GEO-INFO GRAPH SPECTRUM ANALYSIS FOR REPRESENTING DISTANCE RELATIONS IN GIS

M. Y. Duan [a, *], Y. Zang [a], P. X. Pu [a], L. Shi [b]

[a] Chinese Academy of Surveying and Mapping, No. 28, Lianhuachixi Road, Beijing 100830, China - duanmy@casm.ac.cn
[b] The 2nd Institute of Surveying and Mapping of Hebei Province, Shijiazhuang 050031, China

WG VI/4

KEY WORDS: Spatial Distance; Geo-info Graph Spectrum, Spectrum Analysis

ABSTRACT:

Distance is a fundamental concept in spatial sciences. Spatial distance is a very important parameter to measure the relative positions between spatial objects and to indicate the degree of similarity between neighbouring objects. Indeed, spatial distance plays an important role in spatial query, analysis and reasoning. However, how to represent directional relations in a unified form is still an open issue. Indeed, the information of object direction is complicated and relative. It is difficult to describe the directional relations using mathematical method. In this paper, existing representing algorithms for the distance between spatial objects are evaluated and their problems pointed out; then, the concept of Geo-info Graph Spectrum is introduced as a metric indicator for different types of spatial objects. This paper presents a rigorous mathematical methodology that addresses the idea of using directional spectrum for direction analysis between spatial objects. So, the complicatedness can be represented by their spectrums or feature values, and the relatively can be represented by the hierarchies of their spectrums (or feature values). Furthermore, an experiment is given to illustrate the feasibility and advantages of the proposed approach in the paper, explaining how to use Direction Spectrum to do spectrum generation and spectrum analysis as to get the results of directional relationship between spatial objects. Finally, potential applications are discussed.

1. INTRODUCTION

Currently, description and inference of spatial relations are considered as the common interest in the study areas of geography, computer science, and cognitive science, and so on. Spatial relations usually include distance relations, direction relations and topology relations, essentially expressed the characteristics of restraint among the data with different layers separately.

In daily life, distance is a measure of the effort required to reach one place from another. It can be specified in various ways, e.g. travel time, length, or cost (Sharma 1996). Distance is a fundamental concept in spatial sciences. Spatial distance is a very important parameter to measure the relative positions between spatial objects and to indicate the degree of similarity between neighbouring objects. Indeed, spatial distance plays an important role in many areas such as neighbourhood analysis (Chen et al. 2004), structural similarity measure (Veltkamp 2001), image (or object) matching (Rucklidge 1996, Devogele 2002), clustering analysis (Jain et al. 1999), and so on. In GIS, distance is usually utilized as a constraint for spatial query and analysis.

Spatial distance may be defined in different ways. In Euclidean space, distance means the straight length between two given points, which is in fact the shortest distance. However, in a spherical space, the distance along the great circle becomes the shortest distance, which is also called the geodetic distance. In the field of geographic information science (GIS), all such distances are defined in a so-called vector space. On the other hand, in a raster space, the definition of a distance is an approximation of vector distance. The commonly used raster distances are chessboard, city block, octagon, chamfer 2-3, and chamfer 3-4 distances (Rosenfeld and Pfaltz 1968, Borgefors 1986, 1994, Melter 1987, Breu et al. 1995, Embrechts and Roose 1996).

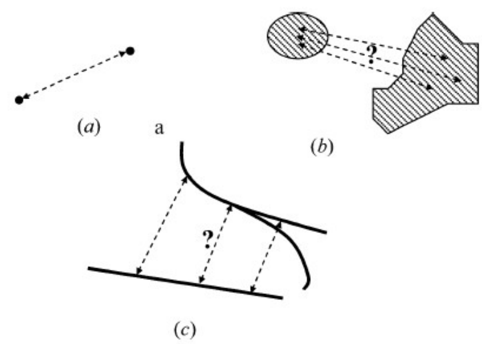


Figure 1 Examples of distances between various objects

As one can imagine, in spatial data handling, there might be point, line, and area objects. Therefore, there is a need to provide a general distance concept so as to accommodate all these kinds of spatial objects (figure 1). From the literature, it can be found that such terms as minimum, maximum, and centroid distances have been in use. These distances do have their applications domains. However, in other applications, these distances may fail to make sense because they have not taken into consideration the position, orientation, shapes and extent of objects. That is, they are incapable of measuring the distance relations of the objects adequately.

