RESEARCH ON TOPOGRAPHIC CONTOUR MAP SOFTCOPY

Ma Fei
(Research Center for GIS, WTUSM, 39, Luoyu Road, Wuhan, 430070, China)

Commission IV, Working Group 3

ABSTRACT:

This paper discusses study and experiments on topographic contour map softcopy. The main aim of this research is to develop a practicable software system on Windows for scanned contour map automatic recognition and then to generate corresponding Digital Elevation Model (DEM). In the system, raster data is organized in the form of quadtree structure, and vector data is manipulated on Freeman code. At first, image processing insists of two aspects, that are frequency domain processing and spatial domain processing. In frequency domain, fast Fourier transforms and wavelet transforms are used to delete noise from optical scanning, and the result helps to contour line detection. In spatial domain, mathematical morphology, a new complete theory on parallel dealing raster data, is used to improve image quality, to segment contour lines from map symbols, and then to implement features of symbols and contour lines. Gray-scale thinning, an efficient algorithm, makes contour thinning and backgrounds optical illumination submerges disappear, and then it is easy to determine threshold to make gray scale image binarilized. As follows, binary morphological thinning makes contour one pixel wide, and short arcs are deleted. Contour gap connecting, which is difficult during the whole procedure and needed in order to make raster data vectorlized, is dealt with on morphological transform and knowledge based deduction. Another difficult question, that is contour height determine, is solved by inference from morphological points and lines. At last, DEM is constructed in the form of triangulated irregular networks. Restore DEM data is discussed as well. Experiments on 1:50,000 topographic contour maps verify the efficiency of the system.

Key Words:

Topographic contour map; Scanned image processing; Digital Elevation Model; Mathematical Morphology; Map recognition

1. INTRODUCTION

The significance of constructing topographic database in a large area such as a district takes attention to those who work on city planning, land management and resource development. The first thing to construct topographic database is to capture enough data for analysis. Conventional geodesy and engineering survey is far from the demand of database revision on neither efficiency nor finance. However, our photogrammetrists around the world have used our own tool to change the dilemma. It is remote sensing image and aerial photographs which contain potential information, and these information are often reflected on map. In general, the method to set topographic database is divided into two parts: one is extract information from existing map, the other followed by the former is to update the data in the database which demonstrates the new situation on image. There are two conventional methods to change map information into topographic database data: one is to digitize the map on the digitizing table with manual, the other is to recognize the map by computer. Obviously, the latter is more efficient than the former. Therefore, more and more people pay their attention to automatic map recognition.

In this paper, a complete procedure for topographic contour map softcopy is presented. The general idea to make topographic contour map digitize and then to generate corresponding digital elevation model (DEM) is introduced in section II. In more detail, scanned map image processing in both frequency and spatial domain is analyzed in section III, and a new automatic raster-vector conversion method is presented in section IV. DEM generation and data restoration are discussed in section V. At last, experiments on a new software system are remarked and the result is concluded in section VI.

2. GENERAL IDEA FOR SCANNED CONTOUR MAP SOFTCOPY

For a long time, toward automatic recognition contour map is an appealing aim to study. This study involves such fields as image processing, pattern recognition, surveying and mapping, geomorphology and database. The difficult is not only extracting information from scanned image but also setting up topological relationship around map and making digitized results meet a given precision. Therefore, the method and tool in the task should be selected carefully and reasonably. At first, image deform arouse from optical scanning must be analyzed to rectify and its characteristics should be fused into image processing. So in the first part of this study, mathematical mapping and Fourier analysis is needed. Secondly, contour lines extraction from a topographic map is not equal to ordinary edge detection due to position precision and the influence of optical shadow. Thus the question insists of two parts: extract watershed lines on gray tone image and disappear background optical illumination submerge. Luckily, gray-scale thinning, an efficient algorithm is found to deal with it powerfully. Thirdly, contour lines in raster form must be changed into vector form to construct a topological database for terrain analysis. It is necessary to set up topological relationship in spite of poor broken contour segments. That's why extracting contour lines from annotation symbols, gap connecting be studied. Furthermore, contour elevation must be determined and a new way by means of geomophological points and lines is presented. Finally, corresponding digital elevation model (DEM) is constructed in the form of triangulated irregular networks(TIN).

3. SCANNED CONTOUR MAP IMAGE PROCESSING

3.1. Data Capture

There are two reasons about contour map scanned in gray tone image rather than binary image: one is about the scanner, the other is about map itself. The point spread function of the scanner leads image to convolution distortion and nonuniform illumination with nonuniform paper reflection make map background submerge. On the other hand, different types of pens and ink might be used on the map to distinguish various kind of symbols.

