Comparison Between Generalization Processes at Large and Small Scales

Jacqueleen Joubran Abu Daoud
Faculty of Civil and Environmental Engineering,
Department of Transportation and Geo-Information Engineering, Technion- Israel Institute of Technology,
32000 Haifa, Israel
e-mail: jacquele@tx.technion.ac.il

Izabela Karsznia
Faculty of Geography and Regional Studies,
Department of Cartography, University of Warsaw
Krakowskie Przedmiescie St., 00-927 Warsaw, Poland
e-mail: i.karsznia@uw.edu.pl

***Abstract*—The presented research touches on the questions of digital cartography, particularly on automated map generalization in advanced geographic information systems. The aim of this article is the investigation and comparison of generalization processes of settlements at different detail levels. Two generalization models are described and verified at two different detail levels. On the basis of this the directions of future research are outlined.**

Keywords- generalization; small-scales; large-scales; generalization models.

I. INTRODUCTION

Although automation of cartographic generalization has been an extensive field of research [3][7][11[14][15], there still remains a lack of a usable holistic generalization method. A holistic process that makes it possible to generalize the whole map including all the layers, to take into account the connections between the layers and to deal with levels of detail in the small and large scales at once. More recently, the demand for automated map generalization, which has been longstanding in the context of conventional GIS (Geographical Information Systems), has been reinforced by the prevalence of geographical information access on the Internet, that make it more complicated. There are several types of public access map-based Web sites that allow a user to zoom in and zoom out of a particular region, but at presently, this is usually based on stepping between independent pre-processed generalized datasets which may differ markedly in their degree of generalization. It would be desirable to be able to change the level of detail on such systems in a smooth and progressive manner rather than the quantum-leap changes that often characterize the current approach [12].

The aim of this article is to investigate and compare generalization processes on different detail levels. The author's intention is to describe and compare two generalization processes of one thematic layer, which are settlements. The first process touches large-scale generalization, from 1:10 000 to 1:50 000 and the second one concerns small-scale generalization, from 1:250 000 to 1:500 000. Two different models of generalization processes are proposed: one based on electric fields theory implemented in MATLAB (by MathWorks) and the second based on mathematical morphology implemented in Clarity (by 1Spatial). The main point of interest is to show the specifics of the generalization process at large- and small-scale elaborations in terms of characteristic of the generalized thematic layer and also characteristic algorithms and tools used at different level of details. The authors want to investigate whether it is possible to build one comprehensive model to manage the on-line generalization process at different levels of detail.

With respect to the goals of the article in section II the authors present the different problems, which a cartographer has to deal with, on different detail levels. In Section III the authors concentrate on some significant differences, as well as some similarities, between generalization processes at large and small scales. In Section IV two generalization models, for both detail levels, are presented and discussed. Finally in Section V the authors conclude their research and they point at the future research directions.

II. DIFFERENT PROBLEMS ON DIFFERENT LEVELS OF DETAIL

Cartographic generalization is a decision-making process aimed at reflecting the purpose of a map or database and emphasizing characteristics and relations of generalized objects. Due to its holistic nature, the generalization can hardly be transferred into a process of sequences of tasks which might be applied in computer environment. The necessary condition of de-composing the generalization process into tasks sequence is a formalization of cartographic knowledge [13].

In many countries, there are specifications of map redaction with additional remarks on the generalization process for topographic maps. Based on them, a cartographer is able to collect and formalize knowledge about the generalization process at the large scales. Those map specifications are the source of important constraints like: threshold values, minimum or maximum values of distances.

Unfortunately, a dominating part of existing elaborations touches both the maps and spatial data generalization expressed in large scales [1], [2]. [10]. By large-scales elaboration, we understand maps and databases at the scales from 1:500 up to 1:50 000 while as a small-

 ISBN: 978-1-61208-178-6

scale elaborations we consider maps and databases from 1:200 000. The reason for that can be placed in the wide practical application of such kind of data. Basic spatial databases from country levels have been expressed just in the scales of 1: 10 000, 1: 25 000 and 1: 50 000 and hence the need for their automated generalization appears.

On the contrary, the generalization at small scales depends, in general, on the experience and knowledge of a cartographer. The result of the process depends on, not always consequent, decisions of its author. As a result, maps in particular scales or a spatial data of the same level of detail may differ from each other both in a range, as well as in a level of generalization. Due to the subjective character of the small-scale generalization process, none of the precise instructions of its redaction were elaborated on until now: what makes it significantly difficult to collect knowledge about the process, its formalization, and implementation in GIS systems.

III. LARGE *VERSUS* SMALL GENERALIZATION

Generalization methods and processes have been changed and improved alongside development in the science and art of cartography and have been surely influenced by progress in computer science [8].

