A FAST METHOD OF SMOOTHING VECTOR GRAPH CONVERTED FROM RASTER IMAGE

YUE Jian-wei[a,b,] *, HUANG Jun[c], WANG Bin[a]

[a] College of Resources Science and Technology, Beijing Normal University, Beijing. 100875, China;
[b] State Key Laboratory of Earth Surface Processes and Resource Ecology, Beijing100875, China
[c] Land and Resource Bureau of Liuzhou, Liuzhou 545001, China

Commission VI, WG VII/5

KEY WORDS: Smooth，Vector，Method

ABSTRACT:

Vectorization of raster data is very important in GIS. Adjacent vertices of vector graphics usually pass though the same vertical or horizontal straight lines after vectorization. This paper proposes a method based on making full use of this character mentioned before. First, the Latent important vertices will be found the and the zigzag will be eliminated. Second, the redundant vertices, which is very close to the adjacent vertexes, will be removed. After these two processes, we can get smooth vector lines obtained from vectorization rapidly.

1. MANUSCRIPT

Remote sensing data is becoming more and more important in GIS application． Data conversion between raster and vector is one of the key technologies for the integration of RS and GIS．A lot of techniques have been used, such as edge tracing(Huang B 1995), directed boundary(Li Z C, 1997), vectorization technique based on "grid technique" (Zhang X C, 2001), transferring technique based on topological theory(Shen Z Q, 1999), the method for tracing run-length outline and so on．By researching the sideline or the boundary gotten by above technologies, we find that the sideline is serration. But the factual sideline is more unbend. How to process the sideline to get a smoother sideline, which seems as the fact one much more, is an important problem in the transformation of raster to vector accordingly.

Interpolation method is often used by many techniques to smooth curves, but these methods can not be use to deal with the sideline apparently. Douglas-Peucker method was used by many researchers (Chen Renxi, 2006)to process the vector and good result was gotten: the data quantum of the vector graph was decreased and the sideline in the vector graphics was smoother. However, the method has some limitation:

    1)　The method needs too much calculation and too much computer memory. So the computing speed will slower when a curve has volume points. For example, the method needs too much time to get the furthermost point when the curve is composed of many points.

    2)　If the curve is close and the start point is changed, the close curves gotten by the method will show different shapes.
To resolve the problems cited above, this paper studied the vector graphs' character converted from raster images and proposed am arithmetic which can smooth curves fast. The computational time will linearly increase only with the nodes number.
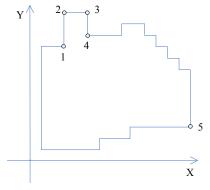

Figure 1. Latent important nodes

2. THE PRINCIPLE OF THE ARITHMETIC

2.1 The basic principle

The edge of the polygon converted from a raster image is serrated, especially the sloping one. The sloping edge brings the polygon some character:

    1)　the angle of the two conjoint line of the polygon is a right angle (Figure 1).

    2)　two neighbour nodes stand on one horizontal line or one vertical line. The distance of the two nodes could be calculated with $|X_n-X_{n-1}|$ or $|Y_n-Y_{n-1}|$. For example, as figure 1, the node number 1and node number 2 are on one vertical line, the distance of them is Y_2-Y1, and the node number 3 and node number 2 are on one horizontal line, the distance of them is X_3-X_2.

    3)　the two end nodes of a vector line of a polygon must be latent nodes, if the direction of the previous line is different

---

*YUE Jianwei, postdoctoral researcher. Main research interests are GIS, E-government, and land and resources informatization. College of Resources Science, Beijing Normal University. E_mail:yuejw@ires.cn.

from the one of the next line.
The node, which is far from not only the previous node but also the next node, must be a latent node, such as the node number 5 in figure 1.

Smoothing a vector graph gotten from a raster image and compressing vector data have the same problem. Compressing vector data requires that draw out the subset A from the data set S. The data in the subset A must be so little as to could not be cut down within the allowable precision and the subset A must be close to the data set S. Basing on above, one smoothing vector graph method is proposed in this paper. The basic principle of the method is as fellows: if the node in a curve of a polygon is very near to the previous node and the next node and the distance is shorter the allowable value, then the node will be deemed to the superfluous one and will be deleted. After this, if the angle formed by the neighbour line is larger than the allowable value, the two line will be look as one line, and the node at the point of the intersection of the two lines will be delete too.

2.2 The process of arithmetic

2.1.1 Eliminate the sawtooth

If the curve is close, a node, selected randomly, will be the first node to compute. The node is marked as Pn. If the curve is unclosed, the second node will be the first one to compute and the first node should be added to the new curve directly.

Computing the distance respective：D n-1,n、D n,n+1、D n+1,n+2 and D n+2,n+3.

D = |Xn+1-Xn| (Yn+1=Yn) or D = |Yn+1-Yn| (Xn+1=Xn)
Computing the direction of the line L n-1, n、L n,n+1、L n+1,n+2：θn-1,n、θn,n+1、θn+1,n+2 andθn+2,n+3.

$$\theta = \begin{cases} 0 & (X_{n+1} > X_n) \\ 90 & (Y_{n+1} > Y_n) \\ 180 & (X_{n+1} < X_n) \\ 270 & (Y_{n+1} < Y_n) \end{cases}$$ (1)

Give the allowable value ε.

