AUTOMATIC GENERALIZATION OF ROADS AND BUILDINGS

Pingtao Wang, Takeshi Doihara

Asia Air Survey Co. Ltd.
Kanagawa 215-0004, Japan
{pt.wang, t.doihara}@ajiko.co.jp

Commission VI, WG IV/3

KEY WORDS: Generalization, Automation, Centerline, Buildings, Block, Road, Cartography.

ABSTRACT:

Map generalization simplifies the details of map representation. Automatic generalization has been a hot research topic for decades, but there does not exist a set of universal rules or algorithms that explicitly defines how generalization should be performed. This paper presents a method to automatically generalize roads and buildings. With the proposed method, road generalization and building generalization are carried out consecutively. Road generalization includes Road Modeler, which converts original road edges to road polygons, and Network Generator, which collapses road polygons to road networks. The created network is topologically connected and suitable for GIS (Geographical Information System), such as car navigation systems. Building generalization is mainly composed of clustering building polygons to building clusters, aggregating a building cluster to a polygon, and simplifying both original and aggregated building polygons. Using the created road networks as the constraints for generalizing buildings leads to the generalized results without contradiction. Some experiments have also been implemented to verify the effectiveness of the proposed method.

1. INTRODUCTION

A map is a resolution-dependant geographical representation of the real world. Map generalization is a complicated process and usually involves a great deal of spatial analysis to decide what and how to generalize, and how to resolve conflicts that might occur during the process. Manual generalization is a time-consuming and skilled work. Cartographers draw a reduced map by hand, and eliminated some unimportant features to simplify lines, to combine adjacent areas, and to resolve conflicts as their judgment (ESRI, 1996). That is to say, manual generalization leads to the inconsistent results because of the difference of cartographer's experiences. Therefore, automated map generalization is desirable and has been researched for decades.

Nowadays, with the development of computer science, some manual processes of map generalization are being implemented on some GIS software. There are many researches about automation of generalization operators such as simplification, displacement, collapse, aggregation, typification, and so on. Lines are major features in a map, and most efforts about automatic generalization have been made in line's generalization or simplification. Furthermore, roads and buildings, which are usually represented by lines, are the most basic objects in a digital map. Therefore, we shall pay special attention to the automatic generalization of road and building data in this paper.

Road generalization is the process of creating and/or updating the road network of a small-scale map from the corresponding road edges of a large-scale map. Line simplification, which may be the earliest attempt for automating generalization (Douglas, 1973), has been widely used to simplify road networks (lines) and other linear features. Kreveld and Peschier presented an approach to generalize road networks by keeping three objectives in mind: not allowing roads to be too close, avoiding detours between important points, and giving priority to bigger roads (Kreveld 1998). Annita and others tried to collapse polygonal road area to road network by using triangulation (Annita 1998). There are also many other papers about the processing of road centerlines or generalization of polygonal roads, but rare researches are found for creating road network from road edges or double-line roads. Considering the fact that roads are widely represented by double-lines, rather than polygons, in large-scale maps for our experiments, further researches are necessary to convert road edges to road networks directly. Here, double-line roads may be the curb lines or the boundaries of the corresponding road area in a large-scale map.

Building generalization involves the simplification of independent building polygons, the aggregation of building clusters and the displacement between the generalized buildings and other features such as roads. For building cluster aggregation, Regnauld developed a method to detect building pattern groups by applying the minimum spanning tree (MST) model from graph theory (Regnauld, 1996). Anders and Sester applied hierarchical cluster algorithm to typify buildings and lakes (Anders, 2000). From the legibility of the entire generalized map, some other objects, such as roads, should also be considered in building generalization.

In this paper, a new framework is proposed to generalize road and building data. The present framework implements road generalization and building generalization consecutively. Road generalization includes Road Modeler, which converts original road edges to road polygons, and Network Generator, which collapses road polygons to road networks. The created network is topologically connected and suitable for GIS (Geographical Information System), such as car navigation systems. Building generalization is mainly composed of clustering building polygons to building clusters, aggregating a building cluster to a polygon, and simplifying both original and aggregated building polygons. Displacement of building polygons is also considered in clustering procedure by using the created road networks to separate building polygons.