3.2. Rectification And Resample

Map image from scanner has two aspects deform: one is on scanner, the other is on map itself. As for the former, a standard lattice is used to test rectification coefficients in the form of quadratic polynomial:

$$x_1 = a_0 + a_1x + a_2y + a_3x*x + a_4x*y + a_5y*y$$
$$y_1 = b_0 + b_1x + b_2y + b_3x*x + b_4x*y + b_5y*y \qquad (1)$$

As for the latter, the kilometer lattice on the map can be used to measure the extend of deform in the form of quadratic polynomial also:

$$x_2 = c_0 + c_1x + c_2y + c_3x*x + c_4x*y + c_5y*y$$
$$y_2 = d_0 + d_1x + d_2y + d_3x*x + d_4x*y + d_5y*y \qquad (2)$$

After parameter a_i, b_i, c_i, d_i can be solved by the least squares method, each pixel on the image corresponds to a new position(x_2,y_2). Because x_2,y_2 maybe not integral, so it must give a correct gray value to each pixel ,that is image resample. The resample method is selected as bilinear interpolation[1].

3.3. Image Filter

Image filter is dealt with in both frequency domain and spatial domain. In frequency domain, Fourier analysis method is used to improve image quality because it corresponds to optical scanner characteristic. In order to preserve contour, it must remain the high frequency component of the image, and restrain the low frequency component of the image. So the homomorphic filter is selected and the coefficients of both high frequency and low frequency are carefully chosen.

Recently, wavelet transform is widely used in image processing. Superior to Fourier analysis, wavelet transform contributes good localization properties in both the spatial and frequency domain. Wavelet transform on the scanned image enhances noise-corrupted edges under the multiresolution decomposition and this helps to detect contour lines.

Traditional image processing methods such as Laplacian enhancement and midvalue filter are not used because they haven't selectivity of spatial localization. These methods smooth the contour and make followed steps more difficult. Instead, morphological filter is used to delete little noise such as grain and short arcs. Selecting 3*3 structuring elements as follows:

0	0	0
1	①	1
0	0	0

i=1,2,...8, rotate $\frac{\pi}{2}$ as soon as i increase 1, and

0	0	0
0	①	1
0	1	0

i=1,2,...8, rotate $\frac{\pi}{2}$ as soon as i increase 1

where "①" represents the centre of structure element.

The recursive opening operations of mathematical morphology [3] are needed to eliminate noise in different size and conditional dilation operations are chosen as assisting tool to control the lose of contour lines information.

3.4. Edge Detection

Gray-scale morphology is natural extend of binary morphology from two to three dimensions, it reveals a surprising landscape of beauty and utility[4]. Gray-scale morphology uses the basic concept umbra to represent grayscale data. The significance of the umbra of the umbra of image processing is that they remain umbra under the usual morphological transformations of union and intersection, dilation and erosion. The difference between grayscale morphology and binary morphology is only that the operations in the first domain replace intersection and union with min and max.

Select flat structure elements $\{L_i\}$ as

0	0	0
.	①	.
1	1	1

i=1,3,5,7, rotate $\frac{\pi}{2}$ as soon as i increase 2, and

.	0	0
1	①	0
.	1	.

i= 2, 4, 6, 8, rotate $\frac{\pi}{2}$ as soon as i increase 2

where "." represents 0 or 1.

Operate

$$\mathrm{GSK}(X) = X\,\mathrm{O}\{L_i\} \qquad (3)$$

where X refers to contour map image, "O" represents gray-scale thinning, result in a gray-scale thinned contour image. Because of the finite width of scanned contour (often several pixel width),the operation convergences quickly after no more than 2 or three recursive times. The result preserves the contour in middle axis, disappear the submerge mostly. Furthermore, the result opened with the structure element H clears the submerge completely. That is

$$\mathrm{GSK}'(X) = \mathrm{GSK}(X) \circ H \qquad (4)$$

where "∘" represents open operation, $H = \begin{matrix} 1 & 1 & 1 \\ 1 & ① & 1 \\ 1 & 1 & 1 \end{matrix}$.

Moreover, the contour now is sometimes in jagged way, and its structure reflects as

$$\begin{matrix} . & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & . \end{matrix}$$

i = 1,2,...8 rotate $\frac{\pi}{4}$ as soon as i increase 1

Design these group feature as structuring elements $\{J_i\}$, by means of thinning transformation as

$$\mathrm{GSK}''(X) = \mathrm{GSK}'(X)\,\mathrm{O}\,\{J_i\} \qquad (5)$$

leads to one pixel width contour.