* Corresponding author. duanmy@casm.ac.cn

2. LITERATURE REVIEW

2.1 A Critical Examination of Existing Distance Measures

In geoinformation science, the most commonly used distance is defined by Euclidean geometry and Cartesian coordinate. In a two-dimensional Cartesian system, the Euclidean distance between two points is as follows:

$$D(p_i, p_j) = \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2} \qquad (1)$$

Where (x_{i1}, x_{i2}) and (x_{j1}, x_{j2}) are the Cartesian coordinates of points p_i and p_j, respectively.

It is well known that the Euclidean distance satisfies the following four properties:

- d1. Non-negativity: $d(p_1, p_2) \geq 0$, for any two points p_1 and p_2;
- d2. Identity: $d(p_1, p_2)=0$ iff $p_1= p_2$;
- d3. Symmetry: $d(p_1, p_2)= d(p_2, p_1)$; and
- d4. Triangle inequality: $d(p_1, p_2)= d(p_1, p_3)+ d(p_3, p_2)$

Therefore, the Euclidean distance is a metric. It should also be pointed out that equation (1) is the distance between two individual points. However, in GIS, there are also line and area objects. In order to make the measure of distance between all types of spatial objects possible, some extensions of this model need to be made. From the literature, it can be found that the minimum (or shortest) distance (Peuquet 1992), the maximum distance, and the distance between the centroids of spatial objects have already been in use. These distances are respectively defined as follows:

- Minimum distance:

$$D_{min}(A, B) = \min_{p_a \in A}\{\min_{p_b \in B}\{d(p_a, p_b)\}\} \qquad (2)$$

- Maximum distance:

$$D_{max}(A, B) = \max_{p_a \in A}\{\max_{p_b \in B}\{d(p_a, p_b)\}\} \qquad (3)$$

- Centroid distance:

$$D_c(A, B) = d(\tfrac{1}{m}\textstyle\sum_{i=1}^{m} v_{iA}, \tfrac{1}{n}\sum_{j=1}^{n} v_{jB}) \qquad (4)$$

In equation (4), v_{iA} $(i = 1, 2, \dots, m)$ is the ith vertex of object A; v_{jB} $(j = 1, 2, \dots, n)$ is the jth vertex of object B. Figure 3 is an illustration for these three distances. One may notice that the differences among them could be very large, depending on the shape of the objects.

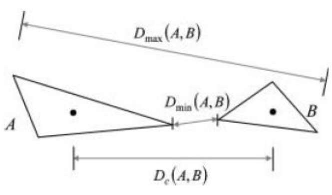


Figure 3 Minimum, maximum and centroid distance

Let us take the minimum distance as example to see the suitability of these distances. In figure 4, the minimum distance between A and B_1 is equal to that of A and B_2, according to equation (2), although B_1 is distinctly different from B_2 in shape and size.

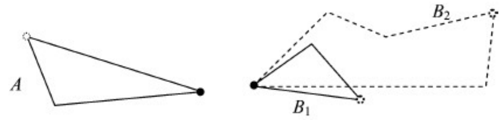


Figure 4 Shape of the whole object not considered in the minimum distance

Of course, the type of distance has its particular applications. For example, in forest precaution, any kindling (a point or an area object) must be further away from the forest (an area object) for a given distance, which indeed employs a minimum distance criterion. However, such a distance may cause a contradiction to human cognition. For instance, in figure 4, one could say that object A is close to object B_1 (or B_2) considering their minimum distance, as shown by two solid circles. With this kind of information in mind, one would naturally expect that on point on one object is far away from the other object. In practice, this is not the case, and one could see that the distance between the points circled with a broken line indicates that they are not so close.

