A COMBINED AUTOMATED GENERALIZATION MODEL OF SPATIAL ACTIVE OBJECTS

J. Joubran Abu Daoud, Y. Doytsher

Faculty of Civil and Environmental Engineering
Department of Transportation and Geo-Information Engineering
Technion- Israel Institute of Technology, 32000 Haifa, Israel
(jacquele, doytsher)@tx.technion.ac.il

Commission IV, WG IV/3

KEY WORDS: Generalization, Cartography, GIS, Automation, Analysis, Data Mining,

ABSTRACT:

Automating the map generalization process has traditionally been a major focus of research in cartography and GIS environment. Many algorithms and models have been developed over the past three decades, starting with the line simplification algorithms and ending with "Clarity" the new environment employing AGENTS, JAVA, XML and Topology. Although automation of cartographic generalization has been a field of extensive research, a method of usable holistic understanding generalization is still lacking. The model for combined generalization described in this paper, is intended to initiate a method that understands and describes the action and behaviour of active objects in the map generalization process. The paper focuses on the object properties analysis in order to determine the "power" of each object in any given map, and the interactions between these powers. These interactions produce "forces" that act on the objects and control their behaviour according to the cartographic constraints.

1. INTRODUCTION

Cartographic generalization aims at simplifying the representation of cartographic data to suit the scale and purpose of the map. Although automation of cartographic generalization has been a field of extensive research (Muller et al 1995, Weibel & Jones 1998, Richardson & Mackaness 1999, and many others), there is still a lack of a usable holistic understanding generalization method. Successful implementation of a generalization process is supposed to produce a good map that satisfies the cartographic requirements, rules and constraints. Recently several methods have been developed based on constraints and rules definition (Raus & Plazanet 1996, Harrie 1999, Sester 2000, Raus 2000, Harrie 2003). However, it is clear that more research on the definition and the formulation of the rules and the constraints to be used is needed. The main issue of the generalization process is to determine where the conflicts are and how to solve them without creating new conflicts. The objects in the map must not be treated in isolation, and the combined generalization should model the relationship between the objects and the way they affect each other. Several authors (Ware & Jones 1998, Sajakoski & Kilpelanen 1999, Sester 2000) have already suggested such holistic process; however, no solution has been presented for the implementation of the complete generalization process in a single continuous step.

The research described in this paper examines behaviour of the map objects and the interactions between them in order to understand the generalization process. The developed model defines several parameters that determine for each object a "power" in the map, and set rules to control the mutual interaction forces between these powers in order to compromise between the constraints and solve the competition between the objects on the limited map area at a reduced scale. The parameters are dependent on the object properties (area, type, stiffness, and shape) on the one hand, and the area properties (density, empty area surrounding an object, the map target, and the map scale) on the other. Each object acts according to its power, computed as a function of its properties and these parameters. Interactions between map objects are expressed by actions of the forces constructed around the cartographic constraints and affected by several parameters depending on the properties of the surrounding objects. The combined generalization model takes into consideration the surrounding objects and defines their properties, such as distance, type, density and topology. As a result, the surrounding objects affect and cause the "weak" objects to change their shape or place. The implementation of this new method requires: (1) determination of quantities and thresholds of each parameter, (2) definition of the rules and the constraints of each force action, and finally (3) translation of the results into one or more of the generalization operators - displacement, aggregation, selection, and enlargement.

Due to the limited scope of this article, we will discuss the interaction among objects belonging to only one layer, buildings. This paper will present the new combined generalization method of cartographic generalization, its implementation, tests on a real data subset, and the results achieved.

2. CARTOGRAPHIC OBJECTS AND THEIR PROPERTIES

The objects in the map are treated according to their properties, their type, and what they represent. The cartographic map generalization at a required scale is a process of competition among the objects over the map area. Each object has its own power as a function of its properties, the surrounding objects, map scale, and map type. This power controls the behavior of the object in the generalization process in accordance with the cartographic rules.

2.1 Basic Object Parameter

The basic parameter in the generalization process is the object area at the scale of the new map. The object symbol has its location, its dimension, and its relative importance among other objects, according to the map type. The same objects may be of different relative importance when plotted on different maps for different uses.

2.2 Shape Analysis Methods

Many cartographic objects represented on large scale maps can be considered to be geometric objects in the form of polygons or closed polylines. These geometric objects could be described by the set of shape parameters, consisting of, but not limited to, numerical values and topological descriptors (Guienko & Doytsher 2003). The shape analysis methods themselves could be formally classified by certain parameters. There are two major groups of shape analysis methods: boundary techniques (external analysis) and global techniques (internal analysis). These methods are applicable to different representations of the same object. This paper, however, focuses on the objects as polygons in vector format, derived from GIS databases.

