly the erosion and accumulation of sediments. These processes can be monitored through the observation of changes of landscape units such as foreshore, beach, and foredune. The foreshore is the area above the closure depth and beneath the low water line, the beach is the area above the low water line and beneath the dune foot (Reineck and Singh, 1980), and the

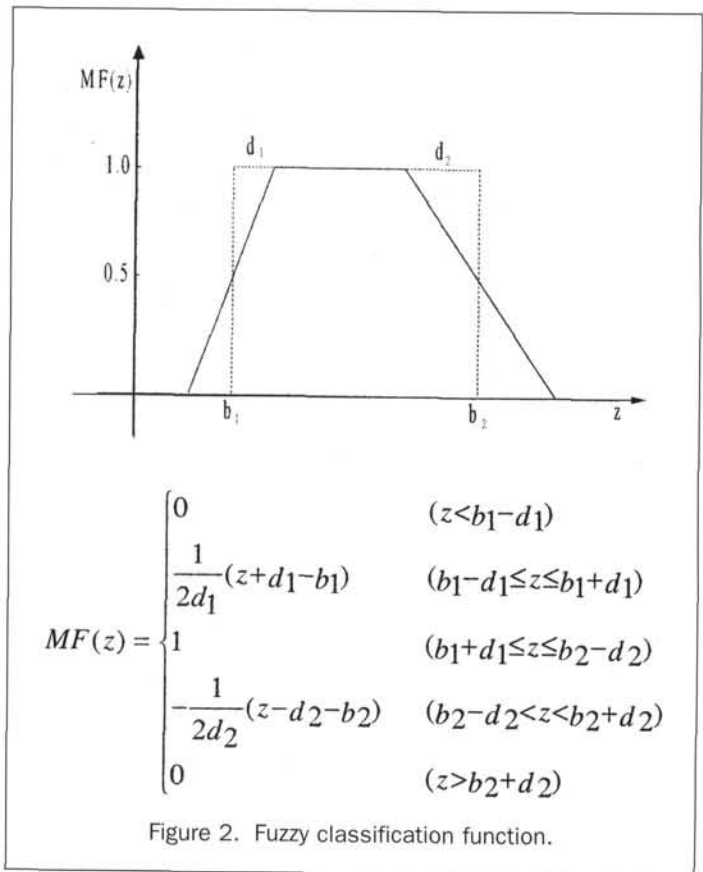


Figure 2. Fuzzy classification function.

foredune is the first row of the dunes inland from the dune foot. According to the definitions given by geomorphologists for the situation of Ameland, the height values of the closure depth, low water line, and dune foot are suggested to be about -6.0m, -1.1m, and 2m, respectively. These definitions appear to be vague because the experts do not agree exactly on these values. Therefore, we adopted a trapezoidal membership function as illustrated in Figure 2 and define the transition zone between the classes related to these landscape units as in Table 1 (Cheng *et al.*, 1997).

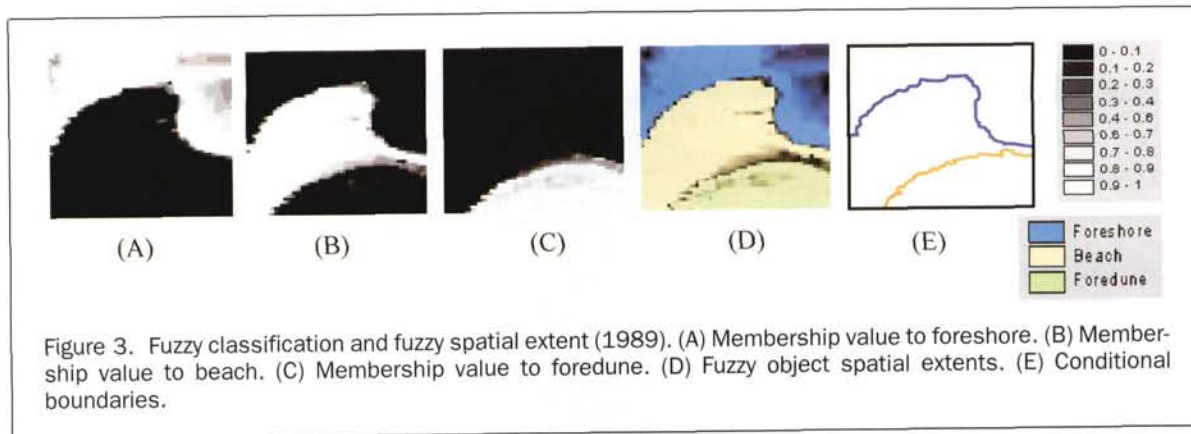
Height observations have been made by laser scanning of the beach and dune area and by echo sounding at the foreshore. These data have been interpolated to form a full height raster of the test area. Experiments show that the uncertainty of the interpolated heights of the raster can be expressed by their standard deviation ($\sigma = 0.15\text{m}$) (Huisling *et al.*, 1996).

As shown in Plates 1A, 1B, and 1C, each grid cell has a membership vector containing a value for each of the three classes. After identifying the most likely class type for each cell, the mapped area has been segmented by clustering the cells belonging to the same class. The areas of different classes represent the fuzzy spatial extent of the objects, which are shown in Plate 1D. The transit boundaries are shown in Plate 1E.

