An algorithm of regional space segmentation on integration of spatial and non-spatial attributes based on weighted-Voronoi Polygons

YAN ZHANG   XIAODONG ZHANG   DEHAI ZHU   ZHIMING HUANG
College of Information and Electrical Engineering
China Agricultural University
Box 698, China Agricultural University, Beijing, China, 100083
CHINA

Abstract: Voronoi Polygons as a simple and effective method has been applied in many fields. This paper first introduces the original Voronoi Polygons generation algorithm, and then proposed a new algorithm which is based on integration of spatial and non-spatial attributes, for regional space segmentation.

The mainly method of algorithm utilizes grid data to displace the traditional TIN to complete Voronoi Polygons. Traditional Voronoi generation process actually just embodies the continental distance of space entities due to neglecting space entity attributes. Therefore, this paper will first normalize non-spatial attributes, and process the value of non-spatial attributes with weight. Then put the attribute characteristic value as weight value, calculate the space distance with weight, namely, consider the influence of non-spatial attributes. At last generate the Voronoi Polygons and combine the non-spatial attributes of the spatial object more efficient.

Finally, in order to analyze the influence of feed enterprises, we try to use this algorithm on the research and application of space segmentation using the feed corporation data of China Anhui as an example.

Key-words: Weighted-Voronoi Polygon spatial attributes; non-spatial attributes; raster date format; feed factor; space segmentation

1   Introduction

Voronoi Polygons was put forward by a Russian mathematician by 1908.And in 1911, Thiessen who was a meteorologist in the Netherlands, tried to use Voronoi Polygons in meteorology. So Voronoi Polygons was called Thiessen Polygon. Nowadays, Voronoi Polygons has been widely used in Space subdivision, Neighborhood search space, Pattern Recognition and so on [1]. It has the geometric characteristics, which play an important role in spatial division.

In actually use, Voronoi Polygons plays an important role in many fields, such as estimating the quality of atmosphere. Data of the contamination of atmosphere usually get from sample points (Observation station), that means data can not cover the situation of whole area. Voronoi Polygons is commonly used to analyze the contamination of atmosphere in the whole area, because Voronoi Polygons can substitute the point for the surface [2]. In the field of geosciences, Voronoi Polygons has been used in searching the adjacent area, geospatial imagery mode. In addition, Voronoi Polygons has been used in mensurating similitude region in spatial database, building a city populated region, establishing schools, shopping centers, and urban public transport stations service region. And Voronoi Polygons which reveals results more intuitively is also used in layout of the major parks in the apartments, confirming the limits of territorial waters and so on. Further more, Voronoi Polygons as a

spatial interpolation method has been used in many other fields. In short, Voronoi Polygons, as a simple method of calculation, has been applied in many fields.

2   Voronoi Polygons definitions and characteristics of the current algorithm

The Voronoi Polygons can be regarded as each point in control points collection is taken as growing core and expanding outwards with the same speed until meet each other. Besides the most outer point which displays an open area, the other points present a convex polygon individually.

At present, there are various ways to model Voronoi Polygons, However, the effectiveness and stability of these algorithms are great different. In a lot of software, Voronoi Polygons are modeled on the basis of Delaunay triangulation. The concrete step is to build triangles by connecting all of the adjacent control points, and then make a vertical bisector for every line of the triangles. Many vertical bisectors around a control point build a polygon. The diagram which is built by these polygons is Voronoi Polygons [3]. It is only consider the spatial attributes of every point to build Voronoi diagram using Delaunay triangulation. That means we consider that the value of unknown region is the value of the nearest control point. However, in practical applications, such as classifying the regional of city in terms of influence, the result is often unsatisfactory, if we build Voronoi Polygons using Delaunay triangulation. It's because that the influence of cities associate with the non-spatial attributes such as the size of the cities, population, transportation, topography and the others. We can't simply classify it in terms of the distance between the city [4].Figure 1, shows the process that Voronoi Polygons generated by using Delaunay triangulation.
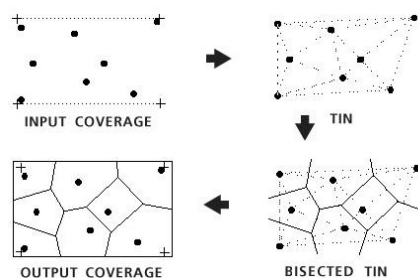