In the following sections, we give an overview of the proposed framework firstly and discuss the details of some algorithms in the framework in section 2. Section 3 gives some experimental results to verify the effectiveness of the proposed method. And section 4 concludes the paper with some discussions.

2. FRAMEWORK AND ALGORITHM

2.1 Notations and Definitions

Some notations used in this paper are listed below.

P_{obj} :	Point set representing obj
L_{obj} :	Line set representing obj
A_{obj} :	Area (polygon) set representing obj
N_{obj} :	Network (topologically connected lines) set representing obj
T_{obj} :	Triangle set representing obj
$TIN(X)$:	Function creating TIN (Triangulated Irregular Network) from X
T :	$=\{T_i \mid i = 1,2,\cdots N\}$, Triangle set

In this paper, we only discuss the generalization of road edges and building polygons. Let L_{road} be original road edges, and $A_{building}$ original building polygons. Then, the original map dataset may be represented as follows.

$$\{L_{road}, A_{building}\} \tag{1}$$

Furthermore, a notation without an apostrophe (') means an original datum or object, such as road edges L_{road} mentioned above, otherwise, a created/output datum or object.

For the convenience of the following explanation, we also clarify some key words as follows.

Road edge:	Line representing the right or left side / boundary of the original road data.
Building polygon:	Polygons representing the original building data.
Road region:	The range between two intersections, excluding any intersections and inside a road.
Intersection:	Range connected by more than two road regions.
Road polygon:	A single road region or connected road regions, the first output dataset A'_{road} in this paper. Road polygons are also connected with each other to cover the entire road.
Centerline:	Line representing the central skeleton of a road polygon.
Road network:	Topologically connected centerlines, the second output dataset N'_{road} in this paper. As compared to road centerlines, road network emphasizes the topological characteristics of the road data.
Aggregated building:	Polygon merged by more than one building polygons.
Generalized building:	The third output dataset, $A'_{building}$. A generalized building is a polygon obtained by simplifying an aggregated building or an original building polygon, which was not aggregated.

2.2 Framework

The proposed framework is shown in Figure 1. Here, road generalization includes two units: Road Modeler and Network Generator (Wang, 2002, Wang, 2003). The input data is a set of road edges L_{road} and building polygons $A_{building}$, and the output includes three datasets: road polygons A'_{road} (result of Road Modeler), road network N'_{road} (result of Network Generator), and generalized buildings $A'_{building}$ (result of Building Generalization). Here, Road polygons is used to create road network inside the Road Generalization unit, and can also be outputted as modeled road data for some GIS systems too. Road network is also used as constraints for generalizing buildings in Building Generalization unit. The framework can also be represented as follows.

$$A'_{road} = \{T_i = TIN(L_{road}) \mid \mathrm{T_i} \text{ is inside the road area}\} \tag{2}$$
$$N'_{road} = \text{topologically connected centerlines of } A'_{road} \tag{3}$$
$$A'_{building} = \text{aggregated } A'_{building} \text{ with constraints of } N'_{road} \tag{4}$$

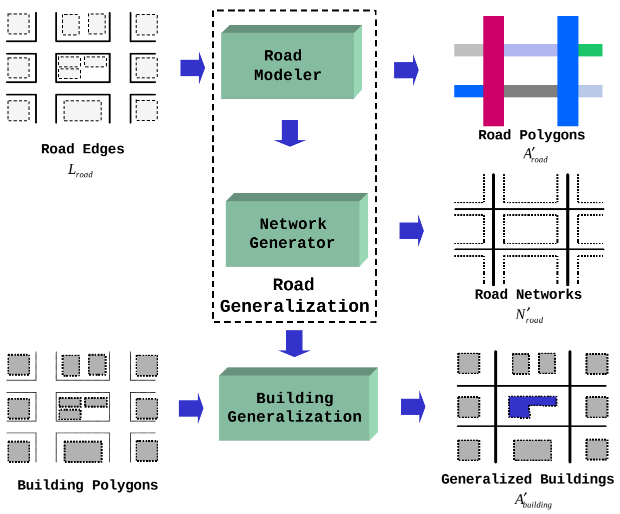


Figure 1. Framework for Generalizing Roads and Buildings

The proposed method works as follows. Firstly, the original road edges L_{road} are modeled as road polygons A'_{road} by Road Modeler, as shown in equation (2). Here, a TIN is created by using road edges as constraints and analysed to merge the triangles inside the road range to generate road polygons.

