The Extracting of 2D Plot Knowledge in Chart Automatic Generalization

To cite this article: Ying Zhu et al 2019 IOP Conf. Ser.: Earth Environ. Sci. 237 032004

View the article online for updates and enhancements.
The Extracting of 2D Plot Knowledge in Chart Automatic Generalization

ZHU Ying¹, LU Yi¹, LI Zhi-wei², WEI Lan¹, CUI Yang³

¹ Naval Institute of Hydrographic and Charting, Tianjin, 300061, China
² China Navy Hydrographic Office, Beijing, 100036, China
³ Department of Hydrography and Cartography, Dalian Naval Academy, Dalian, Liaoning, 116018, China

Corresponding author’s e-mail: suzy_hych@163.com

Abstract: In order to keep the facticity of generalization process, more attention should be paid on the feature of datum. According to the analysis of familiar 2D plot knowledge, the graphic and geographic feature can be pre-digested in the electronic chart automatic generalization process. At the same time, the maintenance of facticity should be considered.

1. Introduction

The extraction of geographical features (except attribute features) is mapped to two-dimensional graphs, which is the recognition and measurement (i.e. quantization). If the comprehensive method only pays attention to the skills of graph transformation or graph processing, but ignores the restriction of the element's own characteristics to the comprehensive method, it will make many algorithms have similar comprehensive results, and this kind of comprehensive result is not the most reasonable. The graphical features and geographical features of the elements themselves should be considered at the same time of the simplification of the graphics, and different comprehensive principles should be adopted according to different geographical contents. Only in this way, the comprehensive results of the elements can conform to the geographical features expressed by them and effectively transmit geographic meanings and knowledge. Thus, extracting knowledge from the graph itself is commonly acknowledged.

2. Overview and analysis of 2d graphics

Two-dimensional graphics are the mapping of spatial information on the plane, which takes up the main components of spatial information. In fact, it is a pattern recognition in a special sense and a reflection of computer vision that makes computers process graphics information automatically. There are many kinds of knowledge and complex content of two-dimensional graphics. This paper only discusses several aspects closely related to cartography synthesis. The knowledge of 2d graphics includes the graphical features of the elements themselves and the spatial relations among the elements. It is mainly divided into the following categories.

2.1 Location

The spatial position is to transfer the individual positioning information of the spatial object by means of the spatial coordinate system. It is transformed to the plane by the projection, and is usually
represented by the rectangular coordinates \((x, y)\).

2.2 Form
Form is the most basic geometric feature of the figure, and it is also the most easily perceived visually. For example, when we look at a line, the first thing we feel is whether it's a straight line or a curve. When we look at a facet, the first thing to feel is whether it is round or square. This shape feature can be directly felt and judged by us. There are other geometric features that must be described by numerical values, such as length of linear elements, bending degree and fractal dimension and axis, centroid, contour circumference, area and convex hull of planar elements, the slope direction, slope degree and structure line of curved surface, too.

2.3 Distribution
Spatial distribution is a kind of group positioning information reflecting similar space objects. The parameters describing spatial distribution include distribution center, standard distance, distribution density and axis. Here, point group target is an important object of spatial distribution analysis, and its purpose is to extract spatial structural information. The characteristic parameters of point group spatial distribution are defined as distribution density, distribution center, distribution axis and dispersion based on statistical analysis method. From the perspective of visual recognition, we can establish three concepts of neighborhood, distribution center and density change of point group distribution. In the field of computational geometry, point group distribution belongs to the study of spatial proximity, and Delaunay triangulation and its dual Voronoi diagram are important analysis tools. This paper uses these two models to analyze the characteristic information of the point group in the two-dimensional figures and simplify the point group through Voronoi diagram dynamic reconstruction.

2.4 Topological relation
Topological relation is the spatial relation between spatial objects without regard to measurement and direction. The figure in GIS uses some map projection in a certain area, which is topological mapping of ground projection onto flat surface. In topological mapping, points are mapped to points, lines are mapped to lines, and faces are mapped to faces. Points within a surface are mapped to remain within a surface. Therefore, points, lines and planes in GIS have topological properties, and they are basic topological elements. In practical application, the topological relation is generally summarized into four types, namely adjacent \((A|B)\), phase separation \((A||B)\), strictly including \((A<B)\), and intersection \((A\times B)\). As shown in figure 1, \(A\) is the boundary of \(A\), and \(A_0\) is the inner domain of \(A\). As we need to distinguish the topological relations between different types of spatial elements in the concrete work, the possible spatial relations between point, line and surface are shown in Table 1. As shown in the table, there are 18 different typologies. For cartographic generalization, the invariability of topological relation before and after synthesis is one of the basic principles of automatic synthesis.

