NORMALIZATION AND GENERALIZATION FOR DIGITAL BASE MAP PRODUCTION: CASE STUDY IN EGYPT

Mohamed Rashad Eldin and Ramadan Khalil
Transportation Department, Faculty of Engineering,
Alexandria University, Alexandria, Egypt.

ABSTRACT

This research has reviewed the conceptual and technical issues of generalization. Some definitions and history of the terminology of Generalization are presented. An actual digital map of some district in Egypt is used in the implementation. It is concluded that generalization is an important issue in the design and operation of digital maps and GIS. Some suitable rules of generalization to be considered in the Egyptian Survey Authority are recommended.

Keywords: Generalization, Normalization, Digital base map, Vector model and raster model

1. INTRODUCTION

(1-1) Definition and history of Generalization:

Generalization is the whole process, by which cartographers derive the geo-data necessary for a concrete compilation task from a set of primary data and express them by means of suitable representation [3]. A more complex form of data reduction involves changes of scale in spatial data. We may need to assemble property ownership records in an area, and the original surveyor's records may be more detailed than we require. There are two obvious options: either accept the level of detail (and thus, incorporate a greater volume of data than necessary, with the attendant increased processing and storage costs), or develop a less precise representation from the original source data. The latter is called Generalization [9].

Research into techniques for simplifying and smoothing line map features has a relatively long and well-established history in the literature of computer-aided cartography [1]. Generalization is more than a set of operations applied wholesale to groups of features classified according to data type. The subtleties of map design mean that generalization techniques are applied both within classes of objects and between different classes of objects. This reflects the fact that objects have their own behaviour which varies according to several factors, but principally map task, scale theme, symbology and contextual content [5].

Research and development in the realm of generalization began in the early seventies. At that time the goal was to support the surveying and mapping administration in the production of revised medium scale topographic map series. During the last couple of decades a number of scientific institutes engaged in the exploration of generalization processes.

(1-2) General:

Automated generalization has been regarded as one of the main challenges in computer aided cartography. Traditionally, the cartographic generalization has also been a difficult problem. The decision about the characteristics of the map objects or the elimination of nonessential features of the map data used to require cartographers to be thoroughly familiar with the maps they were making. However, communication with the graphic symbolism of maps can be learned. Originally, generalization arises from the need to reduce the spatial relations to be more comprehensible. The graphic generalization of the map was separated from the generalization mode in the real world in connection with mapping. However, research into generalization has been more concerned with the technical aspects of the generalization process in recent years [4].

2. KNOWLEDGE - BASED GENERALIZATION

One of the main topics in map generalization today is the establishing of rule - bases for generalization, a task that includes data modelling and representation techniques. Generally, establishing rule - bases is a task that seems to be strongly dependent on the application area. Two structures are currently mainly used for storing the knowledge in rule -based system for conducting generalization operations: rules and parameters tables.

Practically, knowledge - base generalization is still at the beginning. Establishing rule - bases for generalization is an on - going research area at several institutions and universities in the world. The rule - based generalization is strongly application dependent. The way of formalizing the common rules for different application areas have consequences on how general rule - bases it is possible to adopt. It can be said that for generalization tasks only rule - based techniques would not solve the problems.

In the generalization phase, additional, often attribute data is taken advantage of. For example, when storing data on building in the database, additional attribute information concerning which blocks the buildings belong to can also be stored in the database. This additional information can then be applied to aggregation problems. When, for example buildings should be aggregated into blocks. By making use of the additional information the quantity of generalization problems decreases. It has been widely discussed which kind of additional information, e.g. attribute data, should be stored in the GIS databases, as well as what kind of structure the data should have. But there is a lack of studies how we can make real use of this attribute data.

Spatial knowledge is knowledge in which the spatial component is important, especially via locators such as coordinates or place names. In generalization problems, we often suffer from lack of spatial knowledge in its context. Questions that should be asked are, what knowledge do we actually capture, what knowledge do we need and how should we represent this knowledge ? Spatial knowledge can be deduced by logical reasoning. Therefor we need systems to support techniques of this kind. It can be said, on the technical side, encoding spatial knowledge is difficult because logic programming and computational geometry must be combined. The challenge we face is to describe spatial knowledge so precisely that it can be processed by computer.

