Topological Operation Implements for Complex Geometry: A Rasterizing Scanline Approach

Dengji Cui, Mingguang Wu*

Key Laboratory of Virtual Geographic Environment of Ministry of Education, Nanjing Normal University, Nanjing, 210046, P.R.China.

***Abstract* —This paper uses raster scan line as the basic describing unit of point, line and polygon. The definition of seven basic topological relations of raster scan line, formalized definition of topological relations for geometric objects like point, line and polygon are given in the paper. And the line-by-line calculation flow based on topological relations of raster scan line is designed. The test indicates that this method is suitable for collection objects of point, line and polygon and has higher computational efficiency.**

Keywords - Topology; Raster Scan Line; Line-by-line calculation.

I. INTRODUCTION

The conceptual model of topological relations has stimulated wide-ranging study. And 4/9-intersection model [1], RCC model [2], 2D String model [3] and V9I model [4], etc. have been put forward. Comparing with researches in the aspect of formalized description of above topological relations, there are relatively lack of the researches in the terms of supporting algorithm of topological conceptual model, the efficiency of algorithm and stability of numerical calculation on geographical coordinates [5].

The main calculation methods of topological relations currently include analytic geometry algorithm, planar graph algorithm and transformation method of plane sweep, etc. Analytic geometry algorithm [6,7,8] often has high operational efficiency, but it is difficult to integrate with topological conceptual models like 4/9-intersection model, and then to calculate complex topological predicates. Transforming geometric objects into planar graph, planar graph algorithm refers to a large amount of intersection algorithms and has lower calculation efficiency. Besides, it is difficult to calculate mixed spatial relations due to the loss of information like distance and location, etc. in the process of transformation. The researches [9,10,11] of adopting plane sweep method to calculate topological relations has been extensively attracted. Scan conversion method transforms calculation of 2-dimensional topological relations into 0-dimensional pixel points [12], 1-dimensional scanning lines [13] or 2-dimensional polygons [14] to calculate. It can apparently decrease complexity of algorithm and get high calculation efficiency after conversion. However, this algorithm is always proposed for the calculation of topological relation between certain kinds of objects. The method in reference 12 is only fit for the calculation of polygon inclusion relations. And the method in reference 13 is only fit for intersection calculation between lines. But the method in reference 14 is only fit for the calculation of topological relation between polygons. So there is no efficient computing framework of topological relations for complex geometry at present.

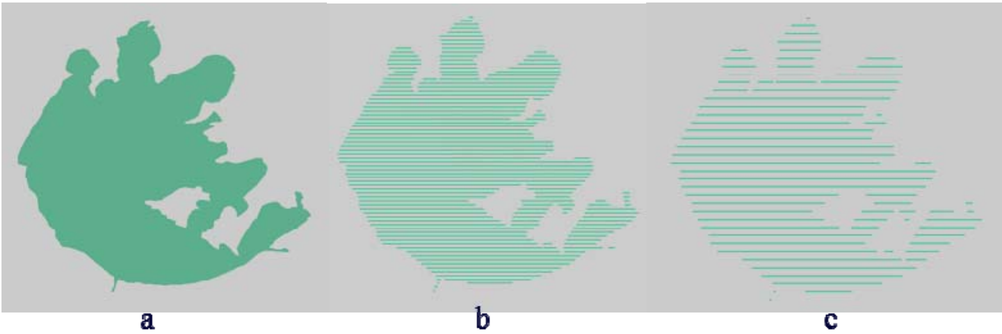
For the calculation of topological relations for complex geometry, this paper proposes one calculation method of topological relations based on raster scan line by learning from the thought of scan line in the process of vector graphics rasterization.

