ACHIEVEMENT AND ISSUES ON THE DESIGN OF DIGITAL MAP GENERALIZATION OPERATORS

Dan Lee
Intergraph Corporation
2051 Mercator Dr., Reston, VA, USA 22091
Phone: (703)264-2330; Fax: (703)264-5620
Email: dlee@ingr.com

Abstract

One of the core components of a digital map generalization system is a set of operators designed to automatically reduce the data complexity for selected features. These operators are not simply tools for altering feature geometry; they should create results that reflect the art and science of cartographic generalization. Scattered specifications for map generalization exist in mapping organizations around the world, but none of them is universally acceptable due to the cultural differences among locations and countries. However, successful map generalization can be guided by one principle --- to ensure the most representative depiction of the geographical area and the cartographic quality for the target output.

This paper discusses cartographic controls that are imbedded in existing generalization operators and describes experience using these operators that are successfully incorporated in a computer-assisted generalization system, MGE Map Generalizer (MGMG), produced by Intergraph. The major focus is how to use the given digital information to identify and produce the most representative shapes and distributions of map features; and how to ensure output data integrity, such as feature connectivity and geographical relationships. Illustrations will be given to show some achievements with several application sources. Issues and remaining questions are addressed as a call for future research and development.

1 Meeting cartographic specifications

The commercial product for computer-assisted map generalization -- MGE Map Generalizer (MGMG) produced by Intergraph Corporation -- has drawn many map makers' and researchers' interest. A question is often asked: What makes the generalization operators in MGMG successfully meet cartographic specifications? The following discussion explains.

A cartographic representation of the reality -- a map -- is usually associated with a plotting scale. At this scale, features are mapped with a minimum size, minimum length, minimum spacing, and minimum density. Only certain feature classes and certain level of information detail are considered significant for inclusion in the map. The operators designed and implemented in MGMG reflect these constraints in many ways. A full description of the MGMG product and functions can be found in an early paper [1].

1.1 Flexible feature selection for inclusion and process

In the initial data extraction from a master database, feature selection determines what feature classes will be shown in the final map and be further processed in generalization. MGMG allows user to select features either by their graphic definitions stored in a plotting feature table or by their geographic attributions in the database. Once the features are specified, they can be extracted from the master database automatically. During generalization, some features need to be processed differently from others using different parameters or operators; therefore, flexible feature selection plays an important role. MGMG provides these fundamental capabilities to meet cartographers needs. For example, one can select county boundaries, major roads, and buildings for Map Content Extraction from a master database; then further select buildings that are smaller than certain size to be Collapsed to point features.

1.2 Parameter setup in relation to map scale

A map delivers geographic information to readers by showing a relevant amount of feature detail with plottable and readable symbology at a given scale. Using scale-related constraints has been a major consideration in MGMG operators design. Many of the MGMG operators involve parameters such as cluster tolerance, minimum distance, and minimum size to ensure the cartographic quality of generalization outcome. A few examples are given below to illustrate how such parameters are used to produce satisfactory results.

In Disjoint Area Aggregation, the two major parameters are:

Threshold Tolerance -- determines if areas are close enough to be aggregated;

Zone Tolerance -- ensures the connecting region (zone) is wide enough.

Figure 1 shows the Disjoint Area Aggregation using the Non-orthogonal algorithm. The above two parameters were applied to combine areal features within a Threshold distance. Data source: USGS DLG data, Herndon, Virginia.