Secondly, road network N'_{road} is generated from the road polygons by Network Generator, as equation (3). To generate road networks, a centreline, the skeleton connected two intersections of a road polygon, is calculated for each road polygon, and all centrelines are topologically connected to generate road network consecutively. As the attribute of each centreline, the width of the corresponding road polygon is also calculated in Network Generator. At last, Building Generalization works to cluster and aggregate the original building polygons and to simplify both aggregated buildings and left original building polygons to form generated building $A'_{building}$. In this unit, the road network created above is used as the constraints to separate the original buildings.

Some details of the proposed framework are discussed with the corresponding processed results in the following subsections.

2.3 To Detect Intersections

In our experiments, an intersection is actually defined by one or more connected triangles, where three edges of every triangle are not original road edges. Here, The triangles are parts of the created TIN from the original road edges in Road Modeler. Figure 2 shows some detected intersections from a set of real road edges. Here, the regions between each pair of intersections are road regions.

Figure 2. Samples of Detected Intersections
Intersections = filled regions,
road regions = ranges between each pair of intersections

2.4 To Create Road Polygons by Connecting Road Areas

Generally, there are three or more road regions around an intersection. To connect two principal road regions, probability connection function between two regions is introduced as follow.

$$
\begin{aligned}
PCF(A_i, A_j) \; = \; & a\,\frac{\left\| D_i - D_j \right| - 90 \right| \bmod 90}{90} \\
+ \; & b\left(1 - \frac{\left|W_i - W_j\right|}{W_i + W_j}\right) \\
+ \; & g\,\frac{2S(A_i \cap A_j)}{S(A_i) + S(A_j)}
\end{aligned}
\qquad (5)
$$

where, $D_i = [-90{,}90)$, direction of A_i
W_i = width of A_i
$S(A)$ = area of A
$0 \le a, b, g \le 1$ and $a + b + g = 1$

a, b and g are weights for direction, width and **overlapped area** between two road regions to be checked, respectively. Here, the width and direction of the rectangle are equal to those of the corresponding road region, and the length is an external parameter. Moreover, a boundary rectangle is created for each road region at the borderline between the intersection and the corresponding road region. Then, the overlapped area between two road regions means their common region. Figure 3 shows some overlapped boundary rectangles. With equation (5), the pairs of road regions to be connected are obtained and shown in figure 4.

Figure 3. Overlapped Boundary Rectangles

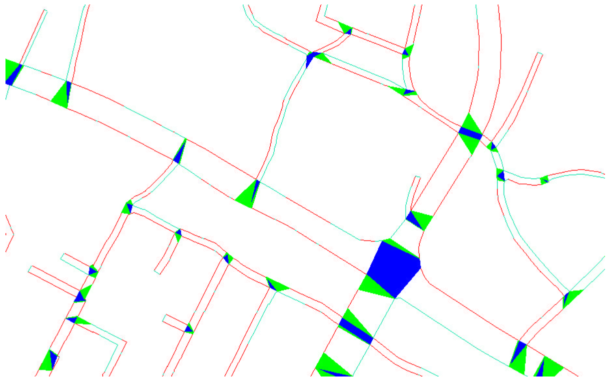


Figure 4. Pairs of Road Areas: There are two green triangles for each intersection. Two road regions including green triangles will be merged as a road polygon.

2.5 To Create Road Network

Two steps are introduced to create road network from road polygons. One is creating centerlines for all road polygons, and the other is extending the centerlines to generate road networks. Both steps are implemented in Network Generator. A centerline is created by connecting the middle points of inner edges of all triangles one by one for a road polygon. Here, the created centerlines are not connected with each other, because each centerline is limited to the inner range of the corresponding road polygon. Road network is created by extending each centerline to the centerlines of connected road polygons. Figure 5 shows the created road networks overlapped with intersections and road polygons, which are represented by triangles.