2.2.1 Major Geometric Shape Parameters

The following major parameters can be used to describe the shape of geometric objects in polygon form: area, perimeter, centroid, major axes and angle, elongation, compactness, solidity and convexity (Guienko & Doytsher 2003). The first five parameters could be calculated using moments. Using moments for shape description was initiated by (Hu, 1962), who proved that moment based shape description is information preserving. This study concentrates on the second moment, the moment of inertia that can be used to determine the principal axes of the shape.

Moment inertia can be computed in respect to the object shape by its vertices or by its edges. Using the length of the edges of a polygon rather than its vertices is preferable as it is independent of the number and density of vertices and is a function of the polygon shape only (Doytsher 1979). Thus, the moment inertia computed hereafter is based on edges, with weights proportional to their lengths. Therefore, the moment for the two major axes in this study is calculated as follows:

$$(1) \quad Ixx = \frac{\sum (\bar{y} - y_c) \cdot dx_i}{A \sum dx_i} \qquad , \ \bar{y} = \frac{y_{i+1} + y_i}{2} \ , \ \ dx_i = |x_{i+1} - x_i|$$

$$(2) \quad Iyy = \frac{\sum (\bar{x} - x_c) \cdot dy_i}{A \sum dy_i} \qquad , \ \bar{x} = \frac{x_{i+1} + x_i}{2} \ , \ \ dy_i = |y_{i+1} - y_i|$$

In order to obtain effective information on the shape of the object, the ratio between the moments is also calculated (ratio between the larger and the smaller moments). When the numerical value of the ratio is close to 1, it indicates that the shape is approximately a square, and it is more stable than a prolonged shape (with a numerical ratio figure much higher than 1). Other parameters like compactness and solidity, can also provide useful and important information about the object's shape (Guienko & Doytsher 2003). These parameters are calculated as follows:

$$(3) \quad compactness = \frac{4\pi \cdot Area}{perimeter^2} \quad , \quad perimeter = \sum sqrt(dx^2 + dy^2)$$

$$(4) \quad solidity = \frac{area}{convexarae}$$

2.3 Spatial Analysis

One of the well-known methods of spatial data analysis is spatial data mining; a field dealing with producing new spatial data from existing data. Data mining is facilitated by utilizing shapes or properties which are not explicitly expressed in the original databases, and is performed without changing them (Kang 1997). Databases kept in GI systems constitute a valid potential for data mining due to the vast and diversified data stored in them. In this study spatial data analysis is attained by implementation of the Delaunay Triangulation and Rings analysis. Both methods allow producing large quantities of information about the data, density and shape, even though no previous explicit knowledge existed regarding these properties.

2.3.1 Delaunay Triangulation

Generally, triangulation of a planimetric surface is the method of dividing the surface into a finite number of triangles. This method is the key tool for handling problems requiring solutions based on area division according to the principles of finite element theory.

There are number of methods for performing triangulation, of which the Delaunay triangulation is preferred for cartographic purposes, since it supplies triangles with the shortest edges. At the first stage, a constrained Delaunay triangulation is performed in order to include the polygons or line edges, describing the objects as part of triangle edges. The triangles in the triangulated surface are divided into two groups: 1) triangles contained within objects, and 2) triangles stretched between objects in the intermediate space (Joubran & Gabay 2000).

It is the edges of the second group triangles, or their area that serve as an indicator of data density in the region surrounding a given object. The area of the empty triangles surrounding each object and the average length of their edges determine its density and the surrounding free area. The surrounding objects between which and a given object these triangles are stretched, affect the object's density and behavior, and are therefore defined in this study as "neighbor objects".