3. CONTENTS OF CARTOGRAPHIC GENERALIZATION KNOWLEDGE

The cartographic generalization have a good gradation and sequence. Firstly, the cartographer needs to determine the contents which the map will show based on the map use. This process is the selection of map feature; second, he also needs to determine the object of selection. In this process the rule of "priority' of feature must be considered; third, simplify the graphic which had selection. In this process the rules' priority' and 'major characteristic" must be related to. And at the same time the characteristic of quantity and quality are generalized. This process is realized by classify and different levels; last, use the displacement to deal with the relations of all features. This process will be related to the rules of "priority" [10]. The identification of rules and their implementation into a system which can simulate the work of a traditional cartographer is one of the most difficult challenges facing the GIS research today [7].

4. THE COMMON OPERATORS OF GENERALIZATION:

(4-1) *General:*

There are about twenty operators as follows. 1. simplification, 2. smoothing, 3. area to line collapse, 4. merging, 5. aggregation, 6. reselection, 7. scale change, 8. refinement, 9. enhancement, 10. abbreviation, 11. graphic association, 12. area to point conversion, 13. dissolution thru classification, 14. combine, 15. selective omission, 16. amalgamation, 17. local displacement, 18. masking, 19. symbolization, 20. exaggeration [6].

(4-2) Computer - assisted generalization:

The whole process of cartographic generalization can be divided into seven elementary procedures as follows.

1. Simplification of object geometry (e.g. middle axes, contours).
2. Enlargement of cartographic representation for the sake of visualization of map symbols.
3. Displacement of symbols with less priority near the larger ones.
4. Combination of clustered symbols into a representative one.
5. Selection of objects which are important for the cartographic purpose.
6. Classification, i.e. assignment of certain objects to an abstract class according to their semantic information including the consequent variations of geometric information or even the change of original geometric types of objects as well as the corresponding alteration in their representation (e.g. transition to the symbol representation).
7. Emphasis through enlargement of a representation based on the evaluation of semantic object information.

5. GIS, SCALE AND GENERALIZATION:

Computer - assisted generalization of spatial information is a most important function of a GIS software package for two main purposes. 1) in order to derive a digital model of reality in a purpose oriented resolution of geometric and semantic information from a genuine model, 2) in order to display digital spatial information cartographically, i.e. visualization by some cartographic rules.

GIS are established and used in order to support spatial analysis in geosciences, assessment and planning of environment, areal statistics etc. Generalization processes are necessary in both cases: model generalization and cartographic generalization [2].

A fundamental issue is to decide at which scale the information should be generalized. Ideally, it would be useful to be able to vary the scale according to the level of precision required.

6. VECTOR - AND RASTER - MODE GENERALIZATION:

Vector and raster models express different perceptions of geographical space. Whereas vectors represent geographical objects qualified by their attribute location. raster cells are spatial containers without regard for any objects, attributes, or properties of space within them. Therefore, it is expected that the underlying principles of vector - mode and raster mode generalization will be different.

(6-1) *Vector - mode generalization:*

Vector - mode generalization focuses on the simplification, selection and enhancement of linear objects. The lines may be open or closed depending on the topology of the objects they describe (e.g. rivers versus administrative units). Eighty per cent of all objects (points, lines and areas) which are found on a typical medium - scale topographic map consist of lines. This, in part, explains the considerable interest raised by the issue of line generalization. Numerous algorithms to perform line generalization have been published.

The amalgamation of smaller polygons into larger polygonal units requires a selective elimination of arcs and represents a special case of vector - mode generalization. The operation is based on the statistical generalization of the associated attributes. The generalization of buildings and built-up areas received much attention among Cartographers. Efforts have concentrated on the development of computer - assisted procedures for the generalization of large - scale topographic land survey maps up to about 1 : 5000. The generalization of linear features (e.g. streets) and areal features (e.g. building blocks) can be largely automated, Figure (1). Recent efforts, however, include structural - conceptual generalization as well as using pattern recognition and expert systems for more complex cartographic generalization models [7].

(6-2) Raster - mode generalization:

Whereas vectors lead naturally to an object - oriented approach to generalization, rasters provide the framework for generalization of the attribute component of the data. This is largely a difference of points of view, however, since operationally generalization of objects and their attributes is closely intertwined. One emphasizes data representation while the other is concerned with classification of phenomena. Both affect each other, as has been illustrated in the case of polygon filtering: generalization of the attribute leads to a generalization of the object and vice versa.