In the process of cartographic generalization at every level of detail we can point at four main stages:

- Selection of categories of objects (object classes) presented on a map and their classification.
- Selection of objects within particular categories.
- Change in a cartographic method of representation – replacing an outline of area feature with a signature [9].
- Simplification of built-up area's outlines.

With respect to the holistic character of a generalization process, in this article, we concentrate on one thematic layer - settlements. The reason for that is that most cartographers consider this thematic layer as the most important and at the same time the most difficult one.

The above mentioned four generalization stages will be applied differently at large and small scales elaborations. Both in large and small-scale generalization processes, the context of geographical information and topologic relationships are to taken into consideration.

A. Characteristics of the generalization processes at large- and small-scales

There are significant differences between large- scale and small-scale generalization processes. First of all at large scales (in this article we treat 1:10 000 scale as a source) settlements are presented as separate buildings, while at small-scale elaborations, all settlements are presented as signatures, and additionally the ones of them which are highly-populated are presented also as outlines (these are built-up areas). So, in the small-scale data model settlements are placed in two thematic layers.

In the small-scale generalization process in order to achieve desirable cartographic results, the following operators need to be applied. First and the key generalization stage is a selection of information which concerns both settlements presented by signatures and presented by outlines. A selection operator is based on attribute information concerning the importance of particular settlements, like population, administrative meaning, and area for built-up areas. It is also often connected to the spatial characteristic of a whole settlements' network like density of settlements.

On the other hand, while generalizing at large scales, the goal of the process is to show the presented data with more details without spatial conflicts that destroy the correctness and the reality of the data. A selection operator is based on attribute information concerning the importance of a particular map object which is affected by its geometric and attributes properties and its topologic relation with other objects.

Another important operator during the small-scale generalization process of settlements presented by outlines is aggregation. At the small scales the goal is to put together all parts of the same locality (city) presented by outlines. The parts of a locality are usually aggregated based on two conditions: the distance between them and the name of the locality. In this article, we use aggregation operators originated from mathematical morphology (erosion/dilation) to keep the characteristics of shape of the outlines after the generalization process. At large scales this operator also plays a very important role as we aggregate the relevant buildings, based on its functions and the distance among them, in order to create built-up areas. At the same time those two aggregation processes differ in kind. At large scales, buildings will be aggregated by moving them to each other in order to decrease the presented area and increase the free area between them, while at such a scale the movement distances are small relatively. Another important difference, especially according to its relation to other thematic layers like roads is that in the large-scale generalization process, during building aggregation it is not allowed to aggregate together the buildings lying on the other side of the road. While at the small scales aggregation we are allowed to aggregate built-up areas (parts of one city) even if they are shared by the roads.

One of the last-used operators in the generalization of settlements, in both large- and small-scales processes, is a simplification operator. It is used in the generalization process of the outlines of built-up areas at small-scales but also for building simplification. The difference here lies in the application of a relevant algorithm of simplification. At small-scales where the goal is to simplify the irregular outline of built-up area we need to apply a simplification algorithm which lets us do the shape simplification while at the large scales we can, for example, use both simplification and squaring algorithm for buildings generalization.

IV. LIMITATIONS OF PROPOSED APPROACHES

The main issue in the automated generalization process is the formalization of rules and cartographic constraints definition, which makes it possible to obtain correct cartographic results.

 ISBN: 978-1-61208-178-6

A. *Small-scale generalization model*

The concept proposed by Karsznia [5], [6] comprises the collecting of cartographic data, its formalization and implementation in a form of knowledge base in the Clarity system.

In the first stage, cartographic knowledge concerning small-scale generalization of settlements was collected and formalized in a form of a rules sequence for 1:500 000 level of detail. As a result, a knowledge base concerning generalization process of settlements was elaborated.

The knowledge base in the Clarity system [5][6] consists of generalization activities together with their implementation in a form of either algorithms or respectively, generalization tools. For that purpose, the available system functions have been used as well as new algorithms (by Java programming) and spatial analyzes tools have been proposed.

An important element of a knowledge base in this system was the development of the algorithms *cluster settlements* and the application of *action polygon erode* algorithm (supplied by Ordnance Survey), made it possible for more correct results (from the cartographic point of view) to be obtained from the aggregation of part of the built-up areas presented as outlines. In Clarity it is possible to aggregate built-up areas on the basis of distances between them, by using the algorithm *clustering*. Unfortunately, this may lead to connection of even a few different localities. The modification of the built-up areas aggregation algorithm in Clarity made it possible to connect selected parts of localities under the condition that the same name be used, and the limitation of an assumed distance between them applied. As for the algorithm *action polygon erode* deriving from mathematical morphology, its application allows the proper shapes of objects to be kept, after generalization (Fig. 1).