Fig1 the process that Voronoi Polygons generated by using

Delaunay triangulation

In view of this shortage, many scholars have proposed lots of improved algorithm. The most effective algorithm is proposed by Japanese scholar Sugihara which can be successfully achieved 10 million occurred points [5]. Aurenhammer and Edelsbrunner proposed an algorithm which can built a weighted-Voronoi Polygons based on vector [6].In addition, the Voronoi Polygons generation algorithm based on grid has gradually matured. However, these algorithms still can not combine spatial and non-spatial attributes appropriately. Therefore, this paper proposed an algorithm based on the Voronoi Polygons of spatial and non-spatial attributes integration.

3   The achievement of the algorithm of integration of spatial and non-spatial attributes

3.1 Algorithm

Because TIN as a vector data format can not reflect regional differences in spatial attributes, we have to abandon the original vector data format. In this article, we chose raster data format to build Voronoi Polygons instead of the traditional method. Then we take attributes as weight, to correct the distance of space, namely, we build Voronoi Polygons under the impact of non-spatial attributes. The algorithm includes two core issues. One is using of raster data to generate Voronoi Polygons, and the other is generation of Weighted-Voronoi Polygons.

### 3.1.1   Using raster date format to build Voronoi Polygons

Traditional vector data format can not show the non-spatial attribute of a spatial object on the map very well, so this algorithm adopts the method of raster date. The algorithm utilizes the characters that the distance from each point in Voronoi Polygons to its control point is smaller than the distance to the other control points. As the characters of Voronoi Polygons is essential to spatial division, each point in polygon can be approximately regarded as having the same feature with its control point.

Setting up a control point collection $P_i(i = 1, 2, , n)$ on a plane. Thus the polygons of random p can be defined as:

$$V_n = \bigcap_{i \ne j} \left\{ p \middle| d(p, p_i) < d(p, p_j), \right\} \quad (1)$$

Here, d is the distance of Euclidean. In the raster spatial, the unit is pixel, coordinates value of two points $p_1(i, j)$、$p_2(m, n)$ are integer and the results are generally calculated using integer approximation instead. So the distance between two cells can be defined as:

$$d(p_1, p_2) = f(i, j, m, n) = \sqrt{(m-i)^2 + (n-j)^2} \quad (2)$$
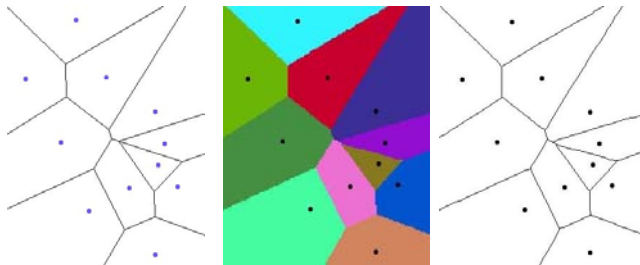
(a) Based on vector data (b) Based on raster data (c) Convert raster data

to vector data

Fig2 Voronoi Polygons based on vector data and raster data formats

Figure 2 (a) shows a Voronoi Polygons based on vector data formats using Arcview3.2, Figure 2 (b) shows a Voronoi Polygons based grid data format for the purposes of the research and development. Figure 2 (c) show an imagine picture which transform Figure 2 (b) into vector data formats using Arcview3.2. Compare the effect of Figure 2 (a) and 2 (c). Voronoi Polygons based grid data format can be a very good approximation of the Voronoi Polygons generation based on vector data formats. That means we can use the Voronoi Polygons based on grid data formats to simulate the Voronoi Polygons based on

vector data formats, which provide a scientific basis for simplifying the Voronoi Polygons generation algorithm.

