RESEARCH OF FUNDAMENTAL THEORY AND TECHICAL APPROACHES TO AUTOMATING MAP GENERALIZATION

WU,Hehai
School of Resources and Environment Sciences of Wuhan University.
129 Luoyu road , Wuhan China. 430079
E-Mail: RHCHEN@PUBLIC.WH.HB.CN

ABSTRACT

What in the paper is mentioned concerns only the generalization of geo information which is stored in cartographic database or spatial database of GIS. Author's basic opinion is that in cartography and GIS there is lacking essential theory and effective algorithms which could be able to support the automatic generalization in digital environment by computer. In this paper the author proposed the (5W+1H) Conceptional Model of generalization("What is the essence of ?", "What is the object of ?", "Why ?", "When?", "Where?" and "How?"). The former 3W represent the scientific category of generalization, the latter (2W+1H) belongs to the category of technology. The conception model is implemented by the "Basic Structured Generalization Model" which consists of tree submodels: Global selection model(problem about "how many"); Structured Selection Model (problem about "what/which"). This is a map / database "configeration" model and Entity Handling Model which is entity composition model.

1. INTRODUCTION

Spatial data processing is a specific and important field of data handling realm. Its key pecularities lies in: mass volume of data, correlativity and complexity of geo information, variety of geo entities, regionality of distrbution and the necessity to handle geo information integratedly and so on. Map information is the main support component for spatial information handling and at the same time the dominant outputs of GIS processing results are represented in map form, especially in the form of thematical maps.

What in this paper is mentioned concerns only the generalization of geo information which stored in cartographic database or spatial database of GIS. Author's basic opinion is that in cartography and GIS there is lacking essential theory and effective algorithms which could be able to support the automatic generalization in digital environment by computer. Without true theory it will be impossible to develop correct technical methodology and to make it to have universal meaning.

2. BASIC THEORETICAL CONCEPTION MODELS

Author proposed the (5W+1H) **Conceptional Model** of generalization("What is the essence of ?", "What is the object of ?", "Why ?", "When?", "Where?" and "How?"). The former 3W represent the scientific category of generalization and the main part belongs to map /database design. The latter 2W+1H belongs to handling a given map feature in principle.

2.1 What is the essence of generalization ?

This problem of **"what is"** can be expounded by two main categories:

From the philosophical aspect to view, the author has an opinion that the term "generalization" is not owned by cartography and GIS themselves only, but it is a universal cognitive methodology by which human being can properly understand real world.. Therefore, from the macroscopic point of view, the carto / geo generalization should be considered as a particular situation of application of scientific cognition rules "abstraction / summarization " (the synonym of term "generalization"). That is to say the necessity of generalization will be forever because it is an epistemological category. This statement here mentioned can also answer the question **"why"** in philosophical aspect: to understand real world scientifically and properly.

From the technological aspect to view, author proposed a viewpoint of geo-info transformation as the procedure essence of generalization. In the digital environment, the geographical information has its equivalent DLM (Digilal Landscape Model) which describes the geo reality through graphics, attributes and relations among geo entities. DLM is geo-entity oriented, independent of concrete graphic representation. Therefore, it complies the common demands of multi kinds of users. The DLM view of generalization is equivalent to model generalization. Therefore, map generalization in digital environment became DLM tansformation or spatial database generalization: i. e. from DLM1 to DLM2.

At the same time the last statement has answered also the question **"what is to be done"**, namely the object of generalization: DLM or spatial database.

2.2 Why generalization is necessary ?

Beside the above mentioned answer of **"why"** and the traditional task to compiling maps in different scales, here we summarize the **why**s from other aspects. In the digital environment there are many new tasks which require geo info generalization: creating multi-resolutional spatial databases to serving different departments to support their planning, managing and decision making; compressing the spatial data volume to raise the speed and therefore to reduce the cost of data transfer. But in each case there is a key requirement: the receivers should be able to get maxmal information through the compressed spatial data subset. At last, for the data analyses in GIS itself the need of generalization is obvious, because direct using the full mass data set to perform spatial data analyses seems to be unreasonable or impossible.