The gray tone contour image now reveals a surprising by-product, that is, the gray level histogram concentrates on the maximum gray value. So luckily, it is easy to determine threshold to segment the contour from the image, The difficult segment problem is solved!

4. AUTOMATIC RASTER TO VECTOR CONVERSION

4.1 Data Structure

The linear quadtree is a data structure which is based on the regular decomposition of a square into quadrants and sub-quadrants[6]. It represents spatial relations only by leaf nodes with numeric keys. Contour map after thresholding is organized in such structure not only compact raster data in a high rate, but also improve the operation speed on morphological transform.

Freeman code is often used to describe a vector data in the form of number series. The number representing the dot near the center is less than 9 when the data are organized in 8-connect form, such as

		top		
	4	3	2	
left	5	0	1	right
	6	7	8	
		bottom		

Using this code to describe a contour helps to learn the properties of a contour such as length, direction and connectivity. Obviously these features are important in contour line identification. The dot on a contour can be caught by recursive morphological thinning with structure elements $\{E_i\}$ such as

$$\begin{matrix} 0 & 0 & 0 \\ 0 & ① & 1 \\ 0 & 0 & 0 \end{matrix}$$

i = 1,2,...8 rotate $\frac{\pi}{4}$ as soon as i increase 1

and the Freeman code of the dot is equal to the number i of series $\{E_i\}$.

4.2 Contour Line Extraction

To extract contour line from annotation symbols on a contour map is the first step of raster to vector conversion. In general , the length of contour is far from the length of symbols, and contour is not crossed by themselves or each other. Although this way can separate most contour from the other part of map, some little arcs which are broken from the contour by manual annotation are filtered also, but they are often important to show features of the contour. So the method should be improved.

In order to provide a precise mathematical description of the important shape-size of a scanned contour map image under consideration, the pattern spectrum is utilized. The pattern spectrum PSK(X) of a set X in terms of the structuring element B is given by

$$\mathrm{PSK}(X) = \begin{matrix} \mathrm{Card}[Y_k(X) - Y_{k+1}(X)],\ \text{for } k >= 0 \\ \mathrm{Card}[\Phi_{-k}(X) - \Phi_{-k-1}(X)], \text{for } k <= -1 \end{matrix} \qquad (6)$$

where k ∈ Z and Card[X] denotes the cardinality of a set X, Y represents open operation and Φ represents close operation. The pattern spectrum provides a measure of similarity between an image and the collection of all possible structure elements kB. So selecting suitable structure elements to describe contour line not annotation symbols such as

$$\begin{matrix} 1 & 0 & 0 \\ 0 & ① & 0 \\ 0 & 0 & 1 \end{matrix}$$

i = 1,2,...8 rotate $\frac{\pi}{4}$ as soon as i increase 1, and

$$\begin{matrix} 1 & 0 & 0 \\ 0 & ① & 0 \\ 0 & 1 & 0 \end{matrix}$$

i = 1,2,...8 rotate $\frac{\pi}{4}$ as soon as i increase 1

contour line contributes to lower size part of the pattern spectrum, where annotation symbols contribute to higher size part of the pattern spectrum. Therefore the contour line is distinguished from annotation symbols and is extracted easily. As you may know, the method is also suitable for annotation symbols recognition.

4.3 Contour Gap Connecting

It is necessary to deal with the "bottle-neck" problem of contour gap connecting, in order to set topological relations among contour lines and then to realize raster to vector conversion. It is easy to find the whole end points and their stepped segments recorded in Freeman code. So we define a morphological window rather than rectangle window such as in [5] that contains end points and is surrounded by connected contour lines and their normal lines. Generally, the window contain the matched pair of end points and their relationship (direction, distance) can be determined easily. Then the matched pair of points can be inferred from knowledge.

As you know, contour lines couldn't cross each other, and their elevations are determined. Elevations of contour lines on a monotonous slope is increased one by one between a pair of summit and sink. So the end points appear in pairs in a morphological window; after connecting two end points belong to the same contour, the window is divided into two parts, each part contain other end points in pairs also. After gap connecting, the contour elevation is determined no matter from what direction to infer.

4.4 Contour Line Raster To Vector Conversion

Contour map raster to vector conversion includes geomophological points and lines determining, monotonous slope division, contour elevation deduction and discrete points selection.