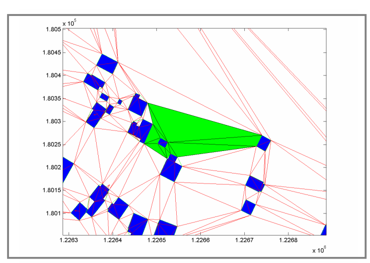

Figure 1. Sample of Delaunay triangulation - surrounding triangles stretched between an object and its neighbors

2.3.2 Ring Analysis

Spatial analysis of the surrounding area of each object is performed by dividing the area into rings around the minimal circumscribing rectangle (Doytsher 1988). The ring analysis method is based on giving a larger weight to rings closer to the object. It describes the effect of surrounding objects and the importance of the distance between them. We thus started with the ring closest to the minimum rectangle circumscribing the object. The effect of each ring on the density determination is calculated as follows:

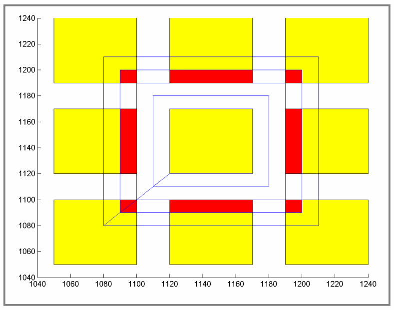

Figure 2. Ring analysis method

$$(5)\quad Density = \sum \frac{A_i W_i}{ring\_area_i \sum Wi}$$

Where:
A is the area of the objects contained in the ring
W is the weight of a certain ring

This equation, when used in the example below, produces the required results, where objects in the middle of a given area have higher density values than objects on the boundaries of this given area.

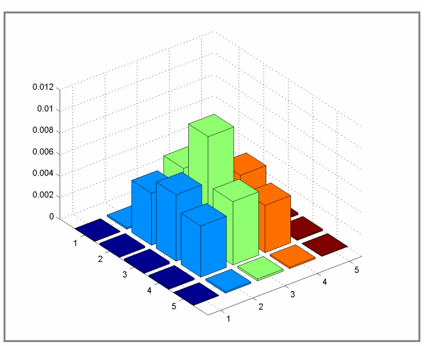

Figure 3. Density values of set of buildings

In order to be more precise about the way the objects are scattered in a given ring, the rings were divided into four parts with the free area and the minimum distance between other object are calculated in each part. The following examples demonstrate the need for such sub-analysis method:

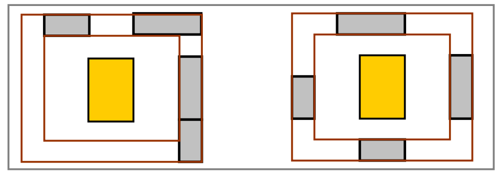

Figure 4. Side analysis

Each object and its minimum circumscribing rectangle is analyzed as demonstrated in Figure 5.

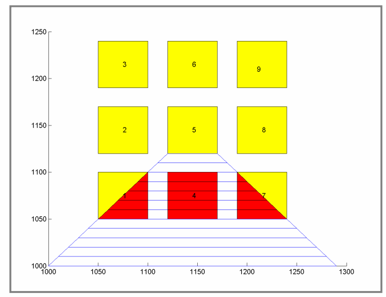

Figure 5. Ring analysis by 4 different sides

It was found that the width of the rings should be defined as identical to the cartographic tolerance based on the scale of the desired map. The number of the rings is a function of the rings' width and the area of the surrounding free region for each object.

$$Rings\_area = object\_area + object\_free\_surroundin\_area$$
$$Rings\_area = (ab + 2 * Rings\_number * rings\_width)^\wedge 2$$
$$ab = (a+b)/2$$
$$(6)\quad Rings\_number = \frac{\sqrt{Rings\_area} - ab}{2}$$

It was found by several experiments with a satisfactory result that according to this equation the number of required rings around each object is almost the same. Thus, more or less the same weight was assigned to each of the rings surrounding the different objects.

$$(7)\quad \left[\begin{array}{l} a_1 = Rings\_number \\ a_n = 1 \\ first\ \ ring\ \ wieght = \dfrac{a_i}{\sum_{j=1}^{n} a_j} \end{array} \right]$$

3. METHOD OF PERFORMANCE

Several requirements must be fulfilled in the generalization process. A possible framework for automatic generalization is to formulate these requirements as constraints and let them control the process (Beard, 1991). The major difference between rules and constraints is that the rules state what is to be done and constraints state what results should be obtained (Harrie, 2003). Since it is difficult to formalize the generalization process in the form of rules, several authors have proposed and used constraints in the generalization process (e.g. Brassel & Weibel 1988, Ruas & Plazanet 1996, Harrie 1999, Ruas 2000). In this study, the generalization process is controlled by the power of objects. These powers have been determined and thus affect and act according to the process rules. The forces that are "developed" in each object as a result of the powers action are "translated" according to its value and direction to suit the generalization operator in respect to the process constraints.