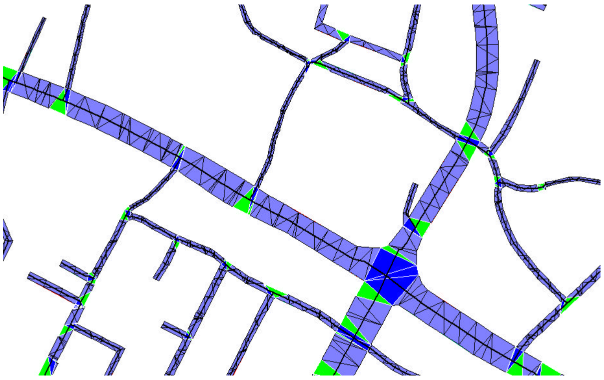


Figure 5. Road Networks

2.6 To Generate Building Polygons

To generalize building polygons, there are three steps: clustering building polygons, aggregating building clusters, and simplification, in the proposed method.

Clustering building polygons is implemented to decide which buildings are near enough to be aggregated. Figure 6 depicts the clustering procedure. Here, to avoid erroneous clustering, road networks should be used as constraints for separating buildings compulsorily. The details are as follows.

1. Defining links between two different buildings. A link's length is equal to the minimum distance of two linked buildings (Figure 6). Here, each link begins at a building and ends at another one. The numbers near each link is the minimum distance between linked two buildings.
2. Sorting all links by their lengths in ascending order. A sorted result is also shown on the left-down corner of Figure 6.
3. Grouping links from the first (shortest) link. If one of two linked buildings is included in the current group, another building will be appended to the current group. Otherwise, a new group is created with the current link's two buildings. A new group is also created if a link's length is larger than the parameter for clustering.
4. Repeating the grouping until all buildings are grouped.

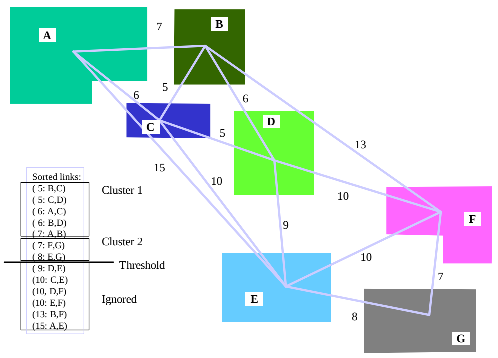


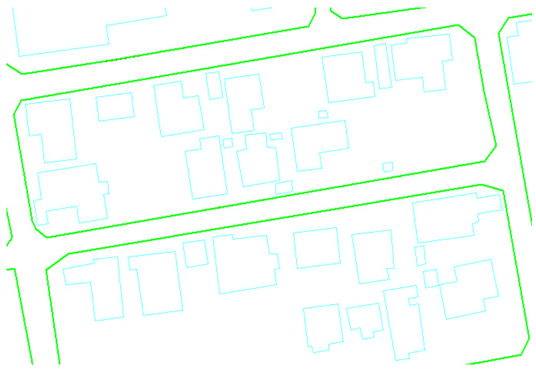
Figure 6. Link's network for clustering buildings

With the above procedures, we can obtain the clustered results of Figure 6 with different thresholds for clustering buildings, shown in Table 1. The left-down corner of Figure 6 is same with the greyed areas in Table 1.