This is because the minimum distance in equation (2) takes into consideration only a single point from each object but has nothing to do with the position, shape, orientation, and spatial extent of the spatial objects at all.

2.2 Hausdorff Distance as Measure of Distance Between Two Spatial Objects

Given two point sets A and B, the Hausdorff distance between A and B is defined as (Rucklidge 1996):

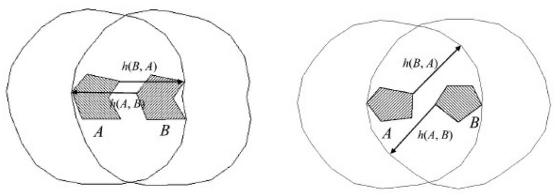
$$H(A, B) = \max\{h(A, B), h(B, A)\} \qquad (5)$$

Where:

$$h(A, B) = \sup_{p_a \in A}\{\inf_{p_b \in B} \|p_a - p_b\|\} \qquad (6)$$

$$h(B, A) = \sup_{p_b \in B}\{\inf_{p_a \in A} \|p_a - p_b\|\} \qquad (7)$$

and sup $\{\cdot\}$ represents the least upper bound of a set; inf $\{\cdot\}$ represents the greatest lower bound of a set; and $\|\cdot\|$ some underlying metric defined on the points of A and B.



(a) (b)

Figure 5 Illustration of the Hausdorff distance between A and B. (a) B obtained from A with a triangulation. (b) Two general area objects A and B.

2.3 Extended Hausdorff Distance between Spatial Objects

By now, it is clear that h(A, B) represents the largest distance from any boundary point on A to its nearest point on B, and h(B, A) represents the largest distance from any boundary point on B to its nearest point on A. As a result, the Hausdorff distance is sensitive to the shape of the two objects. It is noticeable that even a single 'outlying' point may greatly affect the value of Hausdorff distance. That is to say, the Hausdorff distance is not robust with respect to some outlying portions. Figure 6 shows such a case. In this figure, object B in (b) has a long tail, which leads to a large difference in the Hausdorff distance between A and B. Therefore, a more robust distance is desirable.

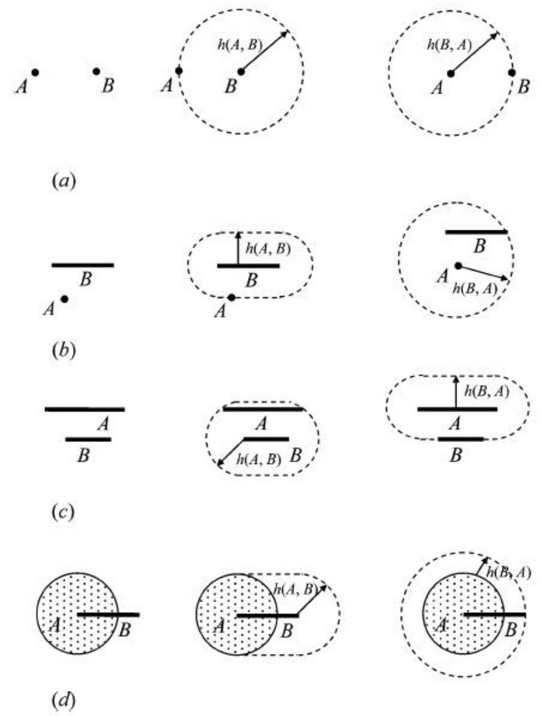


Figure 6 Hausdorff distance between spatial objects

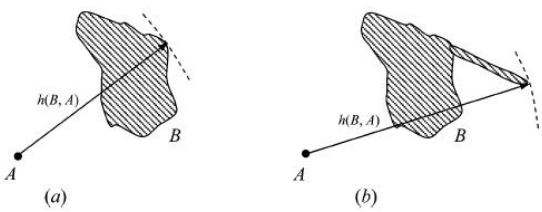


Figure 7 Effect of locally outlying portions on the Hausdorff distance. (a) h(B, A)=4.1cm. (b) h(B, A)=5.2cm.