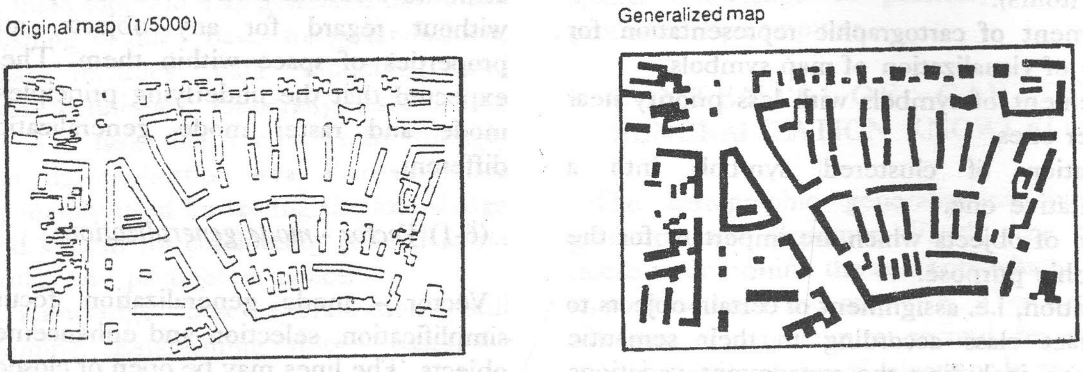

Figure 1. Computer-assisted simplification of building outlines.

7. IMPLEMENTATION AND ANALYSIS:

Two study cases are introduced: Firstly, different levels of generalization of an original coastline can be illustrated in Figure (2). The number of points indicates the levels of compression of the digital file originally used to describe the coastline [8].

Secondly, a case study is presented for Edco district, Behaira governorate, Egypt. The cadastral data is measured by using total stations. Old fixed points were related to some recent fixed points positioned by GPS. These measurements were attained in march, 1996. Some existing maps were also used. Measurements were analyzed and adjusted by using least - squares software. All data were input and drawn by using the Auto - CAD technique. The results were obtained in output by Laser printer. This work was done in Egyptian Surveying Authority, Damanhour branch.

The complexities of map generalization can best be illustrated through the presented application. Seven maps of the same area at scales of 1 : 2500, 1 : 3000, 1 : 5000, 1 : 10000, 1 : 20000, 1 : 25000 and 1 : 50000 are shown in Figure (3). In the analysis it is apparent that many of the buildings will become imperceptible, some of the polygons will become too narrow, as will linear features. In the present solution, many factors should be taken into account. For example the importance of the roads and the importance of conveying the notion of a district with a name associated with it. Some text information should be sacrificed at the expense of exaggerated features, while other salient text has to be enlarged. These and other rules of generalization are summarized in the following points:

1. Reduction in content;

   Omitted: Some boundaries, names, selective buildings, roads.

2. Increase in content;

   Added: Numerical information.

   Classified: The important roads have to be apparent.

3. Changes in symbology;

   Shades which are used to color the fields have to be altered.

4. Exaggeration;

   Symbol footprint of buildings has to be enlarged so as to appear visible, as have other features such as the roads.

5. Reclassification and symbolization;

Houses which are represented as point symbols have to be grouped and represented as area fills between roads.