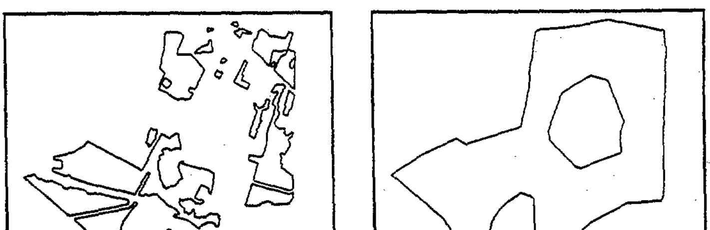
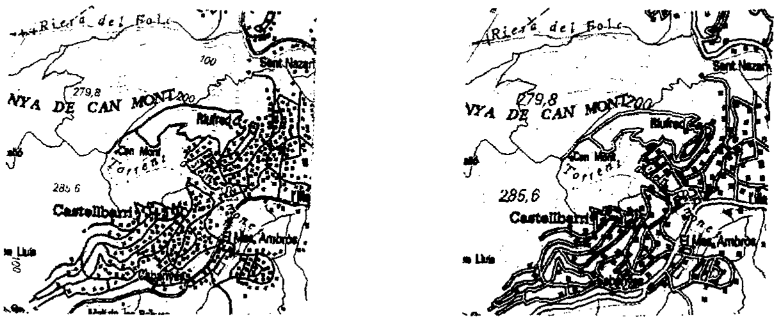


Figure 1: Area Aggregation Left = before; Right = after

The example in Figure 2 was taken from the MGMG evaluation report by Institut Cartografic de Catalunya in Barcelona [2, pp. 21]. Building density at 1:50,000 scale was reduced for a target scale of 1:100,000 using the Point Set to Point Set algorithm in Point Typification operator in MGMG. The result was plotted at 1:50,000 scale for comparison purpose. The *Cluster Tolerance* parameter is used to retain the maximum number of points while ensuring that no two points are within the specified tolerance.



Original data set at 1:50.000 **Automatic generalization at 1:50.000**

Figure 2: Point Typification

Using the Tree Leveling algorithm in Line Typification in MGMG, the complexity of a hydrographical network can be easily reduced. The use of the Maximum Retention Level and the Minimum Length Tolerance eliminates river branches by level and by length. Figure 3, provided by Institut Cartografic de Catalunya in Barcelona [3, pp. 14] illustrates the result of Line Typification.

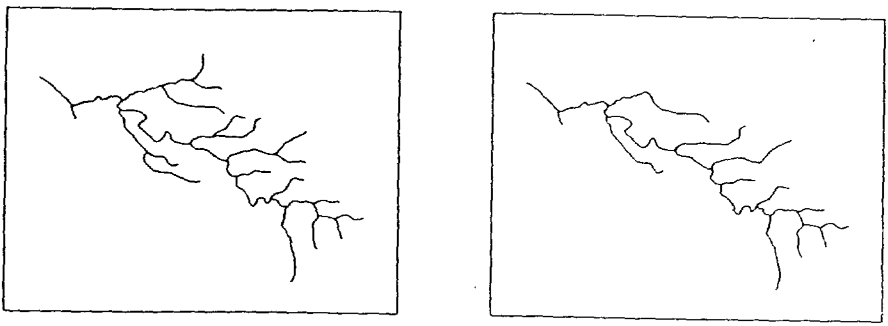


Figure 3: Line Typification Left = before; Right = after

More illustrations can be found in a previous paper [4, pp.327-334], such as the Minimum Length and Minimum Area Tolerances in Area Clarification -- a simplification algorithm, the Minimum Spur Length and Minimum Area in Area-to-Line Collapse, and so on. These parameters are directly related to the output scale and are usually fixed values. It is very important to understand the role of these parameters in order to properly use MGMG operators.

2 Ensuring data integrity

During generalization, features change their appearance and positions, but the data integrity should not be destroyed and feature relationships should be maintained. For example, without considering feature relationship, Area-to-Line Collapse operator in MGMG generates a centerline from an areal feature along the longest path. With the option of Preserve Feature Connectivity in Area-to-Line Collapse operator, if an area is connected to a linear feature, the connection will be retained, although the new linear representation of the area may not necessarily be the longest path found. This applies to the case where a river enters a lake and runs out of the lake. If the areal feature that represents the lake is totally collapsed, the resulting line should be connected to the river on both ends.

Hole Retention option in some MGMG operators is another way of ensuring data integrity. As shown in Figure 1, there seems to be an open space in the middle of the original areal feature cluster. By using a Minimum Hole Size parameter, the algorithm automatically creates the hole if its size is greater than the specified size. A similar parameter is also used in Point Aggregation.