It can be seen from the above figure, Hausdorff distance algorithm is sensitive to the shape, size and orientation of the objects, results can be greatly affected by even the single point. As a result, it is necessary to develop a common rigorous mathematical methodology for direction analysis between spatial objects.

3. METHODOLOGY

3.1 Geo-info Graph Spectrum

The Geo-information graph spectrum (Carto-methodology in Geo-information, CMGI) is brought up by academician Chen Shupeng, which is a kind of methodology in GIS field, supported by such advanced technologies as Remote Sensing (RS), Geographical Information System (GIS), Virtual Reality, Cartography and Internet Communication by computer, etc. Geo-Info graph spectrum is the space-time compound body of geo-information. The fundamental function element of geo-information graph spectrum is the element of geo-information. The geographical feature element is as same as spectrum element is also multi-rank and multi-dimension can make classification according to the space-time dimension diversity, the diversity studying purpose.

3.2 Procedure of Distance Spectrum Analysis

The overall framework of distance spectrum analysis methodology is presented in figure 8.

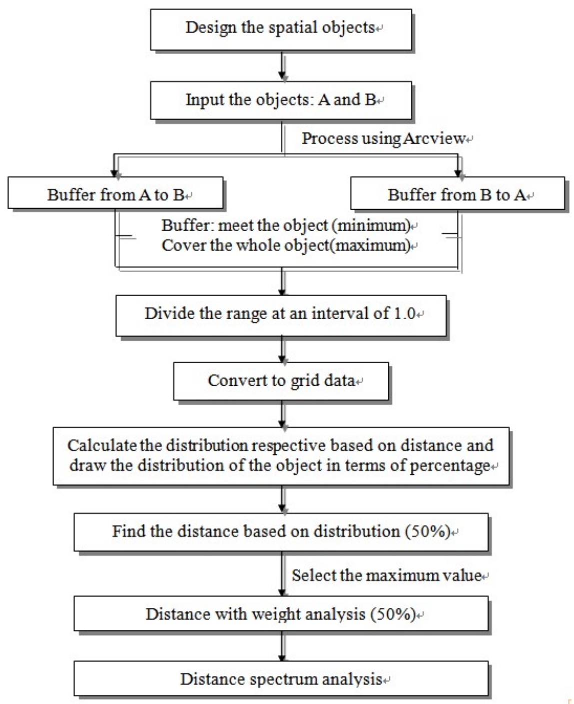


Figure 8 Flowchart of distance spectrum analysis procedure

3.3 Spectrum Analysis

Here we need mathematical factors to define the index of distance spectrum. We use distance spectrum (min., max., μ, H(x), δ) (five parameters) to define this equation. Explanations about these five parameters are as follows:

- Min. and Max. Values stand for the minimum and maximum values of distance.
- μ is the average distance value (Min.< μ <Max.).

$$\mu = \frac{\varepsilon_{min}+\varepsilon_{max}}{2} \qquad (8)$$

- Information entropy of a discrete random variable X, we define:

$$H(x) = -\sum_{i=1}^{n} P(x_i)\log_{10} P(x_i) \qquad (9)$$

- The standard deviation is the most common measure of statistical dispersion, measuring how widely spread the values in a data set

$$\delta = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(p_i - \overline{p})^2} \qquad (10)$$

4. EXPERIMENTS

4.1 Data Preparation

Assume that object A (polygon) represents the island, object B (polygon group) represents two buildings.