II. TOPOLOGICAL EXPRESSION AND LINE-BY-LINE CALCULATION BASED ON RASTER SCAN LINE

A. Graphic Expression Based on Raster Scan Line

The algorithm of raster scan line firstly assumes that spatial range which is attached to geometric object to be calculated is dispersed into square grid unit with same size. Supposing that one line *l* sweeps across the whole plane from top to bottom based on grid unit as spacing, starting from one location higher than the other graphics. In the process of assumed line scanning plane, intersections of scan line and geometric graphics are recorded. All intersections are sorted in the light of increasing orders of horizontal coordinates and matched according to the sequence of the first and second, the third and fourth, etc. One set of parallel and equally spaced scan lines can be obtained and represented geometric graphic after intersection, sorting and match. Compared with described scan line in references from 11 to 13, the stated scan line in this paper is composed of cell, so it is also called as raster scan line. Polygons in Fig.

1(a) are processed with two interval scanning, and there are two groups of raster scan lines as shown in Fig. 1 (b) and Fig. 1 (c) (to illustrate diagram conveniently, the widths of raster scan lines are narrowed to show different scanning intervals).



a b c

Fig.1 Two-Dimensional Eight Shape Model

Through scanning and conversion, raster scan line becomes basic unit of describing geometric object. The unit of raster scan line mainly includes row number, column number and length information of raster scan line. It can be described as triple <*i, j, l*>. Any objects point or line or polygon can be described as one alone set of raster scan line: $S=\{s_0, s1, s2,\ldots\}$. Only one raster scan line is included in set of point object and its attribute value l is 1; sets of raster scan lines with line or polygon object include more than 1 raster scan lines. Raster scan line is one data structure combined with vector and raster, which constructs on the basis of grid cell. Comparing with vector data structure:

1) Raster scan line only includes three attributes < *i, j, l* >. It has united data structure and can describe point, line and polygon targets. Comparing with original scan line, each attribute of triple is integer, which is beneficial for efficient calculation.

2) Operations of raster scan line like plus and subtraction are close operations for closing all domain of discourse.

3) The calculation of topological relationships of raster scan line like intersection, deviation and inclusion,etc only refers to integer Boolean operation.

Comparing with raster data structure:

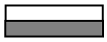
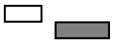
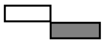
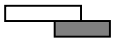
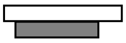
1) Raster scan line doesn't cover "blank" area and amount of data is small.

2) Raster scan line includes attribute l which can avoid data redundancy.

In this paper, raster scan line is basic unit of calculation of topological relations. Seven basic topological relations of raster scan line are concluded in Table I (in order to describe diagram conveniently, the raster scan lines A and B with same value j draw raster scan lines in different lines respectively).

DOI 10.5013/IJSSST.a.17.27.19 ISSN: 1473-804x online, 1473-8031 print

TABLE I. LIST OF BASIC TOPOLOGICAL RELATIONS OF RASTER SCAN LINE

Name	Relation	Diagram	Mathematical description
Equal	A equal B		$A_i=B_i,A_j=B_j,A_l=B_l$
Disjoint	A disjoint B		$(A_j-B_j>1 \Vert A_j-B_j<-1) \Vert A_i+A_L<B_i \Vert A_i>B_i+B_l$
Touch	A touch B		$A_j-B_j=-1,0,1 \&\& A_i+A_l=B_i$
Intersection	A Intersection B		$A_j=B_j \&\& A_i<B_i \&\& A_i+A_l>B_i \&\& A_i+A_l<B_{i+}B_l$
Inside Touch	A inside Touch B		$A_j-B_j=-1,0,1 \Vert (A_i=B_i \&\& A_i+A_l<B_{i+}B_{l)}) \Vert (A_{i+}A_l=B_i \&\& A_i<B_{i)})$
Contain	A contain B		$A_{j=}B_j \&\& A_i<B_i \&\& A_i+A_l>B_i \&\& A_i+A_l<B_{i+}B_l$
Within	A within B		$A_{j=}B_j \&\& A_i>B_i \&\& A_i+A_l<B_{i+}B_l$

