The Spatial Characteristics of a Line and Their Application to Line Simplification

GÜO Qingsheng Christoph Brandenberger Lorenz Hurni

1 Introduction

Map generalization is a classical problem in cartography. Line simplification is regarded by many cartographers as the most important generalization operation. Simplification algorithms can be assigned to five categories\(^1\). With different line generalization algorithms, the authors obtained different generalization results. Therefore, many researchers have evaluated and compared the existing algorithms\(^2\).

Maps are used to communicate geographic spatial knowledge. The result of a map generalization should remain the important geographic spatial knowledge, according to the conditions of the new small-scaled map. Therefore, in the process of line generalization, the spatial knowledge related to a line should be maintained. The spatial knowledge related to a line consists of geometric and semantic knowledge as well as spatial relations between the lines and other neighbouring cartographic features. In the following, the geometric characteristics of cartographic lines are investigated. How to integrate these characteristic points into generalization algorithms for an improved application is one of the main research topics.

2 Spatial knowledge in line simplification

One task in map generalization and cartography in general is to build an abstraction of the geographic spatial information. The well-generalized map should correctly communicate the geographic spatial knowledge. Therefore, the geographic spatial knowledge of a map must be retained in the process of map generalization, according to the requirement of the generalized map, e.g., characteristic points, spatial relations between lines...

---

Project supported by NSF(No.49971068), the National Excellent Young Teacher Fund and Huo Yingdong Fund for Young Teacher in University.

GÜO Qingsheng, professor, Ph. D supervisor, School of Resource and Environment Science, Wuhan University, 129 Luoyu Road, Wuhan 430079, China
E-mail: Qingshengguo@yahoo. com
and other neighboring cartographic objects, etc. In this section, the characteristic points are discussed, and the algorithms for extracting characteristic points are explained.

2.1 Building inflection points and maximal points of a line

In order to extract automatically the inflection points of a line, the determinant of two vectors can be used for detecting these points\cite{3}. For example, in Fig. 1(a), the determinant of two vectors \((X1 \times Y2 - X2 \times Y1)\) is greater than zero, but the determinant of two vectors in Fig. 1(b) is less than zero. This special property of the determinant can be used for detecting the inflection points of a line. In Fig. 1(c), the principle for identification of inflections is shown. The spatial relation of the direction between two neighboring line segments is changing at Point B in Fig. 1(c). But in fact, the flexion lays in the center of line segment AB. This new inflection point must be inserted in the process of calculating the determinant (Fig. 1(c)). Otherwise, the local maximal point may be the inflection, in Fig. 2(a), Point A is a maximal point as well as a inflection point. In order to solve this problem, the inflection point must be inserted. In Fig. 2(b), Point A is an inflection point, and Point B is a maximal point. Points A, B, C, D and E in Figs. 2(c) and 2(d) are evidently points with local mathematical maximum values.

In this paper, three kinds of maximal points are discussed, as can be seen in Fig. 3. In Fig. 3(a), Point A is the point having the largest perpendicular distance to the base line in comparison with other points of this bend. Fig. 3(b) shows that Point A has the largest distance from this point to the middle point of base line. The sinuosity of a line section is a very important characteristic\cite{11}. In Fig. 3(c), Point A has the maximal sinuosity. Thereby, a local sinuosity can be measured using the method of Dutton (1999). The local sinuosity at the vertex A (Fig. 4(a)) is computed by constructing a ratio of distance between the vertex B and C along the line to the length of an anchor line centered at the given vertex A. For example, in this test, the local sinuosity at vertex A can be computed by means of three different line sections in Figs. 4(b), 4(c) and 4(d). The value of sinuosity at the vertex A is the average value of the three values of local sinuosity at vertex A\cite{4}.
2.2 The points on a convex hull and a minimum bounding rectangle

In computational geometry text books, numerous algorithms for building the convex hull and the minimum bounding rectangle are given. In this paper, the algorithm by Laszlo is applied on the test data\textsuperscript{[5]}. Before a line is generalized, points on the convex hull are determined and the minimum bounding rectangle is extracted, and all of them are stored. On a polygon or a polyline, all points on the convex hull and the minimum bounding rectangle must be extracted in the process of line simplification.