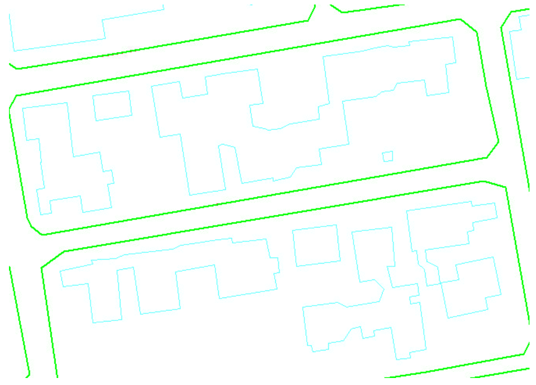
Cell size	Cluster ID	Buildings
5	1	B, C, D
	2	A
	3	F
	4	G
	5	E
6	1	A, B, C, D
	2	F
	3	G
	4	E
7	1	A, B, C, D
	2	F, G
	3	E
8	1	A, B, C, D
	2	E, F, G
9	1	A, B, C, D, E, F, G

Table 1. Results of clustering buildings with different thresholds for Figure 6

Aggregating a building cluster is carried out by using TIN and creates aggregated buildings from building clusters. Some aggregated buildings are shown in Figure 7. Here, the buildings, which are too near to road edges, were also displaced from the road edges before clustering building polygons.



(a) Original building polygons



(b) Displaced and aggregated buildings

Figure 7. Displacing and aggregating buildings

Simplification of buildings includes deleting too small building polygons and eliminating short segments and unimportant vertexes of a building polygon. Figure 8 shows the simplified buildings of Figure 7(b).