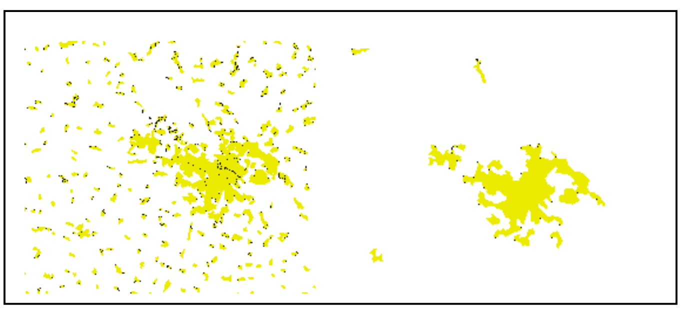

Figure 1. Aggregation results of settlements. Southern Poland

B. *Large-scale generalization model*

In large-scale elaborations where more details are needed and showed, more spatial conflicts arise and need to be solved.

`In [4], an automated process was modeled based on a sub model of a neural network to set relative importance values of the maps' objects taking into account the object's properties, its surrounding area and the map's target. Electric theories were implemented in the MATLAB environment in order to formalize the dynamic maps' object behaviors during the process of generalization.

Each object during the automated generalization process is treated individually; spatial analysis is implemented to define the object's cartographic characteristics and topological relationships with its nearby objects that should remain. The electric model set powers for each object that expressed the relative importance of the cartographic objects in the treated map according to the object properties and the map target and required scale. The interaction between the objects' powers produces forces that act in order to solve spatial conflicts and insure clear presentation of the cartographic data in the required map. The translation of the forces into generalization operators is done with the respect of the cartographic rules and constraints dealing with the objects details and connections. The cartographic rules and constraints were set and formalized by setting a sub-model of neural network that learned previous cartographers and users decisions from the input training datasets. The process of the forces action and the presentation preparing contains of three main stages: 1). simplification of each map's object and deletion of minor objects, 2).clustering close objects of the same type according to the cartographic layer properties, taking into account topologic connections (Fig 2) and 3). movement and reshaping objects in order to solve spatial conflicts, results demonstrated in Fig 3.

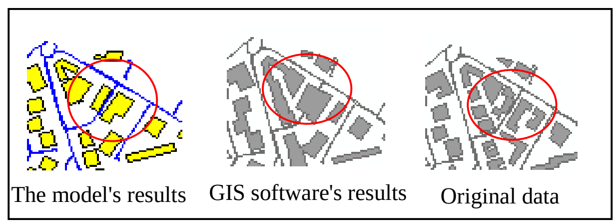

Figure 2. The aggregation results considering the roads layer and compared with GIS software results

Figure 3. The results of conflicts solutions by MATLAB in the lef

V. CONCLUSION AND FUTURE WORK

Successive and satisfying results were produced especially for large scale maps (Fig. 2), where the model elaborated in MATLAB succeeded to generalize the data taking into account the building's properties and clustered buildings of the same kind, the model as demonstrated in the same picture for the two cases preserve the characteristics of the original data. What is more settlements presented by outlines at small scale were also aggregated and simplified successfully There were thematic attributes of the settlements taken into account (aggregate settlements from the same city). Fig 4, in the left side demonstrates the original data of middle Poland, while the right side shows the results of selection stage by MATLAB according to threshold of minimum area 1000000 m, and illustrates the results of aggregation taking into account the city name and the allowable distance 250 m with the respect to the knowledge base built by Karsznia [5], [6] at 1:500000 scale.

The comparison of small-scale generalization results obtained within MATLAB and Clarity makes it possible to formulate few interesting conclusions.

- Aggregation algorithms implemented in Clarity (Fig 4), based on mathematical morphology operations: erosion and dilation, made it possible to obtain more proper results in terms of keeping shape characteristic than algorithms implemented in MATLAB system (Fig. 5).

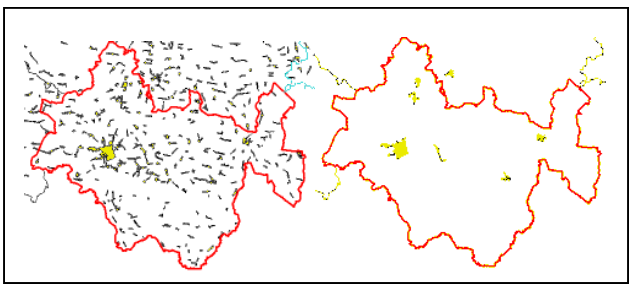

Figure 4. The original data of middle Poland at the left, and generalized ones in Clarity