TABLE 1. FUZZY DEFINITION FOR COASTAL LANDSCAPE UNITS.

Class Code	Landscape Unit	b_1 (m)	b_2 (m)	d_1 (m)	d_2 (m)
1	Foreshore	-6.0	-1.1	2.0	0.5
2	Beach	-1.1	2.0	0.5	0.5
3	Foredune	2.0	25.0	0.5	3.0

Note: b_1 and b_2 represent the cross points of the landscape units; d_1 and d_2 represent the half width of transition zones.



Conclusions

This paper discussed three uncertainty aspects of spatial objects. The emphasis was on the extensional uncertainty of fuzzy spatial objects for which the determination of a crisp boundary was not possible or meaningful. The spatial extent of such objects is extracted from field data, in this case height data, which had limited accuracy and were assigned to fuzzy. A syntactic representation was applied to represent fuzzy objects and their fuzzy spatial extent. A procedure to extract spatial extent and conditional boundaries of objects was proposed, which was illustrated by a coastal geomorphology case.

Conventionally, people first see the boundary of an object and then see the spatial extent of an object. However, when there are no crisp boundaries between objects, the fuzzy spatial extent of objects should be identified first. The conditional boundaries can then be approached. In such cases, the uncertainty in thematic aspects is converted to the geometric aspects of the objects. Because most of existing studies on uncertainty of objects deal with boundary accuracy, this paper provides a different approach to deal with the uncertainty in geometric aspects. It is our hope that this paper will stimulate more interest in the further study of fuzzy objects.

Acknowledgments

The authors wish to acknowledge Marieke Eleveld for preparing the height data of the test area. We thank Theo Bouloucos for the stimulating discussions we had about several issues of this paper. The research is funded by the International Institute for Aerospace Survey and Earth Sciences (ITC) and the Survey Department of the Directorate-General for Public Works and Water Management of the Dutch Ministry of Transport, Public Works and Water Management.

References

- Burrough, P.A., 1996. Natural objects with indeterminate boundaries, *Geographic Objects with Indeterminate Boundaries* (P.A. Burrough and A.U. Frank, editors), Taylor & Francis, London, pp. 3-28.
- Cheng, T., 1999. *A Process-Oriented Data Model for Fuzzy Spatial Objects*, Ph.D. thesis, Wageningen/ITC, Enschede, 163 p.
- Cheng, T., M. Molenaar, and T. Bouloucos, 1997. Identification of fuzzy objects from field observation data, *Spatial Information Theory: A Theoretical Basis for GIS, Lecture Notes in Computer Sciences, Vol. 1329* (S.C. Hirtle and A.U. Frank, editors), Springer-Verlag, Berlin, pp. 241-259.
- Dijkmeijer, J., and S. De Hoop, 1996. Topologic relationships between fuzzy area objects, *Advances in GIS Research II* (M.J. Kraak and M. Mplenaar, editors), Taylor & Francis, London, pp. 377-394.
- Dunn, R., A.R. Harrison, and J.C. White, 1990. Positional accuracy and measurement error in digital databases of land use and empirical study, *International Journal of Geographic Information Systems*, 4(4):385-398.
- Huisling, E.J., R.W.L. Jordans, [remaining authors], 1996. *Framework Digital Elevation Models Dutch Coast - Various Techniques for Topographic Measurement Tested on a Common Site*, The Netherlands Remote Sensing Board (BCRS) Programme Bureau, Rijkswaterstaat Survey Department, Delft, 63 p.
- Ma, Z., and S. Zhao, 1995. Using an object and rule-based merging algorithm to eliminate small regions less than minimum mapping unit, *Geographic Information Sciences, A Journal of the Association of Chinese Professionals in Geographic Information Systems Abroad*, 1(2):110-117.
- Molenaar, M., 1994. A syntax for the representation of fuzzy spatial objects, *Advanced Geographic Data Modelling* (M. Molenaar and S. de Hoop, editors), Netherlands Geodetic Commission, Delft, New Series, No. 40, pp. 155-169.
- , 1996. A syntactic approach for handling the semantics of fuzzy spatial objects, *Geographic Objects with Indeterminate Boundaries* (P.A. Burrough and A. U. Frank, editors), Taylor & Francis, London, pp. 207-224.
- , 1998. *An Introduction to the Theory of Spatial Object Modeling*, Taylor & Francis, London, 246 p.
- Reineck, H.E., and I.B. Singh, 1980. *Depositional Sedimentary Environments, Second Edition*, Springer Verlag, Berlin, 549 p.
- Usery, E.L., 1996. A conceptual framework and fuzzy set implementation for geographic feature, *Geographic Objects with Indeterminate Boundaries* (P.A. Burrough and A.U. Frank, editors), Taylor & Francis, London, pp. 71-85.
- Zhan, F.B., 1997. Approximation of topological relationship between fuzzy regions satisfying a linguistically described query, *Spatial Information Theory: A Theoretical Basis for GIS, Lecture Notes in Computer Sciences, Vol. 1329* (S.C. Hirtle and A.U. Frank, editors), Springer-Verlag, Berlin, pp. 509-510.