![Four basic topological relationships](image)

(1) \(A|B\): \(A \cap B = \partial A \cap \partial B \neq \emptyset\)
(2) \(A||B\): \(A \cap B = \emptyset\)
(3) \(A<B\): \(A \subset B\)
(4) \(A\times B\): \(A^0 \cap B^0 \neq \emptyset\)

Figure 1. Four basic topological relationships
Table 1. Topological relations between points, lines and surface space elements

|                  | points                        | lines                  | surface space elements |
|------------------|-------------------------------|------------------------|------------------------|
| points           | • phase separation            | • phase separation     | • phase separation     |
|                  | • same position               | • namely adjacent      | • namely adjacent      |
|                  |                               | • intersection         | • intersection         |
|                  |                               | • strictly including   | • strictly including   |
| lines            |                               | • phase separation     | • phase separation     |
|                  |                               | • intersection         | • namely adjacent      |
|                  |                               | • same position        | • intersection         |
|                  |                               |                        | • strictly including   |
| surface space    |                              |                        |                        |
| elements         |                              |                        |                        |

3. Knowledge analysis of 2d graphics

Generally, 2d graphics can be divided into point, line, plane, point group, line group and plane group. Since there is a little knowledge about 2d points, this paper mainly discusses the information of their own characteristics and the relationship among them.

3.1 Graphic knowledge of lines

3.1.1 Length The length of linear elements is one of the most basic morphological parameters. In two-dimensional vector data, linear element L is represented as a coordinate string, so the calculation formula of linear element length is as follows:

\[ l = \sum_{i=0}^{n-1} \left[ \left( X_{i+1} - X_i \right)^2 + \left( Y_{i+1} - Y_i \right)^2 \right]^{1/2} = \sum_{i=1}^{n} l_i \]  

(1)

3.1.2 Fractal dimension Fractal dimension can be used to describe the complexity of linear elements. The measured length of the curve varies with the ruler. The smaller the ruler is, the longer the
measurement length is, and the change rate is controlled by the curve fractal dimension H-B dimension (Hausdorff - Besicovitch). See figure 2.

![Figure 2](image_url)

Figure 2 .Measurement of curve length

When the step size d is used to approximate the curve, it is known that the length of the curve is 7d. When the step size decreased by half, the length of the curve changed to 16.5 (d/2), or 8.25d, and so on. The d here is what Mandelbrot calls a ruler. The fractional $D_f$ of the curve has the following relationship with length l:

$$l = n \cdot d = Kd^{-D_f} \cdot d = Kd^{1-D_f}$$  \hspace{1cm} (2)

Take the logarithm of both sides, and you get:

$$\log l = \log K + (1-D_f) \log d$$  \hspace{1cm} (3)

$K$ depends on the initial operating length of the fractal curve and can be regarded as a constant. We can see that equation (3) is a linear equation with $\log l$, $\log d$ are variables and $1-D_f$ as slope. Therefore, as long as the step length is a monotonically increasing sequence, the corresponding length of the curve can be obtained, and then the fitting point pair of linear regression model can be used, then the $D_f$ value can be obtained according to equation (3).

$$D_f = 1 - \frac{\sum_{i=1}^{N}[\log(d_i)\log L(d_i)] - \frac{1}{N}\sum_{i=1}^{N}\log(d_i)\sum_{i=1}^{N}\log L(d_i)}{\sum_{i=1}^{N}(\log d_i)^2 - \frac{1}{N}(\sum_{i=1}^{N}\log d_i)^2}$$  \hspace{1cm} (4)

3.1.3 Curvature and curvature

According to the mathematical analysis, the curvature of a linear object is defined as the rotation rate of the tangent direction angle relative to the arc length. If the curve is defined as $y = f(x)$, the curvature of the curve at any point is:

$$K = \frac{y'}{(1+y'^2)^{3/2}}$$  \hspace{1cm} (5)

The curvature reflects the local bending characteristics of the curve. In order to reflect the overall bending characteristics of the curve, the average curvature of the curve should be calculated:

$$\bar{K} = \frac{1}{n} \sum_{i=1}^{n} K_i$$  \hspace{1cm} (6)

Average curvature reflects the degree of curvature of the curve.