2.3 The monotone line

Many algorithms for identification of monotone line are given in books about computational geometry. Generally a monotone line is defined as a line, where coordinates $Y$ (or $X$) are increasing progressively. In this paper, two new kinds of monotone lines are presented. One is the optimal monotone line (Fig. 5). The other is the monotone line on an oblique axis (Fig. 6). These two sorts of new monotone lines are used to fulfill the special requirements of line generalization.

1) The start point of an original cartographic line is selected, and is linked to the following sequent vertices along this line. Then, the corresponding monotone polylines in $X$-axis and $Y$-axis direction could be identified.

2) By comparing the two identified monotone polylines, the longer one is chosen, and its end point will be stored.

3) The end point of the current monotone polyline is regarded as the start point of the next monotone polyline. Return to the step 1) until the end point of the current monotone polyline is the end point of the original cartographic line.

The algorithm for identifying the oblique monotone polyline is as follows:

1) The start point of an original cartographic line is linked with its second vertex. This line segment can be regarded as an oblique $X$-axis, with the start point as the origin of a local coordinate system (line segments $OA$ and $OB$, in Fig. 7).

2) If the polyline between these two vertices is monotone, the following sequent vertex of the current vertex is linked with the start point of this line, e.g. $B$ is the following sequent vertex of $A$. Continuously whether this polyline is monotone in the current coordinate system or not is tested.

3) If the polyline between these two vertex is not monotone, then the oblique monotone polyline is identified. The polyline from the start point to the neighbored vertex before the current vertex is a monotone line and the enumerated number of this vertex is stored. The current vertex is regarded again as the start point of the next monotone polyline. The next monotone polyline on this original cartographic line will be identified.

4) If the current vertex is the end point of the original cartographic line, the process of identifying the monotone polylines is finished.

2.4 The identification of line bends

A bend (Fig. 8) is already described in detail in
numerous sources about cartographic generalization. But the automatical identification of bends is still a problem which has to be solved. In this paper, an algorithm for identifying bends is shown. Bends are generally regarded as an elementary unit in automatical line generalization as well as in the traditional process. Fig. 8 shows the elementary cartographic unit for line generalization, namely the bend. In fact, there could exist different shapes of bends within a cartographic line. In this paper, a new algorithm for line bends identification is discussed. Using this algorithm, the sought bend conforms with the visual impression of the map reader. The algorithm runs as follows:

![Fig. 8 A bend](image)

**Step 1:** The start point of the original cartographic line is selected. Then this start point is linked with its sequent point (Fig. 9(a)). A is the start point, and B is the next sequent second vertex. C is the point next to Point B. The vertex D is the other neighboring vertex of vertex B.

**Step 2:** Identifying whether vertices D and C are situated on the same side of the line segment AB or not.

**Step 3:** If vertices D and C are on the same side of the line segment AB, then the bend between the two vertices A and B is searched, and vertex B will be fixed as the start point of the next bend (Fig. 9(a)). And the process continuously returns to Step 1. If vertex B is the end point of the original line, then this process will end. And the process switches to Step 5 below.

**Step 4:** If vertices D and C are respectively on an opposite side of the line segment AB, then the following vertex C is tested (Fig. 9(b)). And the process returns to Step 1.

**Step 5:** After carrying out the half-plane, the other half-plane will be handled. If both have been treated, then the entire process will stop. First the start point is identified, which has the maximal curvature of the first bend in all the found bends, e.g. vertex D in the Fig. 9(b). Thus the process return to Step 1.

![Fig. 9 The identification of a bend](image)

3 Improvement of the Douglas-Peucker algorithm

The D-P algorithm of line generalization is classified as a global routine process. During the line generalization process, only the maximal deviations from the base line are considered. But there exists other spatial knowledge for a line, which should be considered in the process of line simplification, e.g. the points on the convex hull and the minimum bounding rectangle. The process of data handling with the improved algorithm runs as follows.