2.3 **"When, Where & How "** can be incorporated into a higher category , namely the so called "context". Among them, the **"when"** can be understood as object own conditions to be evaluated from the viewpoint of generalization on the one hand. At the same time, "when" can also be understood as what or which proximity relationship inter geo entities is taken place on the other hand. "Where" can be interpreted as "context" localization. It can answer "where what kind context is happend". "How" involves very abundant connotation. On the basis of the determining actual situation of the former 5W by which the corresponding "How" can be established.

The (5W+1H) scheme builds up the basic conceptional / theoretical model.

For implementing the model above mentioned here the author proposed a general structurized generalization model.

3. CREATING THE GENERAL STRUCTURED GENERALIZATION MODEL

On the basis of above mentioned the author proposed a general structurized model of generalization. It consists of tree submodels:

3.1 Global Conception Model

Generally speaking, a map or database to be designed may be consist of only a subset of geo objects from the original map or database. Therefore first of all the determination of "which feature classes" and "how many" objects should be carried out. This is a map / database "conception" model. The "how many" problem can be considerd as a problem about the total map load or database capacity.

In the case of traditional map compilation, the solution of this problem is realized empirically. In the digital environment, it is possible to solve the sub-problem (graphic element) automatically through program evaluation of digital map or spatial database.

Because fractal theory especially the extended fractal approach can derive the relationship between the observation measures and the observed results, the derived relationship can be used to determine the "full view" of new map or new database(fig. 1).

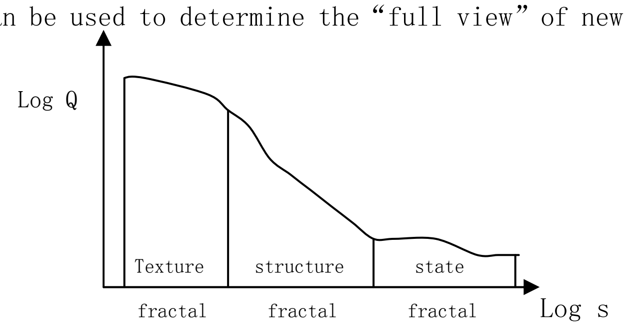


Fig 1 The Relationship Between The Observation Measures And The Observed Results.

Here S is the measures of observation, Q is the observed results.

Using the fractal method mentioned above we can determine approximately the required capacity Q_i of a new map to be designed if the map measure S_i (scale) is given. Such a approach possesses an adaptive property, i.e. the required parameters are derived from map or database itself. Here we use the symbols Q_o and Q_i to represent the map loads of the source map and the current one respectively and we can obtain the coefficient of map load reduction:

$$K_i=Q_i/Q_o$$

The evaluated total map load Q_i should be distributed according to the given feature classes with the following automated ditribution approach proposed by the author (fig 2).

So far we already obtained map load for each feature class. The next step is to create structured model determining which or what objects should be selected according to the distributed map load for each feature class.

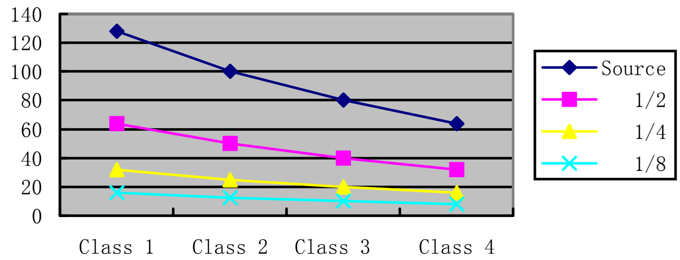


Fig.2 Distribution Of The New Map Load For Each Feature Class

3.2 Structured Evaluation Model

The structured evaluation model has a task to evaluate geo entity set structurely beside the traditional evaluation techniques. Here the author has an opinion to carry out two different kinds of spatial evaluation: global structure evaluation and local importance estimation.

Such a structured evaluation model will be able to determine "what/which" objects should be selected properly. This is a map / database "**configeration**" model. There are two typical examples to demonstrate just mentioned model:

(1) Structured Generalization of Grouped Point Features

For the grouped and scattered point features the maim characteristic is their spatial distribution pecularity. Point set based multi-layer embedded convex hulls is an excellent means to repesent both the global spatial distribution structure and the inner and local characteristics (fig. 3).