3.1.2   Weighted-Voronoi Polygon

Spatial objects' attributes include spatial and non-spatial attributes. It has taken into account the spatial attributes using Voronoi Polygons. However, non-spatial attribute is also one of the main features of a spatial object besides the spatial attributes. Non-spatial attribute data is mainly used to describe the other feature except spatial characteristics [7]. Moreover, in practical applications, more than one non-spatial attributes should be considered. For example, sometime we need to describe the size of city. Populations, areas, GDP and other indicators can be selected to estimate the size of cities. By calculating every eigenvalue of attributes of city, we can describe which city is the bigger one or another is the smaller one, which enable classify the regional city in terms of influence

First, normalization is necessary aim at several non-spatial vectors because of different dimension of attributes. In our algorithm, normalized methods are used in preprocessing the raw data. Then eigenvalue of attributes are calculated using the weighted average value [8] [9]. The formula uses a weighted average of all numerical attributes .Set object x be an object of p-dimension space, the eigenvalue of attributes of x is follows:(3)

$$W_{(x)} = (\sum_{i=1}^{p} w_i x_i) / \sum w_i \quad (3)$$

Here, $W_{(x)}$ is the eigenvalue of attributes of x, w_i indicate the weight of eigenvalue of attributes.

Finally, after treating of non-spatial attributes, Weighted-Voronoi Polygons can be generated taking into account the influence of non-spatial attributes using eigenvalue of non-spatial attributes as weight. Namely, to use the weight of non-spatial attributes to correct the original Euclidean distance. In the grid plane, set P_i (i = 1, 2, , n) are n random points in two-dimensional Euclidean space, thus the Weighted-Voronoi Polygons of random point p can be defined as:

$$V_n = \bigcap_{i \ne j} \left\{ p \middle| \frac{d(p, p_i)}{W_i} < \frac{d(p, p_j)}{W_j} \right\}, (i = 1, 2, \cdots, n) \quad (4)$$

Here $d(p, p_i)$,$d(p, p_j)$ show the grid distance between unknown point p to p_i-point and p_j-point. W_i and W_j indicated the non-spatial attributes eigenvalue of point p_i, p_j. When $W_1 = W_2 = \ldots = W_n$, (4) is equal to (1), and the normal Voronoi Polygons is a special case of Weighted-Voronoi Polygons.

This enable spatial and non-spatial attributes of spatial object well-integration. Moreover, we can see that if characteristics value of the non-spatial attributes is bigger, and the Weight distance is smaller. We can believe that at the regional segmentation, if the Voronoi Polygons with weight generated by the control points is larger, the control points will have greater influence to the around.

3.2 The basic algorithm steps

The specific algorithm to achieve the following steps

a) Data Preprocess. Normalizing and weighting the numerical attributes to eliminate the impact of dimension, so that enable non-spatial attributes comparable.

b) Data Format Translation. Convert the control point with non-spatial attribute vector data to grid data format. The grid data has non-spatial attributes eigenvalue of the control point on the spatial location of the control points.

c) Eigenvalue Calculation. Load grid file and calculate the distance with weight between each unknown grid $x (i,j)$ and each control points $p (k,l)$, such as formula (5)

$$d(x_{ij}, p_{kl}) = \sqrt{(i-k)^2 + (j-l)^2} \Big/ W_{(p)} \quad (5)$$

i, j, k, l, is the coordinates of unknown grid $x (i, j)$ and control point $p_i (j, l)$. $W_{(P)}$ indicates the value that the non-spatial attribute of control point p after normalized.

d) Classify Unknown Grid. Compare the minimum of distance with weight between each grid and control point. We can get the nearest control point of grid distance with weight, and set the sub attribute of this grid as this control point. In fact that, to make the grid belongs to the control point.