B. Definition of Topological Relations Based on Raster Scan Line

Let A and B be objects of arbitrary point, line or polygon. S_A and S_B are represented as sets of raster scan line of objects A and B, C_a and C_b as number of raster scan lines in sets. a_i and b_j are elements in S_A、S_B. With the reference of Open GIS Simple Features standard [15], definition of topological relations is given (described by words with the beginning of capital);

Equal:

For two sets of raster scan line S_A and S_B, they have equal size and element with each corresponding serial number is equal. It can be described as:

$$C_a = C_b \quad and \quad \forall i, a_i \in S_A, b_i \in S_B \{equal(a_i, b_i) = ture\} \tag{1}$$

Disjoint:

Within two sets of raster scan line S_A and S_B, any two raster scan lines have disjoint relation. It can be described as:

$$\forall i, j, a_i \in S_A, b_j \in S_B \{disjoint(a_i, b_j) = ture\} \tag{2}$$

Touches:

Within two sets of raster scan line S_A and S_B, there exists more than one touch relation without relations such as intersection, interior and inclusion, etc. It can be described as:

$$\begin{cases} \forall i, j, a_i \in S_A, b_j \in S_B \{touch\ (a_i, b_j) = ture\} \\ \sum_i^{Ca} \sum_j^{Cb} intersecti\ on\,(a_i, b_j) * \sum_i^{Ca} \sum_j^{Cb} contain\ (a_i, b_j) * \sum_i^{Ca} \sum_j^{Cb} within\ (a_i, b_j) = 0 \end{cases} \tag{3}$$

Within:

Any raster scan line a_i in S_A can be found out in the other raster scan line which includes a_i in S_B. It can be described as:

$$\forall i, a_i \in S_A, \exists b_j \in S_B (within\ (a_i, b_j) = ture) \tag{4}$$

Contains:

Any raster scan line b_j in S_B can be found out in the other raster scan line which includes b_j in S_A. It can be described as:

$$\forall j, b_j \in S_B, \exists a_i \in S_A(contain\ (a_i, b_j) = ture) \tag{5}$$

Overlaps:

Decide A and B as objects with same dimension. Within two sets of raster scan line S_A and S_B, there exists more than one intersection or equal relation and more than one disjoint relation.

$$\begin{cases} dim(A) = dim(B) \\ \sum_i^{Ca}\sum_j^{Cb} intersection(a_i, b_j) > 0 \quad or \quad \sum_i^{Ca}\sum_j^{Cb} equal\ (a_i, b_j) > 0 \\ \sum_i^{Ca}\sum_j^{Cb} disjoint\ (a_i, b_j) > 0 \end{cases} \tag{6}$$

Crosses:

Decide A as objects of line and plane, B as object of line. Within two sets of raster scan line S_A and S_B, there exists more than one disjoint relation and intersection or equal relation.

$$\begin{cases} dim(A) > 0, dim(B) = 1 \\ \sum_i^{Ca}\sum_j^{Cb} intersection(a_i, b_j) > 0 \quad or \quad \sum_i^{Ca}\sum_j^{Cb} equal\ (a_i, b_j) > 0 \\ \sum_i^{Ca}\sum_j^{Cb} disjoint\ (a_i, b_j) > 0 \end{cases} \tag{7}$$

Intersects:

Within two sets of raster scan line S_A and S_B, there exists more than one intersection relation.

$$\sum_i^{Ca}\sum_j^{Cb} intersection(a_i, b_j) > 0 \tag{8}$$

C. Line-by-line Calculation Flow of Topological Relations

After scanning and conversion, the topological relations between any two geometric objects can be calculated by set of raster scan line. And the calculation of raster scan line sets can be resolved into topological calculation between any elements in these two sets. The seven topological relations of raster scan line as shown in Table I, except for disjoint relationship, occur in the same line (touch relationship can occur in two adjacent lines). While for disjoint relation, if relation of each raster scan line for objects A, B in each line is disjoint, then A, B belong to disjoint relation. Therefore, the topological relations of two raster scan lines can process by adopting line-by-line calculation, and the processes are:

1) Obtain the scan line range of objects A and B, minimum and maximum of attribute *j*. The topological relations of raster scan lines A and B must occur in the interval.