In GIS and thematic mapping, generalization is not only altering feature graphics, but also transforming geographic attributes. Some of the geographic attributions are derivable from the source database. For example, when areal features are aggregated into new areas or collapsed to point representations, the associated statistical information should be updated to reflect the new features. Currently, MGMG offers several methods to derive geographic attributions. These methods include summation, maximum value, minimum value, mean, most prominent, least prominent, etc. This ensures the input and output data consistency and information integrity.

3 Issues for Further Research and Resolutions

There are three main areas that need more research and resolutions. These areas are: further automation of generalization process; analytical tools to analyze and evaluate input and generalized data; formalization of generalization processes.

3.1 Further automation of generalization processes

To pursue further automation of generalization processes, there are two major focuses: one is to derive rules and parameters that guide generalization operations; another is to develop new algorithms that follow the rules and handle more complicated cases.

An interesting and practical requirement came from one of our customers, who produces forest maps at various scales and faces a problem of generalizing switch-back roads. It is not a linear feature simplification problem, but rather a linear pattern typification problem. A resolution can not be developed until rules are made explicit.

Another typical generalization issue is feature displacement which requires a full set of regulations to guide the process. When features are in conflicts, what features should be moved away from others? Or in what case all the features should be evenly moved away from a point? When a part of a line needs to be displaced, over what distance the displaced line should be gradually merged and connected to the rest of the line on both ends? Defining these types of rules is very essential to automation of generalization processes.

3.2 Analysis and Evaluation of data

It is easy for a cartographer to visually identify areas and features that need to be generalized. However, there is still a lack of powerful means to evaluate input data and to measure generalized results. Some MGMG operators have imbedded analytical ways of identifying feature clusters, feature density, and feature significance. Other methods of data analysis have been attempted by a number of researchers, such as describing lines by structure signatures [5] and characterizing lines by segmentation analysis [6]. An overall judgment of generalization quality still remains a research subject.

3.3 Formalization of generalization processes

Generalization is probably the most complicated process in cartography. Although the subjective nature of map generalization makes the formalization of the process difficult, certain principles can be followed to build application-oriented workflows. Some processes are feature class dependent (e.g. hydrographical and man-made); others are feature dimension dependent (point, line, and area). Logical operational workflow should be formalized to minimize the reduction of spatial accuracy and to maximize process efficiency. The interactive generalization system such as MGMG provides us a helpful learning environment. Through case by case studies, some initial operational sequences have been examined. Further formalization of the processes will still be a major target in digital map generalization.

4 Conclusions

A significant achievement of designing and implementing cartographic generalization operators in an interactive map generalization system, MGMG, has been made. Great challenge is still ahead. Our future effort will focus on the unresolved issues and make the system more efficient in map production.

References

[1] Dan Lee, 1992, Cartographic Generalization, Intergraph internal technical paper.

[2] Blanca Baella, Josep Ll. Colomer, Maria Pla, May 1994. Map Generalizer: A Practical Experience. Unpublished report for Intergraph Corporation internal use.

[3] Blanca Baella, Josep Ll. Colomer, Maria Pla, September 1994. Experiences on Map Generalizer at the ICC. Presentation report at OEEPE WG on Generalization Meeting, Edinburgh.

[4] Dan Lee, 1995. Area Features in Digital Map Generalization. ASPRS/ACSM Annual Convention & Exposition, Technical Papers, Vol.1, pp. 327-334, Charlotte, NC.

[5] Barbara P. Buttenfield, 1991, A Rule for Describing Line Feature Geometry, *Map Generalization: Making Rules for Knowledge Representation*, edited by Buttenfield and McMaster, Longman Scientific & technical, pp.172-186.

[6] Corinne Plazanet, 1995, Measurement, Characterization and Classification for Automated Line Feature Generalization, ASPRS/ACSM Annual Convention & Exposition. Technical Papers, Vol.4, pp. 59-68, Charlotte, NC.