e) Repeat steps c), and Step d), until the entire grid are traversal.

f) Smoothing Processing. Because of the weight distance that is an approximate value integrated from floating dot data during calculating course, there are some deviations in allocating attribute, which bring some " burr " and" saw tooth" on the border of the Weighted-Voronoi Polygons based on raster data. In order to clear off the "burr", this paper smoothes the border by using eight-neighbor method and get an ideal Weighted-Voronoi Polygons at last.
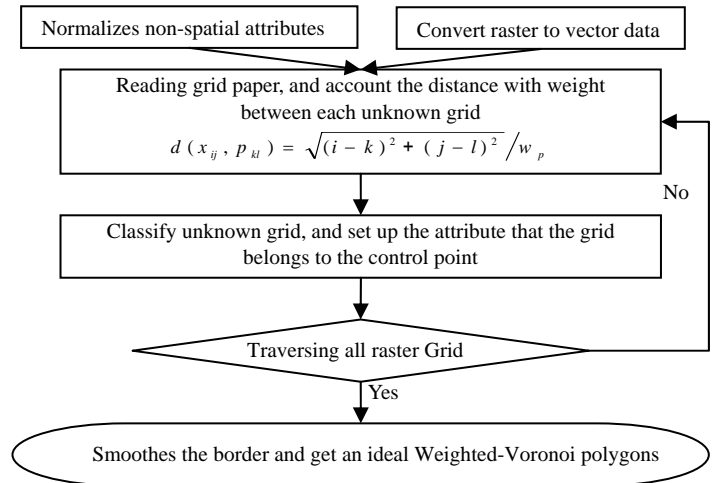
Figure3 is the flowchart of the algorithm

Fie3：Algorithm flowchart

4   Application

Using this method, we take the market area of feed Anhui province in China as an example to segment using the method above. In terms of the fact that the feed enterprises greatly rely on the development of upstream and downstream industry, their developments have already been affected by the scale of feeding and the level of production in the area. Moreover, they are also influenced by many factors, such as the structure and the scale of area planting, the quantity and the scale of the competitor, society, transportation and consumption status of animal food. The raw data of feed industry in Anhui province is normalized to get the eigenvalue of attributes, for calculating the centricity intensity of feed enterprises[10] [11].

At present, many software programs in geographic information systems can achieve the

mutual conversion between vector data and raster date. For reducing the load of programming, we make use of VB programming to carry on Secondary development in virtue of Arcview3.2. First, convert the vector data with coordinate into raster data using "convert to Grid" in Arcview3.2.Then read raster into a two- dimension array which is the original code for raster. Calculate the weighted distance between each raster cell and control point (the center of city) using formula (5) and put the shortest weighted distance as the code of that raster cell. Do this process until the entire raster cell has been confirmed. Put new raster cell into a two- dimension array, then write the data in array to a new file and process it to a weighted-Voronoi Polygons using 'Import Data Source' in Arcview3.2. In the process, we should notice the content of head file which is required in converting data.

Based on the analytic result of feed enterprises centricity intensity in Anhui Province China, we take centricity intensity as weight of feed enterprise in Anhui Province and use the principle and methods of Weighted-Voronoi Polygons to construct the marketing area. (See Fig.4)
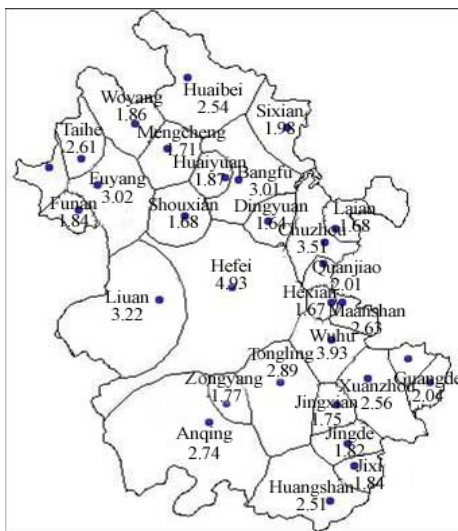