6. Displacement;
   Some of the building point symbols that remained have to be displaced.

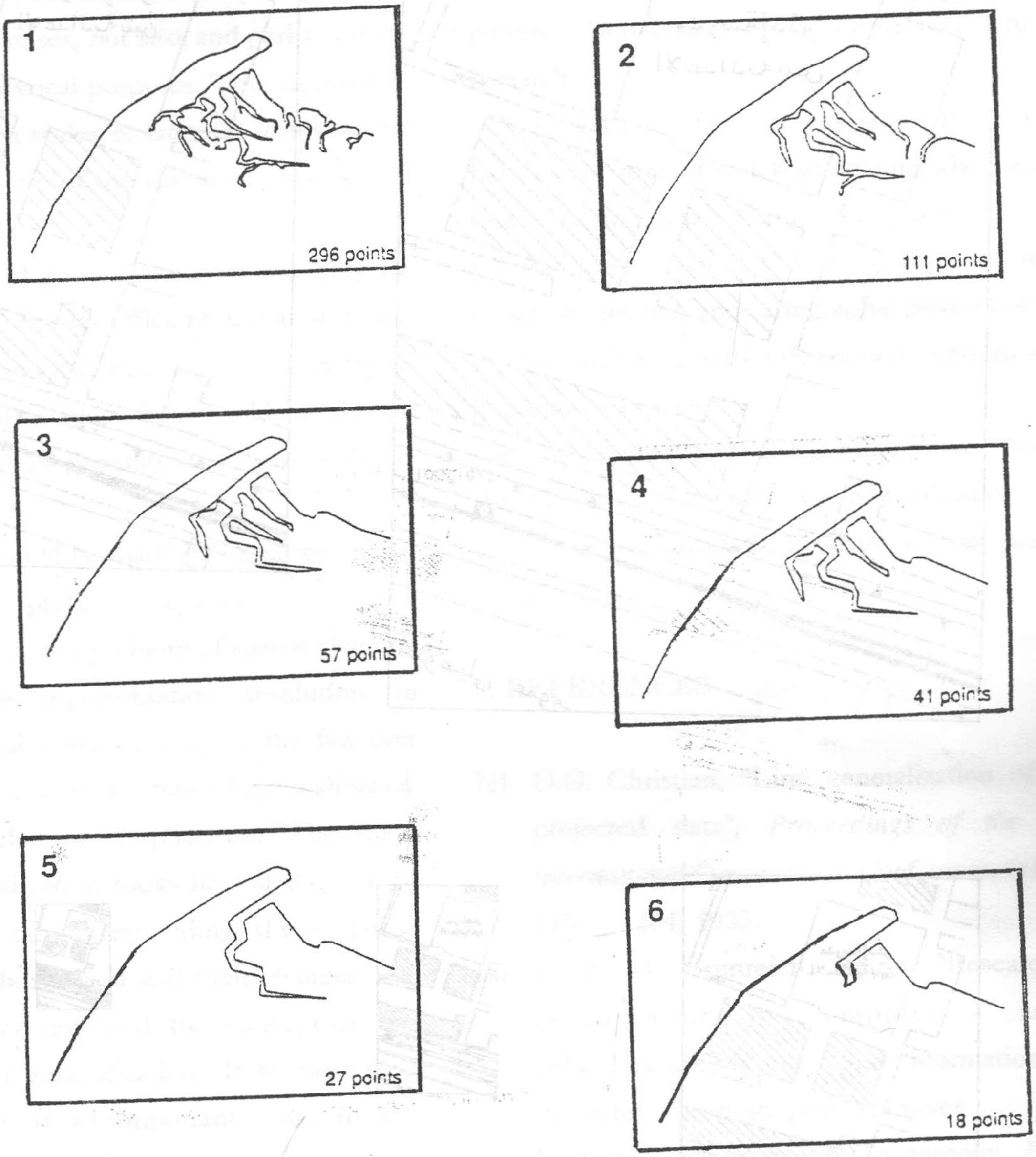

Figure 2. Different levels of generalization of the original map.