Estimating and managing them as fellows:

1) There is a line, and two line are joined by it If the direction of the previous line is different from the one of the next line, i.e. θn-1,n ≠ θn+1,n+2，the node n, n+1 should be added into the new curve. The compute node should be moved to the node number n+1 and the computing will be continued. (Figure 2(a)).

2) if D n,n+1 $> \varepsilon$:
if D n-1,n $> \varepsilon$,the node n will be added into the new curve(Figure 2(b)). if D n+1,n+2 $> \varepsilon$，the node n+1 should be added into the new curve(Figure 2(c)). The compute node should be moved to the node number n+1 and the computing will be continued.

3) if D n,n+1 $< \varepsilon$:

a) If D n-1,n $> \varepsilon$, and D n+1,n+2 $< \varepsilon$, the node n will be added into the new curve(Figure 2(d)).If D n-1,n $< \varepsilon$, and D n+1,n+2 $> \varepsilon$, the node n+1 should be added into the new curve(Figure 2(e)). The compute node should be moved to the node number n+1 and the computing will be continued.

b) If D n-1,n $> \varepsilon$,and D n+1,n+2 $> \varepsilon$, andθn,n+1 =θn+2,n+3 , the midpoint of the node n and n+1 should be added into the new curve. The compute node should be moved to the node number n+1 and the computing will be continued. (Figure 2(f)).

c) If D n-1,n $< \varepsilon$,and D n+1,n+2 $< \varepsilon$, andθn,n+1 =θn+2,n+3 , the midpoint of the node n and n+1 should be obtained. If the distance from the midpoint to the node which was added into the new curve latest is bigger than the value ε , the midpoint should be added into the new curve. The compute node should be moved to the node number n+1 and the computing will be continued. (Figure 2(g)).
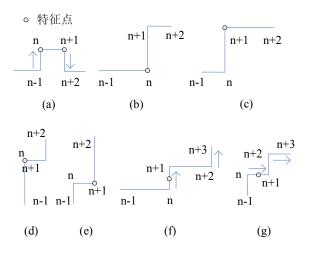

Figure 2. Several instances of eliminating sawtooth

2.1.2 Reducing redundant points

Points that have little effect on shape of circle are produced though the process of removing zigzag shape. If the curve is non closed curve, select the second point as the start point in the curve (if the curve is closed curve, any point in the curve can be appointed as start point); compute the slopes of the first line that is composed of this point and its previous point in curve and second line that is composed of this point and its next point in curve. If the subtraction of two slopes less than pre-given value, this point should be deleted from the curve. During this process redundant point in the curve can be deleted effectively.

2.1.3 Combination of short-distance characteristic points

A condition whether the characteristic points only have short distance is not taken into account in the process described above. That may causes the distances of each neighbouring characteristic points are all less than the pre-given value. This condition has less effect on polygon which has crooked

boundary, but toward to polygon which has straight boundary these points should be disposed. This step is basis of practice.

2.1.4 Experimentation of algorithm

Realization and experimentation of algorithm is based on VC#.NET. As shown in figure3a, any polygon processed after removing indention, variety of area changes between pre-processed and after-processed is 0.5%. Though removing indention, Reducing redundant points and Combination of short-distance characteristic points, trapezia shown in figure3b only has four characteristic points. Variety of area changes is 0.3%. Ellipse in figure3c does the same process as trapezia, variety is 0.2%.
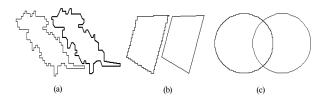

    (a)   (b)   (c)

Figure 3. Experimentation of algorithm

3. CONCLUSION

This method which was proved in practice has the advantages of rapider speed, less RAM possessed and much more excellent ability of obtaining vector image compare with non-processed figures. It is appropriate for regular and irregular polygon extraction, especially for the auto-process of regular object process, and can meet needs of processing this kind of figure.

This article only concentrates on smoothing method after vectorization, topology relationship is not taken into account. That will be researched in the future.

REFERENCES

Huang B．Chen Y. New Approaches for Mutual Transferring of Vector and Raster[J]. Remote Sensing Technology and Application, 1995, 10(3):61-65.

Li Z C, Liu C Y. An Efficient Directed Boundary Method for Vectorization of Dot Matrix Image[J]. Computer Applications and Software, 1997, 14(3):48-51.

Zhang X C, Pan Y H. Vectorization Technique for GIS Grid Data Based on "Grid Technique"[J]. Journal of Computer-aided Design & Compter Graphics, 2001, 13(10):895-900.

Shen Z Q, Wang R C. Study on a New Method for Transferring Grid to Vector Using the Topological Theory[J]. Journal of Remote Sensing, 1999, 3(1):38-42.

Chen Renxi, Zhao Zhongming, Pan Jing. A Fast Method of Vectorization for RS Classified Raster Map. Journal of remote sensing, Vo1．10，No．3 May，2006:326-331.