Fig 4 The marketing area of feed enterprises in Anhui province based on the Weighted-Voronoi Polygons

5   Conclusion

This article first point out some problems of the vector Voronoi Polygons which is based on the analysis of GIS vector method, and then propose an algorithm based on weighted-Voronoi Polygons. The algorithm is in considering the impact of spatial and non-spatial attributes, and can be used in region. At last, we take the influence of feed Anhui of China as an example of this division to testify the algorithm.

A.   The result of the algorithm is satisfactory using Weighted-Voronoi Polygons method. The result can reflect the situation of feed factory in Anhui province and can analyze the impact of the distribution efficiently.

B.   The integration of spatial attributes Voronoi Polygons generation algorithm improve the traditional method. Because Voronoi Polygons has manifested itself in the spatial attributes. And we use raster data format for better integrate non-spatial attributes. The algorithm is more effective in integration of spatial attributes, so that the result is more directly reflect the actual situation. We can basically achieve the needs of a wide range of geographic analysis.

However, because of the drawbacks of raster data, the boundary of the polygon is not smooth curves but a lot of broken line. So the algorithm needs to be further optimized.

References

[1]   Wang Xinsheng, Liu Jiyuan, Zhuang Dafang, et al, The application of Voronoi diagram on spatial organizing of urban influence regions, Journal of Central CHINA Normal University (Nat. Sci.), Vol.37, No.2, 2003, pp.256-259.

[2]   Wang Xianhong, Tian Kaocong, Zhou Yanrong et al, The preliminary exploring of application of GIS on the description of distribution of sulfur dioxide in some city zone, Journal of biomathematics, Vol.1, No.18, 2003, pp. 57-60.

[3]   P. Cignoni, C. Montanil and R. D. Scopigno, A fast divide and conquer Delaunay

triangulation algorithm, Computer Aided Design, Vol.5, No.30, 1998, pp. 333-341.

[4] Wang Xinsheng, Li Quan, Guo Qingsheng, et al, The generalization and construction of Voronoi diagram and its application on delimitating city's affected coverage, Journal of Central CHINA Normal University (Nat. Sci.), Vol.36, No.1, 2002, pp. 107-111.

[5] Okabe A, Boots B, Sugihara K, et al, Spatial Tessellations : Conceptsand Applications of Voronoi Diagrams (Second Edition) , New York, John Wiley and Sons , 2000.

[6] Aurenhammer F, Edelsbrunner H, An optimal algorithmfor constructing the weighted Voronoi diagram in the plane, Pattern Recognition, Vol.2, No.17, 1984, pp. 251-257.

[7] Chen Yanfei, Du Pengfei, Zheng Xiaojin, et al, Evaluation on ecological applicability of land construction in Nanning city based on GIS, J T singhua Univ (Sci & Tech), Vol.46, No.6, 2006, pp. 801-804.

[8] G. Karypis, E. H. Han and V. Kumar, CHAMELEON: A hierarchical clustering algorithm using dynamic modeling, Computer, Vol.8, No.32, 1999, pp. 68-75.

[9] X. Y Li, X. Q. Zheng and H. W. Yan, On spatial clustering of combination of coordinate and attribute, Geography and Geo-information Science, Vol.2, No.20,2004, pp. 38-40.

[10] He Xiao, Zhang Xiaodong. Study on the Marketing Area's Building Method based on Voronoi Polygon-An Example of Feed Industry Enterprises of Anhui Province. IGARSS, 2005(International Geoscience And Remote Sensing Symposium)

[11] The office of national feed industry, The association of Chinese feed industry, The Feed Industry Yearbook (2001), China Agriculture Press, 2002.

Corresponding Author：Xiaodong Zhang Correspondence should be addressed to Xiaodong Zhang