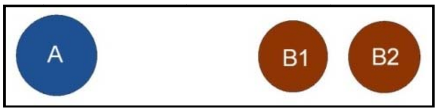


Figure 10 Data Preparation

4.2 Buffer from Object A to Object B (B1, B2)

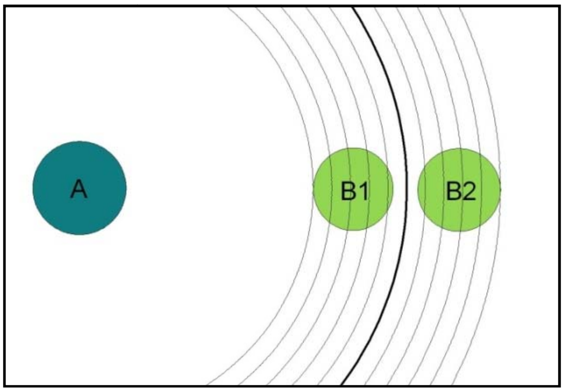


Figure 11 Buffers from Object A to Object B (B1, B2)

From figure 11 and table 2, we can see that object B (B1, B2) is divided into 10 parts by the group of rays from object A. 17.570m is the minimum distance that the groups of rays just meet object B (B1, B2). 28.540m is the maximum distance that the groups of rays finish covering the object B (B1, B2). The thick line in figure 11 shows that the distance covers 50% length or area. Assume the whole area is 100%, object B (B1, B2) is divided into equal parts by dash lines, and the distance covers 50% length or area.

Attributes of Test02-03.shp

Shape	Id	Buffendis	Code	Area
Polygon	0	550.0000	1	6077.65
Polygon	0	600.0000	2	10158.43
Polygon	0	650.0000	3	10907.92
Polygon	0	700.0000	4	8641.32
Polygon	0	750.0000	5	1064.48
Polygon	0	800.0000	6	1694.22
Polygon	0	850.0000	7	8477.79
Polygon	0	900.0000	8	10772.60
Polygon	0	950.0000	9	10537.65
Polygon	0	1000.0000	10	6887.01

Table 1 Attribute table in ArcView GIS

Code	∑ Area of each detected part	% of detected area	Distance distribution
1	6077.65	8.08	17.570
2	10158.43	13.51	18.666
3	10907.92	14.50	20.858
4	8641.32	11.49	21.954
5	1064.48	1.42	23.060
6	1694.22	2.24	24.156
7	8477.79	11.27	25.252
8	10772.60	14.32	26.348
9	10537.65	14.01	27.444
10	6887.01	9.16	28.540
Total	9993.01	100.00	

Table 2 Distribution table of Area distribution

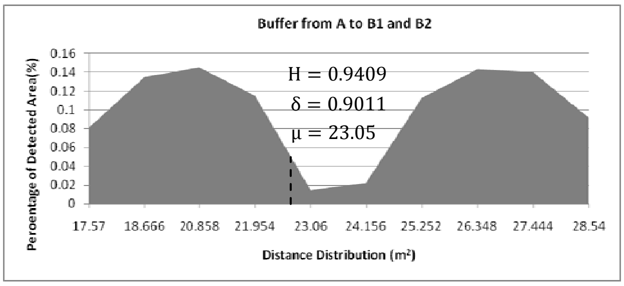


Figure 12 Description graph of buffering from object A to Object B (B1, B2)