1) Identifying whether the line is an unclosed polyline or a closed polygon.

2) Looking for the points on the convex hull of this line, the inflection points and the maximal points.

3) According to these points, the original cartographic line is divided into many polylines. If the original cartographic line is closed, then the coordinate sequence must be adjusted, according to two points that have a longer distance on the boundary than other points. These two points are the endpoints of the new line to be handled.

4) Based on the line segmentation, every polyline is handled with the help of the D-P algorithm.

5) On the entire generalized line, the D-P algorithm is again used to get the final result.

Evidently the monotone polyline is a simple line that should not be self-intersecting after being treated with the D-P algorithm. To avoid this effect, an algorithm for identifying oblique monotone polylines is used for the line segmentation process. Monotone lines are treated then with the D-P algorithm. This process is continuously repeated until no more points of this line will be eliminated. Finally the entire line is once again handled with the D-P...
algorithm.

The polyline \( BDC \) in Fig. 10 is the result with the original D-P algorithm, and with the same thresholds, the polyline \( BAC \) in Fig. 10 is the result with the improved D-P algorithm taking monotone lines into account.

Fig. 11 shows the original test data which are downloaded from http://crusty.cr.usgs.gov/coast/getcoasts.html.

In Fig. 12, a comparison of generalized lines with an original D-P algorithm and its improved version is represented. Figs. 12(a), 12(d) and 12(g) are the results from D-P algorithm, Figs. 12(b), 12(e) and 12(h) are the results from improved D-P algorithm considering points on a convex hull and maximal points. Figs. 12(c), 12(f) and 12(i) are the results from improved D-P algorithm considering the monotone lines. The given thresholds are the same in Figs. 12(a), 12(b) and 12(c). The thresholds of graphs in Figs. 12(d), 12(e) and 12(f) are also equal, so are those in Figs. 12(g), 12(h) and 12(i). From these graphs (Fig. 12) and the test, it can be seen that the different effects are clear.

4 Improvement of the Visvalingam-Whyatt algorithm

The basic principle of the Visvalingam-Whyatt algorithm is that the minimal triangles built by two consecutive line segments are repeatedly eliminated until the area of the minimal triangle is greater than a given threshold\(^{6}\). This method is based on the natural visual phenomenon, when a spatial object is observed from near to far. However, this progress shifts from the local feature to the global, without considering the global line characteristics. In order to resolve this problem, global characteristic points are added into this generalization algorithm. The process of data handling with the improved algorithm is as follows.

1) Looking for the characteristic points on convex hull and other characteristic points, e.g. maximal points of the original line data in Fig. 11.
2) Based on these characteristic points, the original cartographic line is segmented.
3) On every segmented original polyline, the Visvalingam-Whyatt algorithm is used, in order to get an intermediate result of line generalization.
4) Finally the same process is used once again for the entire generalized line.

In Fig. 13 a comparison from the results of the Visvalingam-Whyatt algorithm and its improved version is represented. Fig. 11 contains the test data. Figs. 13(a) and 13(d) are the results having the same degree of generalization. But Fig. 13(a) is the result from the original Visvalingam-Whyatt algorithm, Fig. 13(d) is the result from the improved Visvalingam-Whyatt algorithm, Figs. 13(b) and 13(e), as well as Figs. 13(c) and 13(f) are similar with Figs. 13(a) and 13(d), but with more points eliminated.
5 Line simplification algorithm based on line bends

Bends are elementary units in line generalization. In Section 2, the algorithm for identifying line bends has been shown. The line generalization algorithm based on bends and that by Visvalingam-Whyatt are quite similar, but the minimal unit is a bend, and not a triangle as in the Visvalingam-Whyatt algorithm. The main principle of this algorithm is as follows.