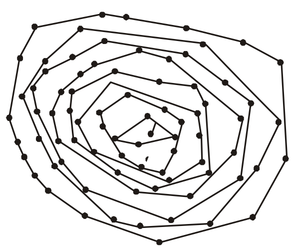


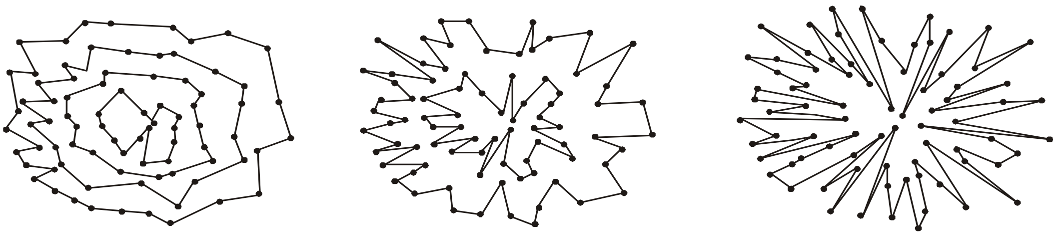
Fig.3 The Struture Description of Grouped Point Features

The further generalization steps are as follows:

Simplifying the distribution structure, i.e. reducing the number of embedded convex hulls though merging the several neighbouring convex hulls into general non-intersected each other polygons (fig.4).

So far the grouped point features are structured very well. Here we completed a transformation: transferring the area form object (point group) into a set of line objects (polygons). This transformation provides a possibility to use reasonable line generalization method to perform point set generalization.

Point set based Voronoi graph can give an auxilary information to help better evaluate the point object: the bigger the area owned by a point object, the importanter the corresponding point object.



a. Into 4 polygons b. Into 2 polygons. c. Into 1 polygon

Fig.4 Merging the neighbouring convex hulls into embedded polygons

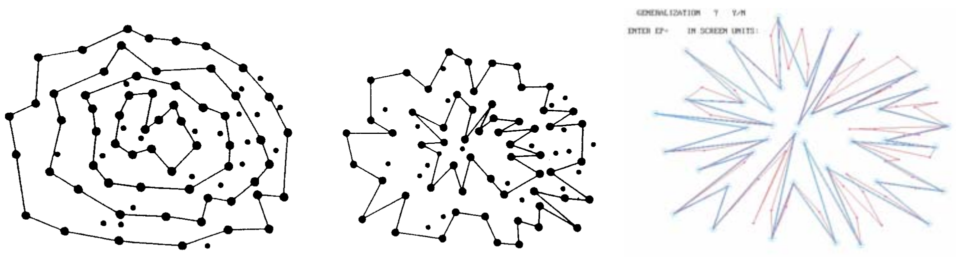


Fig.5 Different Levels of Generalization.

(2) Structured Selection of Rivers

To evaluate the river importance it is necessary to create the tree structure of river network firstly. Tree structure provides hierarchy - rank information. That is very usefull for river selection because in the feature selection process the key attention should be paid to the leaf nodes of tree structure. But this is not to say that all rivers being located at the tree leaf nodes will have the same importance. To further differenciate the leaf node rivers requires an auxiliary method which can distinguish the more importanter. That auxiliary method is the line feature Voronoi graph. Bigger values of Voronoi polygons apearce in the following three cases: the river is longer or the distance between the neighbouring rivers is bigger or the river locates in the periphery of a drainage area. In these cases the river has a priority to be selected.

These two submodels are feature class oriented. The following third submodel will handle information of an entity itself.

3.3 Entity Handling Model

This is an object “composition” model. For a geographic entity there are 3 kinds of information (semantic, metric and relational information) should be handled properly:

Attribute generalizstion;

Graphic generalization;

Relational Information handling explicitly or implicitly.