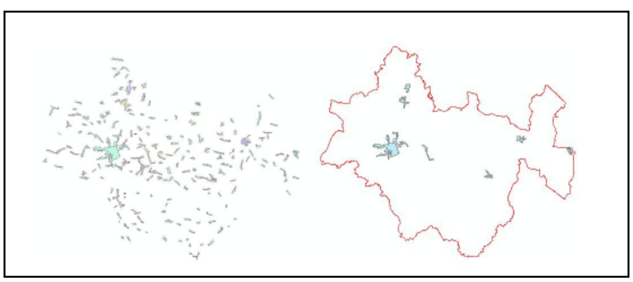

Figure 5. The original data of middle Poland at the left, and generalized ones in MATLAB

- Both aggregation algorithms, implemented in Clarity, based on mathematical morphology and also in MATLAB, made it possible to obtain more proper results in terms of keeping shape characteristic than other GIS software
- Automated generalization process by MATLAB succeeded to take into account the features properties and topological connections and produce satisfying results according to other GIS software.
- The specific character of both generalization process of small-scale and large-scale maps demands in many cases different solutions of the same problem depending on the context and objects' surrounding. In this context, important constraint is a lack of algorithms and tools having a context-like character which would make it possible to implement generalization steps on a higher conceptual level.

The conducted experiments proved that building one comprehensive generalization model to manage on-line generalization process at different levels of detail is a difficult but at the same time, challenging task.

In order to do more detailed comparison between large and small scale generalization on the way of building comprehensive generalization model, another experiment in being carried out. The cartographic knowledge of large scale generalization process collected in polish map specifications is being formalized at the moment and it will be implemented within Clarity based on Israeli data at 1:10 000 scale.

ACKNOWLEDGMENT

The authors want to thank to the Foundation for Warsaw University for financial support of presented research and to 1Spatial for Radius Clarity licensing as well as supporting during the research work.

REFERENCES

[1] Bildirici O., 2004, “Building and road generalization with the Change generalization software using turkish topographic base map data”. Cartography and Geographic Information Science, Vol. 31, No 1, pp. 43–54.

[2] Hardy P., and Lee D., and Van Smaalen J., 2008, “Practical research in generalization of european national framework data from 1:10K to 1:50K, exercising and extending an industrystandard GIS”. ICA Workshop on Generalization and Multiple Representation, Montpellier. http://aci.ign.fr/montpellier2008/papers/19_Hardy_et_al.pdf

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

[4] Joubran, A.J. and Doytsher, Y., 2008. “An Automated Cartographic Generalization Process: A Pseudo-Physical Model”. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII. Part B2. Beijing, China, pp. 419-424.

[5] Karsznia I., 2011a, “Methodical principles of automation in the generalization of selected *general geographic databases elements”*. Proceedings of the XXV International Cartographic Conference, Paris.

[6] Karsznia I., 2011b, “On automatic generalization of General Geographic Database in the applications of geographic information systems”. Geographical Journal. (in press).

[7] Kilpelainen, T., 2000. “Knowledge Acquisition for Generalization__Rules”. Cartography and Geographical Information System, Vol. 27, No. 1, pp. 41-50.

[8] Mackaness W. A., and Ruas A., and Sarjakoski L. T., 2007, “Observations and research challenges in map generalization and multiple representation”. In: W.A. Mackaness, A. Ruas, L.T. Sarjakoski (Eds.), Generalisation of geographic

information: cartographic modelling and applications. Oxford: Elsevier, pp. 315-323.

[9] Ratajski L., 1967, *Phenomenes des points de generalisation*. „Intern. Yearb. Of Cartography” , Vol. 7, pp. 143–151.

[10] Revell P., 2005, “Seeing the wood from the trees: generalising OS MasterMap tree coverage polygons to woodland at 1:50 000 scale”. ICA Workshop on Generalization and Multiple Representation, A Coruna, Hiszpania. http://aci.ign.fr/Acoruna/Papers/Revell.pdf

[11] Sester, M., 2005. “Optimization Approaches for Generalization and data Abstraction”. International Journal of Geographical Information Science, Vol. 19, No. 8-9, pp. 871-897.

[12] Sester, M. and Burrener, C., 2005. “Continuous Generalization for Visualization on Small Mobile Devices”. Proc. Conference on Spatial Data Handling –Springer, pp. 355-368.

[13] Steiniger S., 2007, “Enabling pattern – aware automated map generalization”. PhD thesis, University of Zurich.

[14] Stieniger, S. and Weibel, R., 2007. “Relations among Map Objects in Cartographic Generalization”. Cartography and Geographic Information Science, Issue 34 (3), pp. 175-179.

[15] Weibel, R. and Jones, C. B., 1998. “Computational Perspective on Map Generalization.“,GeoInformatica, Vol. 2, No. 4, pp: 307-314.

 ISBN: 978-1-61208-178-6