The curvature of the curve is another parameter that describes the bending degree of the curve, as shown in figure 3.

![Figure 3](image_url)

Figure 3 .Curvature of the curve

The ratio of the length of line segment between the curve length and the end of the curve is defined as:

$$S = \frac{L}{l}$$  \hspace{1cm} (7)

It is meaningless to study the bending degree of linear objects abstractly. In practical application, bending degree S is not mainly used to describe the bending degree of linear objects, but to reflect the curve’s circuitous characteristics.
3.2 Graphic knowledge

3.2.1 Area, circumference and shape
Area and perimeter are the most basic morphological parameters of planar objects. In GIS (vector data), planar objects are generally represented by polygons formed by their contour boundary. Where, $P = (p_0, p_1, \ldots, p_n) = ((x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n))$, and $p_i = (x_i, y_i) \in P$, $p_0 = p_n$, then the area of the surface domain whose contour boundary is $P$ is calculated as:

$$S = \frac{1}{2} \sum_{i=0}^{n-1} |x_i y_{i+1} - y_i x_{i+1}|$$  \hspace{1cm} (8)

The purpose of taking the absolute value is to free it from the direction of the polygon loop. The calculation of the circumference is calculated according to the length formula (1) of the above curve. For the quantitative description of the shape of planar elements, there are many expert research methods, such as scale-based description, regression analysis, and bi-axial Fourier shape analysis, all of which analyze the shape of planar elements from the pure geometrical meaning. For automatic synthesis, the main shape features of the surface should remain unchanged before and after synthesis, and because the contour line reflecting the shape of the surface itself is a closed curve, it can be completely transformed into the extraction of feature points of the closed curve, which will be discussed in detail below.

3.2.2 Surface directivity
Also known as spatial ductility or orientation, it can be described by its smallest external rectangle. As shown in figure 4, the trend of surface $P$ can be determined by the size and direction of its long and short axes $A_1$ and $A_2$. Where, $A_1$ and $A_2$ are determined as follows: suppose the center of gravity of $P$ is $G$, and the line between the most distant two points in $P$ (as shown in the dotted line in FIG. 4) is shifted to cross $G$, and $A_1$ is obtained. The four extreme points of $A_1$ and $P$ (maximum x, maximum y, minimum x and minimum y) are used to determine the external rectangle, so $A_2$ is the line segment perpendicular to $A_1$ and falling within the external rectangle.

![Minimum external rectangle for a simply connected surface domain](image)

3.2.3 Surface skeleton line
The skeleton line of the surface is the extraction of the one-dimensional features of the opposite shape elements, which can be regarded as the inverse solution of buffer problem in GIS. Skeleton line can be used as an important basis for surface simplification, especially when it is tree-like or obviously extended on the face; it is the positioning reference line after graphic simplification and dimension reduction. See figure 5.

![Surface skeleton line](image)

3.2.4 The convex hull
For a complex planar object, it is sometimes necessary to use relatively simple shapes to describe it generally, among which, the smallest convex shell is the closest to the original polygon compared with the external rectangle or circle. See figure 6.
3.3 Figure knowledge of point groups

3.3.1 Distribution range The distribution range of point groups is a problem with uncertainty. Generally speaking, polygons representing the distribution range of a group of points are not unique, and the fuzzy principle of selection is to conform to people's visual habits. As shown in figure 7, it is clear that figure b is more reasonable than figure a, and figure c is more reasonable than figure b. In the traditional method, it is the most common and easy to use convex hull to express the range of point group. However, it can't be applied to any point group. If the point group has a large area of concave part, the convex hull will produce greater distortion, as shown in fig.c. In this paper, the Delaunay method is proposed, that is, to detect the triangle associated with the boundary of the Delaunay triangulation. It is worth mentioning that because only boundary triangles are detected, it is independent of the density of internal point group distribution and will not cause the phenomenon of inner cavity due to the limit of edge length.

3.3.2 Distribution density The distribution density of a point group is usually defined as the number of points in a unit area. By using reverse thinking, when taking each point as the basis, we can also take the space of a certain number of points as the measurement of the distribution density. For a point group, we can determine its average density, as a basis for comparison with other point groups, and also determine its own distribution density of high and low.

In point group, we can think of each point for its survival space, if not the nature of the differences between point and point, that is between them is equal, no weak points of the, it's the space between adjacent points should be to a fraction, such as the idea in conformity with the Voronoi diagram in computational geometry, Voronoi diagram polygon yeand just expressed it contains some space.
3.3.3 Distribution center  

The distribution center can be roughly represented as the overall position of the distribution, which can be divided into four categories: arithmetic mean center, weighted mean center, median center and extreme center. 