4. CONCEPTION FRAME ABOUT BASIC APPROACHES OF GRAPHIC GENERALIZATION

4.1 Generalization In Space Domain

The most generalization methods belong to the space domain. Here the main pecularity

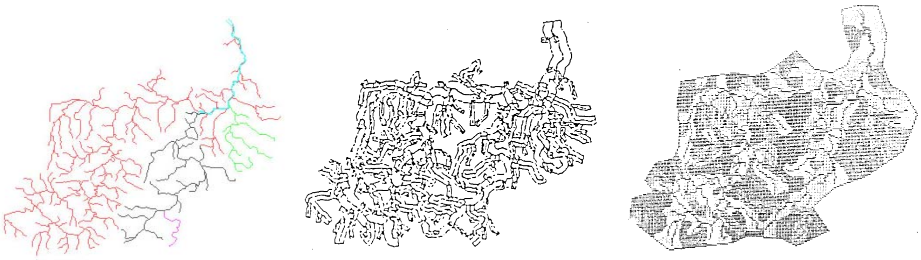


Fig 5 River System ,River’s Buffer Zones and River’s Voronoi Graph

lies in immediate handling geo information entity by entity.

(1) Entropy Approach Based Generalization

This approach of generalization is the best one in theory, but is the diffcultest one in practise (W. Weber,1982).

(2) Filtering Approach Based Generalization

A. Line Data Filtering

This approach gives global smoothed information. Therefore, the local but very characteristical details, for example capes, fiords etc. ,will be smoothed out.

B. DEM Data Filtering

Here the situation almost is the same as the Line Data Filtering mentioned above. The mountain peaks will be flattened out, and the depressions will be filled up.

(3) Heuristic Approach Based Generalization

The main conception of heuristic approaches has a basic proposition that for the creative activity of human being there is not existing general methods. But it is possible to provide several rules or schemata. Although these measures can not guarantee to achieve the goal but can raise the success possibility. The most realization means of heuristic approach is simulating the traditional generalization procedures. The most recent generalization methods belong to this category.

A. Point Set Generalization

Here the key problem is to selecting the point object structurely in global and with considering the proximal importance of individual point object in local.

B. Line Feature Set Generalization

Line feature generalization is studied very much, because the main component of map content is line objects.

a. Natural Line Feature Generalization

For raising the rationality of generalization it is necessary to perform a automatic segmentation of nature line features (e.g. rivers, shore lines) , using qusi optimized statitstic division method. This process gives homogeneous segments. The line homogeneity can ensure the effectness of line generalization algorithms.

b. Man-made Line Feature Generalization

The key chatacteristic of man-made line objects is they have less complexity than natural ones.

C. Area Feature Generalization

There are two types of area features: scattered areas and connected polygons. Here is studied only the city planimetric graphics generalization which involves two main sub-processes:

a. Affine transformation of city block information

Because the representation of streets on the map is enlarged generally, the affine transformation of all objects in the city block will be inevitable (fig.6).

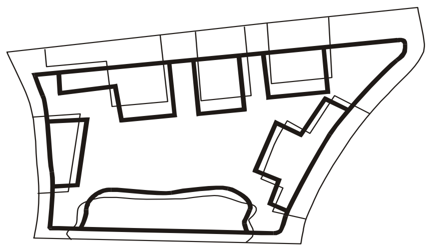


Fig 6. Principle of Affine Transformation

b. Building merging

The author proposed and implemented a building merging algorithm based on the "attractive force direction" principle. There are 6 quantified levels of building merging for 2 closed buildings which have irregular steel beam form(fig 7).

D. Reliefform Generalization

There are two categories of landform generalization: structured approach (reliefform ridges , valleys oriented) and global filtering approach. Their pecularities have been explained as before.

(4) Fractal Approach Based Generalization

Using the fractal geometry a very essential parameter - fractal dimension value of an object itself or of object set - can be obtained. This parameter describes variation rate with the change of scale of observations in preserving self- similariry. Because fractal dimension value describes structure aspects both for individual object and for object set, therefore it can be used in structured generalization.

For point, line and area feature generalization, the fractal dimension value can be estimated by coarsening the view units, e.g. increasing the grid size to count the non-empty cells. Because fractal theory especially the extended fractal approach can automatically provide an adaptive solution , based on this theory a more structured generalization principle can be established.