Let X represents a contour map after gap connecting, X^c is background of X, SK(X^c) is skeleton of X^c ,then

$$\mathrm{SK}(X^c) = X^c \mathrm{O}\{L_i\} \qquad (7)$$

delete short arcs of SK(X^c), then

$$\mathrm{SK}_z(X^c) = \mathrm{SK}(X^c)\mathrm{O}\{E_i\} \qquad (8)$$

Thus, saddle points set S_1 can be determined as

$$S_1 = \bigcup_{i=1}^{32} (SK_z(X^c) \oplus Q_i) \qquad (9)$$

here Q_i represents three-cross points set in 8-connect, $\oplus$ refers to dilation operation. Then the area A contain summit and sink can be determined as

$$A = (SK_z(X^c))^c \ominus \{H\}; X \qquad (10)$$

where $\ominus$ refers to erosion operation; and summit and sink points set S_2 as

$$S_2 = A\mathrm{O}\{D_i\} \qquad (11)$$

where structure elements $\{D_i\}$ as

$$\begin{matrix} 0 & 0 & . \\ 0 & ① & 1 \\ 0 & 0 & . \end{matrix}$$

i = 1,2,...8 rotate $\frac{\pi}{4}$ as soon as i increase 1, and

$$\begin{matrix} 0 & . & 1 \\ 0 & ① & . \\ 0 & 0 & 0 \end{matrix}$$

i = 1,2,...8 rotate $\frac{\pi}{4}$ as soon as i increase 1.

Then, a monotonous slope can be determined as a area between two geomophological points (saddle, summit or sink). Provide $p \in (S_1 \cup S_2)$,

$$\begin{aligned} W &= p \oplus \{H\}; [N(W(S_1 \cup S_2)) < 2] \\ W' &= W \oplus \{H\}; X^c \end{aligned} \qquad (12)$$

where $N(W(S_1 \cup S_2))$ refers to the cross points between W and $(S_1 \cup S_2)$, thereafter the contour map is divided into monotonous slopes $(W_1', W_2', \cdots, W_n')$.

Extracting one of contour line on a monotonous slope relies on geomorphologic points also. Provide A be a monotonous slope, geomorphologic point $p \in A \cap (S_1 \cup S_2)$,

$$A_0 = p \oplus \{H\}; X^c \qquad (13)$$

then the nearest contour line of slope A from p is

$$A_1 = (A_0 \oplus H) \cap X \qquad (14)$$

replace p with $(A_0 \cup A_1)$,repeat

$$\begin{aligned} A_0' &= (A_0 \cup A_1) \oplus \{H\}; X^c \\ A_1' &= (A_0' \oplus \{H\}) \cap X \end{aligned} \qquad (15)$$

until

$$(A_0' \cup A_1') \oplus \{H\}; X^c \cong A \qquad (16)$$

where " $\cong$ " represents images in left and right are the same. Therefore, the contour is extracted one by one.

Due to terrain rise and fall, the contour elevation monotonously changed near saddle. In order to inference contour elevation, all of summits and sinks must be found; their elevations are deduced from control points elevation annotation. As a provided condition, there is one and only one saddle between two near summit constrainly. The inference method insists of two steps: firstly to determine the number of contour lines of a monotonous slope between one point of summit or sink and one point of saddle, and to set them a sequence in the order of increase, that means, to set a sequence for all the summits or sinks to infer elevation; secondly to infer elevations of contours on a slope from summit or sink to saddle, meanwhile this saddle is no longer valid in the next inference.

The vectorlized contour is stored in the form of discrete points. The density of discrete points is determined with the terrain roughness and interpolation precision[7]. In order to describe the rise and fall of the terrain, three neighbor points keep up a slope, that is, the distance from the middle point to the line connected by the other two points should be more than a threshold which relates to precision.

5. DEM CONSTRUCTION AND DATA RESTORATION

The method of generating TIN to represent DEM is described as follows:

5.1. Homotopic Sequential Thinning For The Skeleton.

Let $X \in Z$, the skeleton S(X) can be described as follows:

$$\mathrm{SK}(X) = X \odot \{L_i\} \qquad (17)$$

where " $\odot$ " represents to thickening operation.

5.2. Generating The TIN By Individual Points

According to the definition of Thissen polygon, the skeleton $S(X^c)$ is just the Thissen polygon corresponding to X in raster domain. With the algorithm of partial painting, it is easy to find the two points in two different Thissen polygon which share a same side. The constructed triangulation through connecting all of these two points is just Delaunay triangulation by definition.

5.3. Generating The TIN In Consideration Of Geomorphologic Points And Lines.

In order to describe terrain exactly, geomorphologic points and lines must be contained in the TIN. That means, geomorphologic points should be triangular points and geomorphologic lines must be preserved as triangular sides.