1) Identifying all bends on the line to be generalized by the method in Section 2.4. In this process, curious phenomena could occur. It is possible that the same bend can be located on both sides of the line segment. See Points A, B, C and D in Fig. 14(a). The bend from Point B to Point C is curious. If the process of identifying bends on the line starts from the other end point of the original line, then the other bends on this line can be found (Fig. 14(b)). In Fig. 14, Point O is the start point. Points A, B, C, D, E and F in Fig. 14(b) are the curious points. But generally the bends on a line should be similar to bends in Fig. 15, e.g. Bends 1-6. Points A, B, C, D, E, F, G and H in Fig. 15 can be identified by the method in Section 2.4.

2) Looking for the minimal bend which has the smallest area among all bends. This bend is then eliminated if its area is smaller than a given area threshold. In every step, the new generalized line is newly composed. And bends of this line will be identified again (Fig. 16). If the area of a smallest bend is greater than the given area threshold, then the entire process will terminate. For example, a result of a generalized line with the bend-based algorithm can be seen as reference in Fig. 17. In Fig. 16, Point O (left-top end point) is the start point of this line, and the change of bends in the process of line generalization is shown.
References

1. Gore A (1998) The digital earth: understanding our planet in the 21st century. Los Angeles, California, http://www.digitalearth.net.cn
2. The First European Conference on Precision Agriculture (1997) Warwick University, U.K., 8-10
3. National Research Council (1997) Precision Agriculture in the 21st Century, Geospatial and Information Technologies in Crop Management, Washington D.C.: National Academic Press.
4. Ouyang X G, Zhang J (2001) Some thoughts on the transformation of agricultural information. Review of China Agricultural Science and Technology, (3):76-80
5. Xu G H (2001) Strengthen rely on innovation of science and technology to bring about more rapid advance of agriculture in western China. In: Proceedings of China Agricultural Hi-tech Forum[C]. Shanxi: Shanxi Propels Press. 15-24
6. Zhang X (2001) Strengthen international cooperation and exchange to promote China's agricultural Hi-technology. Proceedings of China's Agricultural Hi-tech Forum [C]. Shanxi: Shanxi Propels Press. 25-29
7. Xu G H, Sun S (1999) Meeting to the challenge of "Digital Earth". Digital Earth [C]. Beijing: Chinese Science Environment Press. 6-10
8. Cui W H, Li X J (1999) Science Engineering of "Digital Earth". Digital Earth [C]. Beijing: Chinese Science Environment Press. 14-19
9. Cui W H, Li X J (1999) Agriculture science engineering of "Digital Earth". Digital Earth [C]. Beijing: Chinese Science Environment Press. 157-164

(Continued from Page 67)

6 Analysis of the results and the future research work

The aim of the line simplification is the representation of the linear spatial knowledge on a higher level or for small-scaled maps. A line generalization with the improved D-P algorithm leads to a better result than those with the original D-P algorithm. So does that with Visvalingam-Whyatt algorithm. The algorithm for line generalization based on bends sometimes produces strange results as in Fig. 14 and Fig. 17. The algorithm based on the bend should be specially suitable for the natural cartographic linear features. In line generalization the spatial relations are an important factor, and must be absolutely considered.

References

1. Van Kreveld M., Nievergelt J., Roos T., Widmayer P. (1997) Algorithmic foundation of geographic information systems, Berlin: Springer-Verlag, 99-112
2. Visvalingam M., Williamson P. J. (1995) Simplification and generalization of large-scale data for road: comparison of two filtering algorithms. Cartography and Geographic Information Systems, 22(4):263-275
3. Guo Q S (1998) Research on new theory and techniques in automated map generalization; [Ph. D Thesis]. Wuhan: Wuhan Technical University of Surveying and Mapping. (in Chinese)
4. Dutton G. H. (1999) A hierarchical coordinate system for geoprocessing and cartography. Berlin: Springer-Verlag.
5. Laszlo M. J. (1996). Computational geometry and computer graphics In C + +. Englewood Cliffs, NJ: Prentice-Hall.
6. Visvalingam M., Whyatt J. D. (1993) Line generalization by repeated elimination of the smallest area. The Cartographic Journal, 30(1):46-51