2) In the range of minimum and maximum, successively obtain raster scan lines a_i and b_j of objects A and B line by line.

3) Calculate topological relations of a_i and b_j, and judge whether they satisfy topological conditions or not. If yes, then shift to step 2 to calculate topological relations of next raster scan line.

4) If the traversal is to maximum, synthesize former calculation results, if satisfy topological judging conditions, return calculation result true.

Complexity of model about line by line calculation increases to $O(max(C_a, C_b))$ and is finished only by Boolean operation.

III. OPTIMIZATION OF TOPOLOGY CALCULATION BASED ON RASTER SCAN LINE

When parallel raster scan lines with fixed spacing are used to represent geometric objects, the interval size of raster scan line (size of grid cell) not only decides memory usage, but directly influences correctness and accuracy of calculation results of topological relationship. Theoretically speaking, the method of improving calculation precision by increasing spacing of raster scan line only infinitely approaches but can't reach coordinate precision of original geometric data. Furthermore, the method of decreasing grid size certainly will bring about floating-point arithmetic of more spatial coordinates.

Current spatial coordinate system mainly includes geographic coordinate system which uses degree as unit, and projected coordinate system by using kilometer and meter,etc as units. Geographic coordinate system and projected coordinate system can be represented by floating-point number. Comparing with real number field, geographic coordinate system is limited in value range of longitude which is [-180,180]. If multiply floating-point number of longitude and latitude by one certain coefficient and represent by using 32-bit integer, then its precision is: $L/2^{32-1} \approx 2$cm (L represents great-circle distance). The precision can satisfy most of GIS application requirements. So using integer value to represent spatial coordinate value is one reasonable and efficient method of describing spatial objects. The lowest interval to promise integer arithmetic of raster scan line is $I_{min}=w/2^{32-1}$ (w is bounding rectangle range of element).

Spatial data has characteristic of measuring scale. Under the circumstance, because precision of human vision addressability is 0.01 cm, scanning and acquisition accuracy of spatial data are defined with the reference of human vision addressability. So *scale* is used to represent measuring scale of spatial data, then maximal scanning interval is $I_{max}=scale*0.01$. Take the example of 1:50000 scale data, the maximal scanning interval is 5 meters. Therefore, the value range of interval of raster scan line (size of cell unit) is $[I_{min}, I_{max}]$.

In order to take account of computational efficiency and computational precision, this paper put forward computing method of double raster scan lines nestification. As shown in Fig. 2, when calculating topological relations of two line objects A and B, firstly calculate topological relations under raster scan line with precision is m, if satisfy logical condition of topological calculation and does not exist underlying ambiguity of topological relations (For instance,if sampling width is large, maybe it is equal relation; if sampling width decreases, then maybe it belongs to intersection relation.), then operate directly. If exist, then scan the raster the second time with precision for the local of A and B, and make a further calculation of topological relations with this precision. The scanning precision of double raster m and n satisfies threshold range described above: $I_{min} \leq m \leq I_{max}$, $I_{min} \leq n \leq I_{max}$.