5.4. DEM Data Restoration

Ordinary topological relationship of TIN containing points, sides and near triangular is too space consuming to represent. It is necessary to compress TIN relationship data in the case of manipulating and analysis. The method of regular storing TIN is to enlarge TIN to regular networks in topology. As you can image, any midpoint polygon can be thought as constitution of hexagon(s) topological, only have some coincided points and lines.

6. EXPERIMENTS AND CONCLUSION

Based on the above procedure, a new scanned contour map recognition system on windows has been designed. Experiments with the system have been done a contour map at scale 1:50000, automatic raster to vector conversion result and part of its corresponding TIN is shown in Fig. 1 and Fig. 2. On experiments, it is verified that a topological contour map without too much complex can be automatically digitized within 48 hours or so on PC 486 without any other hardware.

REFERENCES:

1. Wang Zhizhuo, Principles of Photogrammetry (With Remote Sensing), Press of Wuhan Technical University of Surveying and Mapping, Wuhan, China, 1990.
2. Li Deren, Chen Xiaoyong, Ma Fei. Applications of Mathematical Morphology in automatic Cartography and GIS. International Archives of Photogrammetry. Volume 30-4. ISPRS Commission IV,1994. pp.127-134.
3. Serra,J. Image Analysis and Mathematical Morphology. London: Academic Press, 1982.
4. Sternberg,S.R. Grayscale Morphology. Computer Vision Graphics and Image Processing, Volume 35, 1986, pp. 333-355.
5. Kamada,K., et.al.: Automatic Digitization of 1:25,000 Scale Topological Maps. Volume 29-4. ISPRS Commission IV, 1990. pp. 86-94
6. Li Deren, Gong Jianya. A Unified Data Structure To Represent Vector And Raster Data In GIS, International Archives of Photogrammetry and Remote Sensing. Volume XXIX, ISPRS, 1992, pp. 768-773.
7. Poul Frederiksen, et.al.: A Review of Current Trends in Terrain Modelling. ITC Journal, No. 2 ,1985, pp. 101-106.

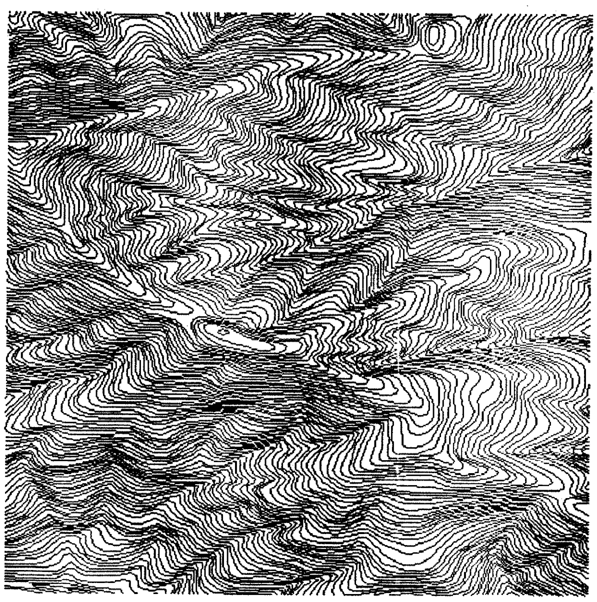

Fig. 1. Contour Map After Automatic Recognition

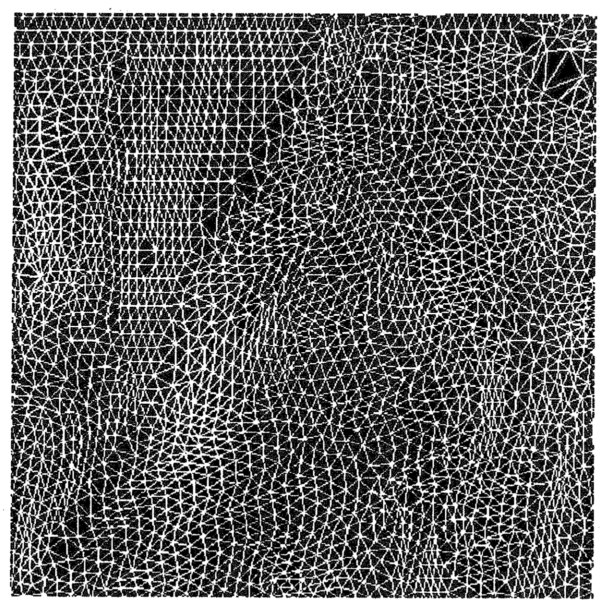

Fig. 2. Part Of Tin Corresponding To The Above Contour Map