3.1 Object Power Determination

An analogy to the interaction among a large number of objects can be found in electric field theory. In an electric field each "object" acts according to its power, affects its neighbors and is in turn affected by them. In this study, it is suggested to implement the electric field theory, assuming that the map generalization process will be based on "powers" of the map's features affecting each other. The "power" is determined as a function of the object's properties, location, and the surrounding area and objects. The action of the power' action controls the object's behavior, thus it has to be calculated carefully, taking into account all affecting elements.

Object Properties

The aim of this research is to establish a model for a "combined generalization", where the powers are calculated and determined in order to be able to highlight the different qualities of each individual object. The area is a very important element in such a process; since a bigger object has a higher power value. Different objects have different factors under the cartographic rules (e.g., trees might be moved easier than buildings). According to the map type each object has its relative importance value (e.g. in a tourist map hotels will be more highlighted than private houses). In a similar manner, high buildings should be "stronger" than low buildings, and the process prefers not to change their shape or move their location. Square buildings should be "stronger" than rectangular or elongated ones.

In analogy to the electric field theory, the power contained in each object will be calculated as a function of the following object properties:

1. Area: calculated at the scale of the map (size of the plotted object or its plotted symbol).
2. Shape: calculated as a function of the compactness, solidity, and second axes moment ratio:

$$(8)\ Ishape = \frac{solidity * compactness}{ratio}$$

$$(9)\ ratio = \frac{\max(I_{xx}, I_{yy})}{\min(I_{xx}, I_{yy})}$$

3. Height: a normalized value, given to 2D objects like roads, and single story houses. The value is increased for multistory buildings.
4. Type: an elastic value for each object describing its "material" according cartographic rules and map content.
5. Importance in the map: normalized values according to the map type.

Area Surrounding an Object

The area surrounding an object affects its behavior as well. Objects can be located in a dense urban area, or "isolated" in a rural area. Objects with a higher density value resulting from more objects in the surrounding area should be "stronger" being practically unable to change their shape or moved from their location.
The values of all these elements were chosen in proportion to the expected power (larger values vs. larger power), and therefore, the power can be calculated as follows:

$$(10)\ power = area * shape * hieght * elastic * \mathrm{Im}\,por\tan ce * density$$

3.2 Forces between Objects

The forces between neighboring objects express the interaction between them. Returning to the electric field theory, each object has its "electric charge", and attraction or rejection forces control their movements. When adopting the same behavior or interaction model, the forces between the objects in the map are computed as follows:

$$(11)\ Force_{a,b} = \frac{G * (P_a - P_b)}{R_{a,b}{}^2}$$

The force between two objects is a direct function of the difference between both powers. Thus, the same style and power objects won't affect each other. However, there is an inverse function expressing the distance between the objects and their effect, with close objects having a stronger effect.

3.2.1 Minimal Distance between Neighboring Objects

It was determined that the approach should be to calculate the distance between objects as the minimal distance between the convex hulls circumscribing the objects as shown in Figure 6.

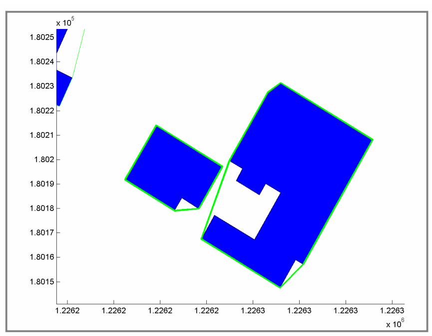

Figure 6. Minimum distance between convex hulls

3.2.2 Direction of Forces between Objects

The critical zones in a map are located where there is minimum distance between neighboring objects, especially if that distance causes a spatial conflict. The goal of the combined generalization process is to increase the distance between objects by moving the weaker objects. In order to achieve this goal, the forces acting from the mass center of the object towards the minimal distance need to be calculated, with the direction of the force issuing from the stronger object and affecting the weaker object.

3.3 Implementing Actions of Forces

The actions of forces on each object control and determine its behavior. A middle object with many forces from surrounding objects is under higher risk of being deleted if the surrounding objects are much stronger. Alternatively, based on the type of the object and its surrounding objects, the object will be clustered with them if they are all of the same type and endowed with more or less the same power level. A spatial conflict is resolved by displacing the weaker object in accordance with the value and the direction of the unified force affecting it.