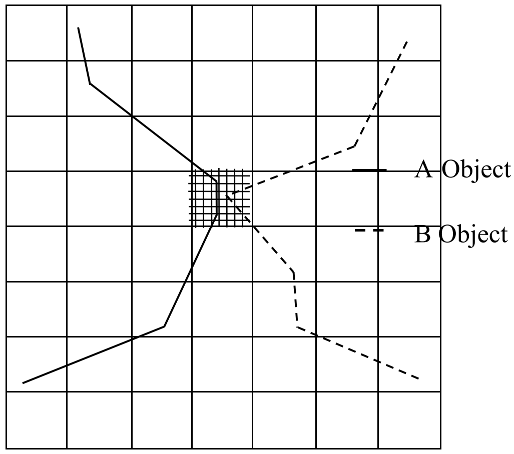


Fig. 2 Double Raster Scan Lines Nestification

IV. EXPERIMENTAL VERIFICATION

In the experiment, correctness, computational efficiency and memory usage, etc, of algorithm in this paper have been experimental verified by using 1:10 00000 topographic map data in a certain range of China. Calculate contains relation by choosing one certain administrative domain and all punctiform settlement places; calculate overlaps relation by choosing one certain basin and all county administrative divisions; calculate Intersects relation by choosing one certain basin data and all county administrative divisions; calculate Cross relation by choosing one certain basin and all linear road data. The above topological relations are calculated by using nested grid line with two precision [1/3600, 1/360000] in the experiment. The algorithm in this paper develops by c++ language. Contrast test is made by using computing software package of open source spatial relation geos 3.1.1. Test 10 times in each time, record mean value of computing time of topological relations and data volume in calculation of topological relations (record 1k if it is less than 1k), and return number of objects. The test environment is: IBM Thinkpad T410s, Inter（R）core（TM）i5 CPU, 3GB memory and Windows 7 operation system. The test results are shown in Table II.

DOI 10.5013/IJSSST.a.17.27.19

ISSN: 1473-804x online, 1473-8031 print

TABLE II. COMPUTING EFFICIENCY OF TOPOLOGICAL RELATIONS FOR COMPLEX OBJECTS

Topological relation	Computing Method	Computing data volume （KB）	Number of return objects	Execution time (ms)
Contains	our method	1/31181	14461	2321.7
	geos	1/31181	14461	5187.7
Overlaps	our method	1/21495	141	218.9
	geos	1/21495	141	788.9
Intersects	our method	1/31181	15039	3117.2
	geos	1/31181	15039	5273.0
Crosses	our method	1/14431	358	783.2
	geos	1/14431	358	2103.9

What can be known in the table II is that for the examples about calculation of four topological relations in the experiment, the method in this paper can get computing results of topological relation in accordance with geos. Calculation of topological relation in geos includes lots of floating-point operations of geographical coordinates. The method in this paper touches upon floating-point operation only when geographical coordinates are converted to integer coordinates. The line-by-line calculation model can further improve computing efficiency. The computing efficiency in the paper is apparently more optimal than geos.

In order to verify memory usage of raster scan line and efficiency of nested raster scan line in this method, objects from one plane and surface-lake data are chosen to calculate contains relation. In the geographical coordinates, nested raster scan lines with four precision (1/360,1/360000), (1/1800,1/360000), (1/3600,1/360000) and (1/7200,1/360000) are used respectively to calculate and record memory usage, structure time and number of raster scan line at first scan, and memory usage, structure time at second scan. The experimental results can be shown in Table III.

TABLE III. CONTRAST BETWEEN SOURCE USAGE AND EFFICIENCY CALCULATED BY CONTAINS RELATION FOR COMPLEX

Scanning precision (degree)	First Memory usage(KB)	First scan time (ms)	First Number of raster scan lines	Second Memory usage (KB)	Second scan time (ms)
1/360,1/360000	38.41/0.15	140	3427/19	1532.74/5.72	4070
1/1800,1/360000	191.98/0.74	490	17132/92	1178.74/2.85	2290
1/3600,1/360000	383.98/1.47	1080	34263/184	893.69/1.32	1590
1/7200,1/360000	768.98/2.92	2470	68525/376	369.16/0.21	590

Table III shows that in the experiment of calculating topological relations of polygon-polygon and line-line, when the precision is low at first scan, there is few memory usage of raster scan line, but the memory usage and structure time are larger at second scan. With the improvement of precision at first scan, the memory usage also gradually improves. Overall, memory usage is far below the memory configuration of current desktop computer or server which can finish in the computer with normal configuration.