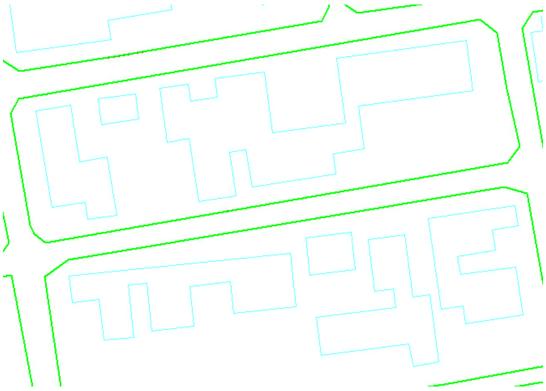


Figure 8. Simplifying buildings

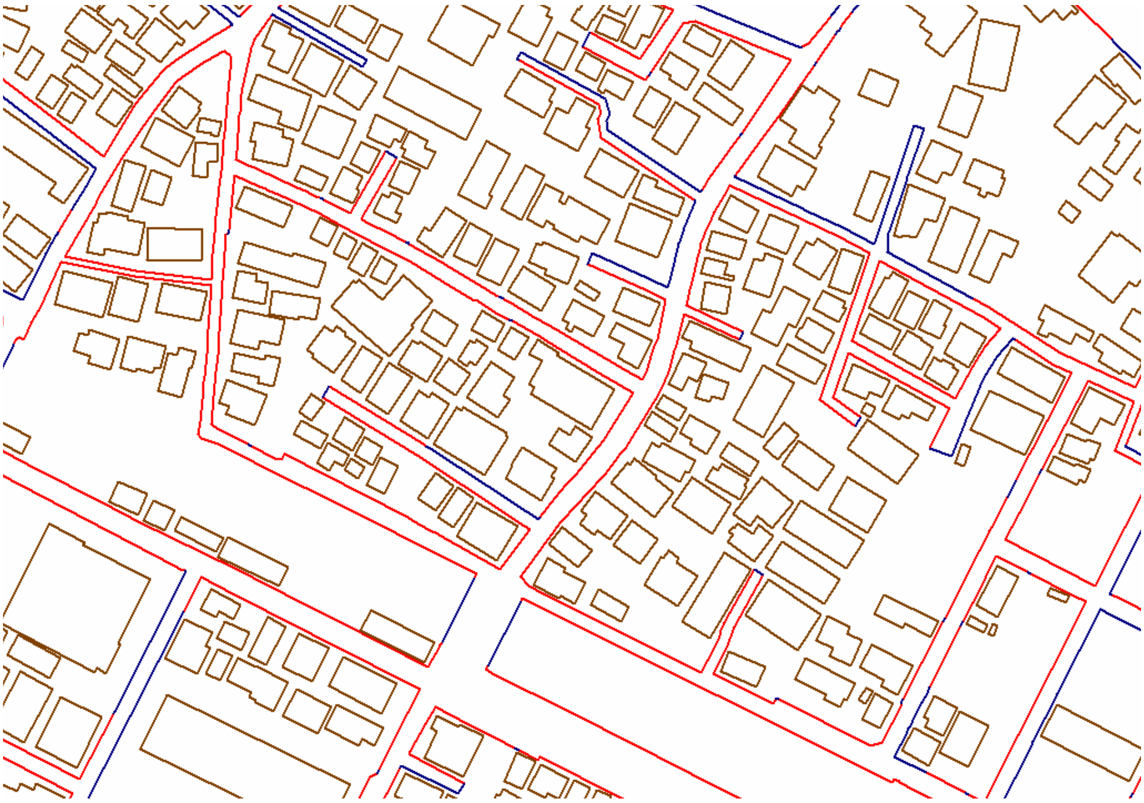
3. EXPERIMENTSAL RESULTS

We implemented some experiments to confirm the principle and the effectiveness of the proposed framework. Figure 9(a) shows the original dataset, including road edges and building polygons. The generalized result is shown in figure 9(b). Here, the modeled road polygons are filled with different colors, and the road networks are represented by black lines inside all road polygons. Also, all near building polygons are aggregated to a single one. And, both aggregated buildings and left original building polygons are simplified and rectified properly.

4. DISCUSSION AND CONCLUSIONS

This paper has presented a framework to generalize road edges and building polygons. In the proposed framework, the generalization of road edges and building polygons were carried out with three modules: Road Modeler, Network Generator and Building Generalization. Road Modeler worked to create road polygons from original road edges; Network Generator is used to collapse road polygons to road networks; and Building Generalization addressed to cluster building polygons, to aggregate building clusters and to simplify both aggregated buildings and left original building polygons. With the algorithms in the framework, the created road networks are topologically connected well and suitable for GIS, such as car navigation systems. Using the created road networks as the constraints for generalizing buildings leads to the generalized results without contradiction. Some experiments have also been implemented and the experimental results showed that the present framework runs well.

The performance of an automatic process usually varies with the quality of the input data. The proposed framework and algorithms here are also partially limited by the input road edges and building polygons. For example, one-side or crossed road edges will influence the processing results of Road Modeler. One-side road edges mean that there are only right or left curbs for some road region. Crossed road edges may exist for overpasses or underpasses, such as a bridge over the other road. For such cases, the input data should be revised before automatic processing. These problems are going to be settled in the future works.



(a) Original Roads and Buildings

(b) Generalized Roads and Buildings

Figure 9. Experimental results

5. REFERENCES

Anders, K.-H. and Sester, M., 2000. Parameter-Free Cluster Detection In Spatial Databases and Its Application To Typification. Proc. of ISPRS, Vol. XXXIII (CDROM), Amsterdam.

Annita N. Wilschut, Roelof van Zwol, Jan Flokstra, Nick Brasa, and Wilko Quak, 1998. Road Collapse in Magnum. ACM-GIS 1998: 20-27.

Douglas and Peucker, 1973. Algorithm fro the reduction of the number of points required for represent digitised line or its caricature. *Canadian Cartographer*, 10(2), pp. 112-122.

ESRI, 1996. "Automation of Map Generalization: The Cutting Edge Technology", ESRI White Paper. http://www.geo.ulg.ac.be/interne/Fichiers_Interet_General/Cartographie/AutoMapGenESRI.pdf (accessed 28 Apr. 2004).

Kreveld M. and J. Peschier, 1998. "On the automated generalization of road network maps", http://divcom.otago.ac.nz/SIRC/GeoComp/GeoComp98/21/gc_21.htm (accessed 28 Apr. 2004).

Wang P.T. and T. Doihara, 2003. Automated Creation of Road Network from Road Edges. In: Proc. of the 24th Asian Conference on Remote Sensing & 2003 International Symposium on Remote Sensing, Busan, Korea.

Wang P.T., T. Doihara and W. Lu, 2002. Spatial generalization: an adaptive lattice model based on spatial resolution. In: Joint International Symposium on Geospatial Theory, Processing and Application, Ottawa, Canada, ISPRS Vol.34.