4. RESULTS

In this chapter the suggested method is demonstrated on a group of polygonal entities. A map of buildings in a certain area is given, each building is described as a closed polygon composed of a known number of vertices. The numeric parameters for each object are calculated – the area of the polygon, its perimeter and its compactness. A convex hull circumscribing the polygon, and the minimal rectangle circumscribing the convex hull are computed. The determination of principal axes enables computing the polygon solidity and thus the orientation of the surrounding rings for the fitting rings analysis. A constrained Delaunay Triangulation is applied by forcing the building edges to become part of the triangle edges formed by the triangulation, as depicted in Figure 7.

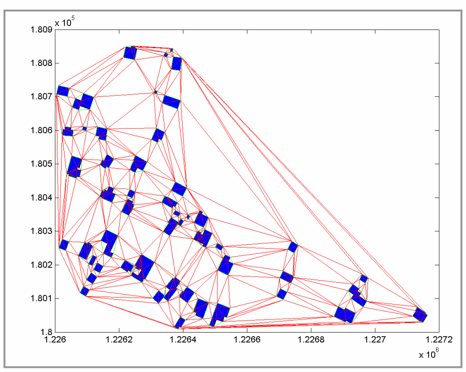

Figure 7. Constrained Delaunay triangulation

The free surrounding area for each object was calculated as the sum of the area of the triangles stretched between a given object and surrounding objects in the intermediate space. The surrounding objects connected to the free area of the given object are defined as its "affecting neighbors". Tolerance may be calculated according to the desired map scale and rings analysis. The width of the rings is equal to the tolerance value and the number of rings is calculated as a function of ring width, object area and its free surrounding area. As depicted in Figure 8, the number of rings needed for the different buildings of the given set as shown in the histogram, is about the same for all buildings.

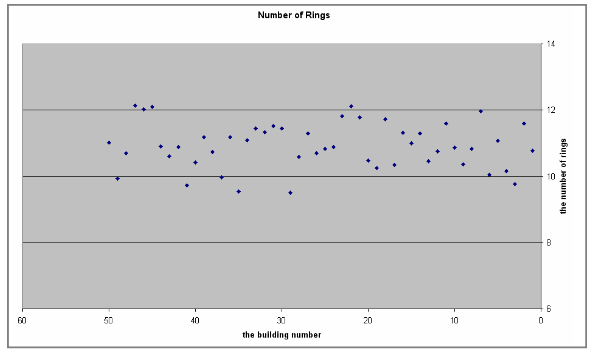

Figure 8. Number of calculated rings

Thus, each object has its power value calculated as a function of the density derived from the rings analysis, and shape parameters. The values of the powers are presented in Figure 9, as a "colored scheme".

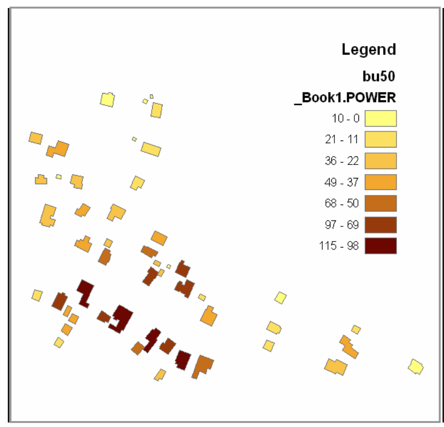

Figure 9. Powers values of buildings

Interaction between the building powers produces and is expressed by forces as shown in Figure 10. Large numerical values of forces evolving between close objects are a warning of potential spatial conflict.

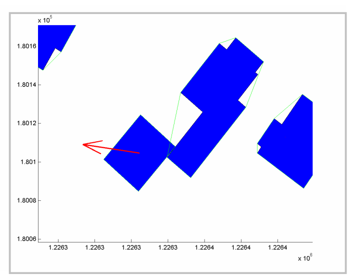

Figure 10. A force solving spatial conflict

The final results are derived from translating the action of forces to the suitable generalization operator according to the value of the balance of forces and its direction as follows (Figure 11):