With n discrete points Pi, its plane position is (Xi, Yi), then:

- arithmetic mean center
  \[
  \overline{X} = \frac{1}{n}\sum_{i=1}^{n}X_i, \quad \overline{Y} = \frac{1}{n}\sum_{i=1}^{n}Y_i
  \]

- weighted average center
  \[
  \overline{X}_w = \frac{\sum_{i=1}^{n}X_i \times W(P_i)}{\sum_{i=1}^{n}W(P_i)}, \quad \overline{Y}_w = \frac{\sum_{i=1}^{n}Y_i \times W(P_i)}{\sum_{i=1}^{n}W(P_i)}
  \]

  Where, \(w\) is the weight of.

- median center
  The median center is similar to the concept of the median, which is determined by the following formula:
  \[
  \text{min}((X_i - \overline{X}_w)^2 + (Y_i - \overline{Y}_w)^2) = \text{min}
  \]

- extreme center
  For every \((X, Y) \neq (X_e, Y_e)\), the following is true:
  \[
  \text{max}((X - X_e)^2 + (Y - Y_e)^2) > \text{max}((X - X_e)^2 + (Y - Y_e)^2)
  \]

  Which, in the sum of the distance from the center point group in every other point to a minimum, the center of extreme point group of each point in the maximum distance than any other point relative to the point of farthest point distance is smaller, geographical significance is to make that point group all the points are not too far, in the general application in actual on the important feature of the location, such as the fire brigade.

3.4 Graphic knowledge of surface groups

Because the face group can be equivalent to the organic combination of face field and point group, the face group has the properties of both face and point group, which are mainly reflected in its structural features and spatial relations. The spatial relationship between surface group and other spatial objects; Voronoi diagram or adjacent correlation diagram of the face group itself.

It should be noted that there is still a lot of knowledge about two-dimensional graphics. In view of the theme and space, this paper only describes and analyzes the most basic and automatic comprehensive needs of the selection part.

4. Conclusion

This paper elaborates the classification of knowledge of two-dimensional graphics, and describes the characteristic parameters of points, lines and planes and their measurements respectively from the aspects of morphological features, distribution states and topological relations. They are: length of linear elements, fractal dimension, curvature and curvature. Area, girth, shape, direction, skeleton line and convex shell of surface domain; Distribution range, distribution density and distribution center of point group; The structural features and spatial relations of surface groups; The spatial relationship between points, lines, and surfaces. Deep understanding and analysis of the above knowledge representation and mathematical description is the basis for the realization of graph feature recognition and measurement.

References:
[1] Wang Jiayao. Integrated Principles and Methods of Digital Map Automatic Mapping [M].
Beijing: PLA press, 1997.

[2] Zhai jingsheng. Digital Chart Production Technology System [R]. Technical Report, 1998.

[3] Du Jinghai. Chart Editing Design [M]. Beijing: Surveying and Mapping Press, 1994.

[4] Guo Qingsheng. Research on The Automatic Synthesis of New Theories and Methods of Maps [D]. Wuhan: Wuhan Institute of Surveying and Mapping, 1998.

[5] Lu Yi, Sun Yan. Special Treatment and Realization of Polygon Clipping [J]. Marine Surveying and Mapping, 1999, (3).

[6] Xu Haitao. Automatic Comprehensive Study on Elements of Digital Chart Sea Department [D]. Dalian: Dalian Naval Ship Academy, 1996.

[7] Wu Hehai. Study on The Automatic Synthesis of Map Information [J]. Journal of Wuhan University of Surveying and Mapping Technology, 2000.

[8] Liu Lian. Basic Problems of Automatic Map Synthesis [J]. Science and Technology, 2010.

[9] Qu Bangding. Status and Development Direction of Automatic Map Synthesis [J]. Shanxi Architecture, 2008.

[10] Wu Fang, Zhu Kunpeng, Deng Hongyan et al. Analysis of Automatic Comprehensive Quality Analysis and Evaluation Criteria For Maps [J]. Journal of Surveying and Mapping Science and Technology, 2007.

[11] Zhu Jianliang, Xu Dingjie. Study on The Principle and Method of Island Automatic Synthesis [J]. China Navigation, 2004.

[12] Wang Huilian. Research on Intelligent Methods for Automatic Integration of Residential Areas [D]. 2005.