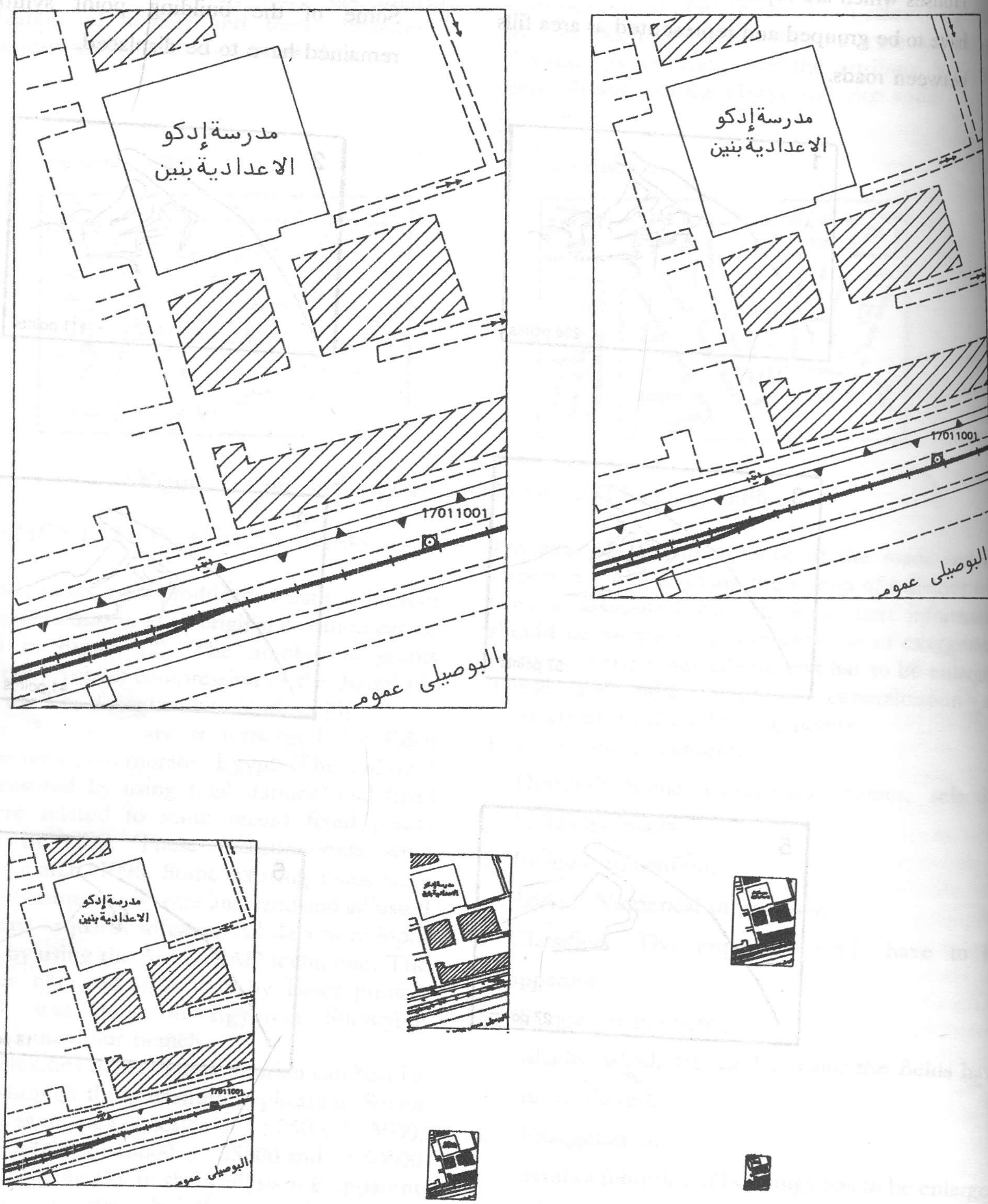

Figure 3. Seven maps of the same area at different scales,

8. CONCLUSIONS AND RECOMMENDATIONS:

Generalization is performed for map display and communication purposes, but also, and perhaps more importantly, for analytical purposes. The necessity to understand at which scales or range of scales spatial processes occur is one of the driving forces behind generalization today.

The importance of generalization is not only for graphic display but also for efficient and appropriate spatial analysis. The challenge we face is to try to find a flexible data model, which could be able to satisfy the demand of generalization and multiple representations.

A greater effort should be made into understanding the generalization problem separately from the visualization context, as a problem of spatial changes varying from one representation resolution to another. Theoretical understanding of the function is essential. The conceptualization of generalization problem should also be emphasized. The main problem for generalization today lies on the task of conceptualization, on understanding the essential nature of spatial phenomena and their changes.

This research has reviewed the conceptual and technical issues of generalization. It is concluded that generalization is an important issue in the design and operation of GIS.

Procedural and a combination of procedural and logical approaches for generalizing geometrical and semantical aspects of geographical features were illustrated by various examples.

The recommendation in the field of computer - assisted generalization is to develop an expert system, with which cartographic representations of spatial information can be purposely derived from the only one 'large scale' digital environment model.

It is strongly recommended that a rule - based generalization system should be subjected to future research.

The traditional view of cartographic generalization has been extended to a boarder one when applying GIS techniques. Computer assisted model generalization is needed as a new, economic method of data acquisition and cartographic generalization for representation of GIS information with increasing importance as well.

It is recommended that Egyptian Survey Authority should consider the presented generalization rules in this research manually or by using of some automatic techniques.