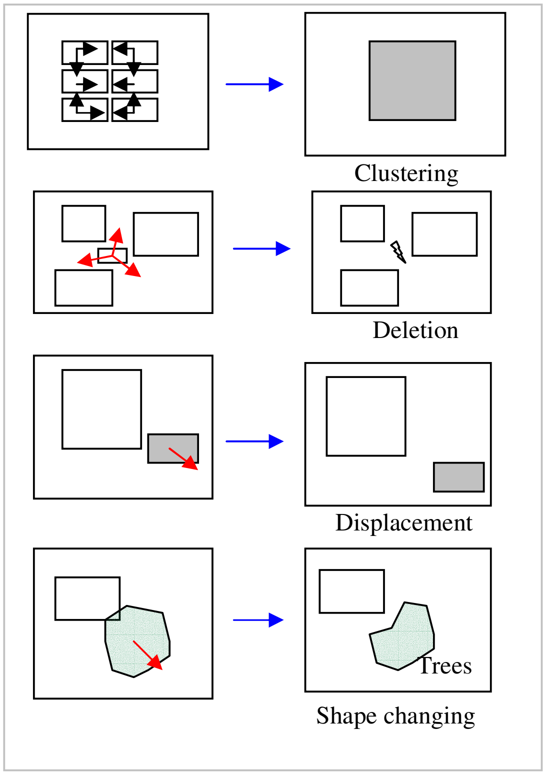

Figure 11. Forces & generalization operators

5. DISCUSSIONS AND FUTURE WORK

The method presented here for a new model of combined generalization, makes use of spatial data mining to understand the properties of objects and of topology in order to determine their behavior in the generalization process. The algorithm examines the generalization process from a new standpoint that views the map as a stage in area warfare. Each object has its power and the forces control the object's final position. Experiment results on a limited level indicated implementation of the method on objects belonging to a single layer of buildings. Additional work is still required. A more thorough investigation of object behavior in the generalization process requires adding additional layers of objects, handling linear objects concurrently with polygonal objects, and dealing with the topological relationship between the different layers.

REFERENCES

Brassel, K. E., and Weibel, R. 1988. *A Review and Conceptual Framework of Automated Map Generalization*. Intennational Journal of Geographical Information Systems, Vol. 2, pp. 229-244.

Doytsher Y., 1979. *Methods and Software for Creating Data-Bases for Town Planning Compatible for Mini-Computers*, 7th European Symposium - Urban Data Management, Hague, The Netherlands.

Guienko G., Doytsher Y., 2003. *GIS Data for Supporting Feature Extraction from High Resolution Aerial and Satellite Images.* Journal of Surveying Engineering – ASCE.

Harrie, L., 2003. *Weight – setting and Quality Assessment in simultaneous Graphic Generalization*. The Cartographic Journal, Vol. 40 No. 3, pp. 221-233.

Harrie, L., 1999. *The Constraint Method for Solving Spatial Conflicts in Cartographic Generalization*. Cartography and Geographic Information System, Vol. 26, No. 1, pp. 55-69.

Hu, M. k., 1962. *Visual Pattern Recognition by Moment Invariants.* IRE Trans. Inf. Theory, 8, 179-187.

Joubran, J., and Gabay, Y., 2000. *A Method for Construction of 2D Hull for Generalized Cartographic Representation*. Proceedings 9th Congress of ISPRS, Amsterdam, Netherlands, PP. 417-424.

Muller, J.C., Weibel, R., Lagrange, J.P., Sagle, F., 1995. *Genralization: State of the art and issues*. In: GIS and Generalization: methodology and practice, J.C. Muller, J.P. Lagrange and R. Weibel eds., Taylor and Francis, Pp. 3-17.

Ruas, A., 2000. *The Role of Meso Objects for Generalization*. Proceedings of the 9th Spatial Data Handling Symposium, Bejing, Pp. 3b. 50-3b. 63.

Ruas, A., and Plazent, C. 1996, *Strategies for Automated Generalization*. Advances in GIS Research, SDH' 96, pp. 319-335.

Richardson, D. E., Mackaness, A. W., 1999. *Computational Processes for Map Generalization*.(Introduction). Cartography and Geographical Information System, Vol. 26, No. 1, pp. 3-5.

Sarjakoski, T., Kilpelaine, T., 1999. *Holistic Cartographic Generalization by Least Adjustment for Large Data sets*, In: "Proceeding ICA", International Cartographic in Canada, Session 39-D, Index 08, pp. 31-38.

Sester, M., 2000. *Generalization Based on Least Square Adjustment*. International Archives of Photogrammetry and Remote Sensing, Vol. XXXIII, Part B4. Amsterdam, pp. 931-938.

Ware, J. M., Jones, C. B., 1998. *Conflict_Reduction in Map Generalization Using Iterative Improvement*. GeoInformatica, Vol. 2, No. 4, pp: 383-407.

Weibel, R., Jones, C. B., 1998. *Computational Perspective on Map Generalization*. GeoInformatica, Vol. 2, No. 4, pp: 307-314.