4.3 Buffer from Object B (B1, B2) to Object A

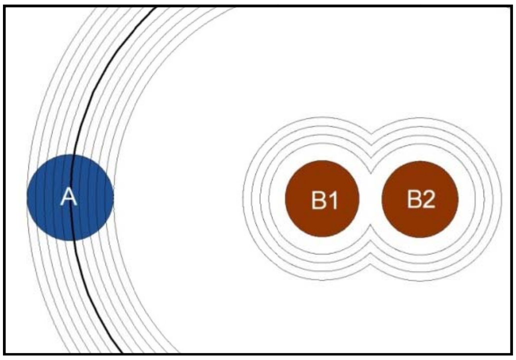


Figure 13 Buffers from Object B (B1, B2) to Object A

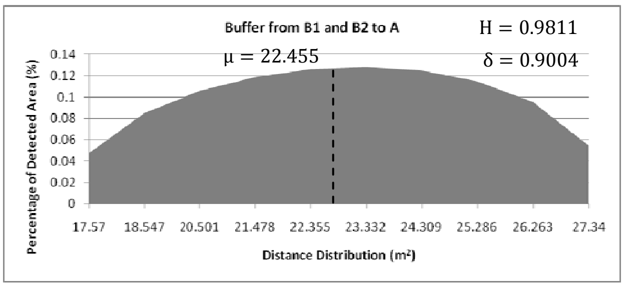


Figure 14 Description graph of buffering from Object B (B1, B2) to object A

4.4 Distance between Object A and Object B (B1, B2)

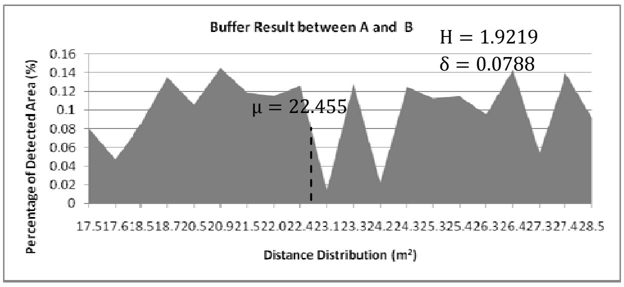


Figure 15 Description graph of buffering between object A and Object B (B1, B2)

5. CONCLUSIONS

Based on extended Hausdorff distance theory and weight analysis, a new methodology of analysis of distance relations has been put forward.

From the above experiments, we can get the conclusion that it is more helpful to be used in the analysis of GIS for integrated generalization. Distance relations can be defined by charts; we can easy find our mathematics result such as minimum, maximum, average and entropy values, and standardized normalization graphs are convenient for mathematical analysis. However, this methodology is complex of processing and time-consuming, and shapes, size of spatial objects have to be taken into account while processing.

Because of the limitation of time and the interval value, the description graph of buffering results should be smoother.

REFERENCES

D. Min, L. Zhilin and C. Xiaoyong, Extended Hausdorff distance for spatial objects in GIS, *International Journal of Geographical Information Science*, Vol. 21, No. 4, 459-475, 2007.

Alexandroff P., 1961. *Elementary Concepts of Topology*. Dover Publications, New York.

Breu, H., Gil, J., Kirkpatrik, D. and Werman, M., 1995, Linear time Euclidean distance transformation methods. *IEEE Transations on Pattern Analysis and Machine Intelligence*, 17, pp. 529-533.

Borgefors, G., 1986, Distance transformations in digital images. *Computer Vision, Graphics and Image Processing*, 34, pp. 344-371.

Burrough P. A., 1986. *Principles of Geographical Information Systems for Land Resources Assessment***.**Oxford University Press, Oxford.

Chen Shupeng, 2001. Prospective considerations for cartography science. *Journal of Science of Surveying and Mapping, 26(1)***.**

Chen Yuanyuan, 2006. *A study of multi-dimensional and dynamic geo-informatic tupu***.** (Master dissertation No P283, Information Engineering University, 2006). Zhengzhou: Information Engineering University.

Chen Yan, Qiqing Wen, Yang Guishan, et al. 2006. Basic theories of geo-info-tupu. *Journal of Scientia Geographica Snica, 26(3).*

Chen, X., et. al., 1995. *Spatial Relations of Distance between Arbitrary Objects in 2D/3D Geographic Spaces based on the Hausdorff Metric***.** LIESMARS'95, Wuhan, China. 30-40 p.

Chen, X., 1996. *Spatial Relations between Sets***.** ISPRS XVIII Congress. Vienna. Austria.

Chen, X., 1995. *Spatial relations of distance between arbitrary objects in 2D/3D geographic space based on the Hausdorff metric.* In LIESMARS'95, Wuhan, China, pp.30-40.

Chen, J Zhao, R.L. and Li, Z. LZ., 2004, *Voronoi-based K-order neighbour relations for spatial analysis. ISPRS Journal of Photogrammertry and Remote Sensing*, 59, pp. 60-72.

ACKNOWLEDGEMENTS

The author would like to acknowledge the Chinese Academy of Surveying and Mapping and Asian Institute of Technology for providing a mini-project that enabled us to undertake this study.