9. REFERENCES

[1] D.G. Christian, "Line generalization of non - projected data", *Proceedings of the 16 th International Cartographer Conference*, vol. 2, pp. 1195 - 1204, 1933.

[2] D. & M. Grunreich, Bernd, "Research and development in computer - assisted generalization of topographic information at the institute of cartography, Hanover University", *Institute of Cartography, Hanover University*, Appelstr. ga, D - 3000 Hannover 1, Germany, 1992.

[3] D. Grunreich, "Generalization in GIS environment", *Proceedings of the 16 th International Cartographer Conference*, vol. 1, pp. 203 - 209, 1993.

8. CONCLUSIONS AND RECOMMENDATIONS:

Generalization is performed for map display and communication purposes, but also, and perhaps more importantly, for analytical purposes. The necessity to understand at which scales or range of scales spatial processes occur is one of the driving forces behind generalization today.

The importance of generalization is not only for graphic display but also for efficient and appropriate spatial analysis. The challenge we face is to try to find a flexible data model, which could be able to satisfy the demand of generalization and multiple representations.

A greater effort should be made into understanding the generalization problem separately from the visualization context, as a problem of spatial changes varying from one representation resolution to another. Theoretical understanding of the function is essential. The conceptualization of generalization problem should also be emphasized. The main problem for generalization today lies on the task of conceptualization, on understanding the essential nature of spatial phenomena and their changes.

This research has reviewed the conceptual and technical issues of generalization. It is concluded that generalization is an important issue in the design and operation of GIS.

Procedural and a combination of procedural and logical approaches for generalizing geometrical and semantical aspects of geographical features were illustrated by various examples.

The recommendation in the field of computer - assisted generalization is to develop an expert system, with which cartographic representations of spatial information can be purposely derived from the only one 'large scale' digital environment model.

It is strongly recommended that a rule - based generalization system should be subjected to future research.

The traditional view of cartographic generalization has been extended to a boarder one when applying GIS techniques. Computer assisted model generalization is needed as a new, economic method of data acquisition and cartographic generalization for representation of GIS information with increasing importance as well.

It is recommended that Egyptian Survey Authority should consider the presented generalization rules in this research manually or by using of some automatic techniques.

9. REFERENCES:

[1] D.G. Christian, "Line generalization of non - projected data", Proceedings of the 16 th International Cartographer Conference, vol. 2, pp. 1195 - 1204, 1933.

[2] D. & M. Grunreich, Bernd, "Research and development in computer - assisted generalization of topographic information at the institute of cartography, Hanover University", Institute of Cartography, Hanover University, Appelstr. ga, D - 3000 Hannover 1, Germany, 1992.

[3] D. Grunreich, "Generalization in GIS environment", Proceedings of the 16 th International Cartographer Conference, vol. 1, pp. 203 - 209, 1993.

[4] T. Kilpelainen and T. Sarjakoski, "Knowledge based methods and multiple representation as means of on - line generalization", *Proceedings of the 16 th International Cartographer Conference*, vol. 1, pp. 211 - 220, 1993

[5] W. Mackaness and K. Beard, "Graph theory and network generalization in Map Design", *Proceedings of the 16 th International Cartographer Conference*, vol. 1, pp. 352- 362, 1993.

[6] W.A. Mackaness, "Knowledge of the synergy of generalization operators in automated map design", *The Canadian Conference on GIS Proceedings*, June 6-10, Ottawa, vol. 1, pp. 525-536, 1994.

[7] J.C. Muller, "Generalization of spatial databases", *Geographical Information Systems: Principles and Applications*, vol. 1, pp. 457 - 475, John Wiley & Sons, Inc., New York, 1991

[8] J.C. Muller, W. Pneg and Z. Wang, "Procedural, logical and neural nets tools for map generalization", *Proceedings of the 16 th International Cartographic Conference*, vol. 1, pp. 181 - 191, 1993.

[9] J. Star and J. Estes, "Geographic Information Systems: an introduction", Prentice Hall, Englewood Cliffs, New Jersy 07632, 1990.

[10] J.Y.F. Wu Wang and Z. Wu, "The research on tools of the cartographic generalization expert system", *Proceedings of the 16 th International Cartographer Conference*, vol. 1, pp. 221 - 227, 1993.