4.2 Generalization In Frequency Spectrum Domain

Some times, the essence of matters is difficult to be discovered in space realm. But the same thing would be easy to realize if this thing were equivalently transformed into frequency - spectrum realm. The representation of this information handling is the Fourier transformation or its new derivations : wavelet analysis.

5. ACKNOWLEDGEMENT

The author acknowledges the support of National Natural Scince Foundation of China for the research.

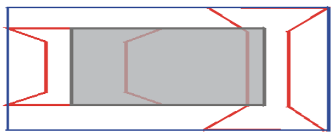


Fig 7a The Strongest Level

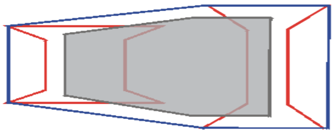


Fig 7b The Stronger Level

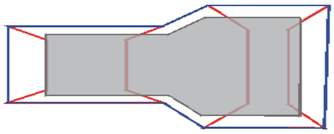


Fig 7c The Normal Level

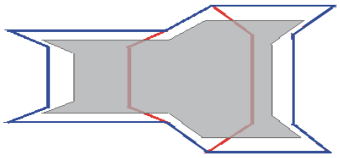


Fig7d Strong Merging of Internal Sides

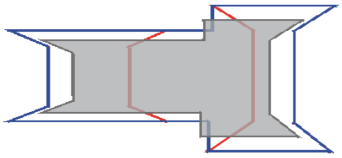


Fig 7e Midle Merging of Internal Sides

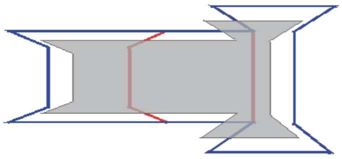


Fig 7f Weak Merging of Internal Sides

Fig 7. Building Merging By Using The Attractive Force Direction

6. REFERENCES

Bader M, Weiber R. 1997, Detection and Revolving Size and Proximity Conflicts in The Generalizaion of Polygonal Maps. Proceedings of the 18th ICA/ICC Vol.3, pp.1525-1532

Buttenfield B P, McMaster R B. 1991, Map Generalization. John Wiley & Sons, New York

Muller J C, J P Lagrange and R Weibel.1995, GIS and GENERALIZATION. GISDATA 1 .

Peng Wanning. 1997, Automated Generalization in GIS. ITC Publication Series.

Weber W. 1982, Automationsgestuetzte Generalisierung. Nachrichten aus dem Karten- und Vermessungswesen, Heft Nr. 88. pp.77-109

Wu Hehai.1981, Prinzip und Methode der automatischen Generalisierung der Reliefformen. Nachrichten aus dem Karten- und Vermessungswesen, 1981.Nr.85 pp.163-174

Wu Hehai. 1995a, The automated construction of tree structure of river network. Journal of Wuhan technical university of surveying and mapping, Vol. 20 Suppl.pp.7-14 (in chinese).

Wu Hehai. 1995b, The construction of tree structure of contoures. Journal of Wuhan technical university of surveying and mapping, Vol.20 Suppl.pp.15-19(in chinese).

Wu Hehai. 1997a, Problem of buffer zone construction in GIS. Journal of Wuhan technical university of surveying and mapping, Vol.22, No.4, pp.358-366 (in chinese).

Wu Hehai. 1997b, Principle of convex hull and its application in generalization. Engineering of surveying and mapping. Vol.6, No. 1 pp.1-6 (in chinese)

Wu Hehai. 1997c, Structured approach to implementing automatic cartographic generalization. Proceedings of the 18th ICA/ICC, Stockholm, Sweden, 1997, Vol.1, pp.349-356

Wu Hehai. 1998, Numerical examination for fractal extension. Journal of Wuhan technical university of surveying and mapping, Vol.23, No, 4, pp.329-336 (in chinese).

Wu Hehai. 2000, Research on fundamental theory and technical approaches of automated map generalization. Journal of Wuhan technical university of surveying and mappint, Vol.25, No.5, pp.377-386 (in chinese).