V. CONCLUSION

This paper gives definition of topological relations for objects point, line and polygon and line-by-line calculation model on the basis of definition of seven one-dimensional topological relations for raster scan unit. Compared with analytic geometry, the method in the paper has unified description and computing framework of topological relation; compared with the calculation method of topological relation based on planar graph, the method in the paper can avoid computational error from floating-point arithmetic itself and has high operating efficiency. Compared with other scanning conversion methods, the method in the paper can be fit for objects point, line and polygon at the same time. Subsequent works include: (1) optimal selection and calculation of precision about double raster scan lines; (2) parallel computing of topological relation based on raster scan line.CONFLICT OF INTERESTThe authors confirm that this article content has no conflicts of interest.

ACKNOWLEDGMENT

This work is supported by the National Natural Science Foundation of China(grant no.41271446).

REFERENCES

[1] W. Kainz, M. Egenhofer, Greasley I. "Modelling spatial relations and operations with partially ordered sets," International Journal of Geographical Information Science, vol. 7, pp. 215-229, 1993 .

[2] A. Cohn, B. Bennett, J. Gooday, et al. "RCC: a calculus for region based qualitative spatial reasoning," GeoInformatica, vol. 1, pp. 275-316, 1997.

the calculation method of topological relation based on

[3] S. Chang, Q. Shi, C. Yan. "Iconic indexing by 2-D strings," IEEE Trans. Pattern Anal. Mach. Intellig., vol. 9, pp. 413-428, 1987.

[4] Jun Chen, Li Chengming, Li Zhilin, C. GOLD. "A Voronoi-based 9-Intersection Model for Spatial Relation," International Journal of Geographical Information Science, vol. 15, pp. 201-220, 2001.

[5] R. Praing, M. Schneider. "Efficient implementation techniques for topological predicates on complex spatial objects," GeoInformatica, vol. 12, pp. 313-356, 2008.

[6] Wang Wencheng, Wu Enhua. "A New Method Deciding Whether a Point is in a Polygon or a Poly hedron," Journal of Software, 2000, 11(12):1612-1617.

[7] Xia Delin, Xiong Xin. "New Algorithm and Principle for Intersection of Line Segments," Journal of Computer-Aided Design & Computer Graphics, vol. 8, pp. 95-100, 1996.

[8] Zhu Yayin, Wang Feng, Wang Huawen. "Algorithm for Determing the Spatial Relationship between Two Polygons," Computer Engineering and Applications, vol. 1, pp. 91-93, 2003.

[9] Vivid Solutions Inc. JTS Topology Suite, Vivid Solutions Inc. [EB/OL]. http: / /www. vividsolutions.com /jts/jtshome.htm, 2007.

[10] J. SACK, J. URRUTIA. Computational Geometry. Netherlands: Elsevier Science, 1999.

[11] Jai Menon, J. Marisa Richard, Jovan Zagajac. "More powerful solid modeling through ray representations," IEEE Computer Graphics and Applications. vol. 14, pp. 22-35, 1994.

[12] B. Zalik, I. Kolinger ova, "A cell based point in polygon algorithm suitable for large sets of points," Computer&Geosciences, vol. 27, pp. 1135-1145, 2001.

[13] Yang Congjun, Ren Yingchao, Li JinPing. "Red-Blue Sweep Line Algorithm Based on Monotone Chains for Connected Line Segment Intersection Problems," Geomatics and Information Science of Wuhan University, vol. 31, pp. 835-838, 2006.

[14] Wang Jiechen, Shen Dingtao, Li Lifan. "Area Topological Computing Based on Sweep Line and Trapezium-split Operation," Geomatics and Information Science of Wuhan University, vol. 35, pp. 773-777, 2010.

[15] G. Open. Consortium. "OpenGIS Simple Features Specification for SQL,"(revision 1.1). 1999.