AN ATTEMPT TO AUTOMATED GENERALIZATION OF BUILDINGS AND SETTLEMENT AREAS IN TOPOGRAPHIC MAPS

M. Basaraner * and M. Selcuk

Yildiz Technical University (YTU), Department of Geodetic and Photogrammetric Engineering,
34349 Besiktas Istanbul, Turkey - (mbasaran, selcuk)@yildiz.edu.tr

Commission IV, WG IV/3

KEY WORDS: Cartography, GIS, Generalization, Object, Database, Scale, Analysis, Processing.

ABSTRACT:

Generalization of topographic maps is a very challenging problem for map producers. Therefore, NMAs are intensively working on the matter in order to make it automated as much as possible for their production requirements. In this paper, we present a case study for automated generalization of buildings and settlement areas in Turkish topographic maps from 1:25K to 1:50K, which is implemented in Laser-Scan LAMPS2 GIS and map production software based on object-oriented database technology. It begins with the issues regarded in their generalization. Then, steps for generalization of surrounding roads of settlement areas are mentioned in a limited focus. After that comprehensive steps for generalization of buildings and settlement areas are given. At the beginning, settlement areas were stored as a whole without any direct interaction with roads. Their independent generalization created some problems and we did not have the possibility of analysing the areas surrounded by roads for the decisions in some building generalization operations. To solve these problems, we create settlement blocks using road segments after creating buffers on roads considering symbology and then partition existing settlement areas according to these blocks. After that voronoi diagrams are created and combined according to building clusters. It enabled us to analyse within the blocks for optimal generalization decisions. First results of this ongoing study are close to solution although some editing is required. It concludes with an evaluation of results, addressing future work.

1. INTRODUCTION

Maps at various scales and types are needed in different fields such as urban and regional planning, geosciences, transportation, natural resource management, environmental protection, defense, tourism, statistics, education, etc. Cartographic generalization is used for this purpose and despite intensive research for last 30 years, a completely satisfying solution could not be found. Among main reasons are initiative component of map design (Weibel, 1995; Spiess, 1995), and necessity of more advanced techniques for spatial data modeling (Ruas, 1998; Ormsby and Mackaness, 1999; Weibel and Dutton, 1999), analysis (Ruas, 1998b), interpretation/mining (Sester, 2000; Anders and Sester, 2000) and processing (Weibel, 1997).

Cartographic generalization has been made by experienced cartographers in many NMAs until today. However, the requirement of building one master database and deriving other lower LoDs or smaller scales from this database mainly due to economical reasons; data updating and map revision problems especially in large countries; national, regional and global SDI activities; intensive demands of society for geospatial data and digital maps together with increasingly widespreading use of GIS, web maps and map-based mobile guides make it essential to automate generalization.

Regarding these requirements and developments, Turkish NMA, GCM, has started a project to obtain 1:100K maps from base topographic maps at scale 1:25K and later 1:50K will be dealt with. In this paper, we will present a case study for cartographic generalization of buildings and settlement areas from 1:25K to 1:50K, aiming at supporting the ongoing studies of generalization and national SDI in Turkey.

2. CARTOGRAPHIC GENERALIZATION OF BUILDINGS AND SETTLEMENT AREAS

Cartographic generalization is responsible for reducing complexity in a map in a scale reduction process, emphasizing the essential while suppressing the unimportant, maintaining logical and unambiguous relations between map objects, and preserving aesthetic quality. The main objective then is to create maps of high graphical clarity, so that the map image can easily perceived and the message the map intends to deliver can be readily understood. Scale reduction from a source map to a target map leads to a competition for space among map features caused by two cumulative effects: at a reduced scale, less space is available on the map to place symbols representing map features, while at the same time, symbol size increases relative to the ground it covers in order to maintain size relations and legibility. These can be resolved by simplifying symbolism, by selecting only a subset of features to depict, and by displacing some features away from others (Weibel and Dutton, 1999).

Buildings and settlement areas are among dominant object types in topographic maps. Building generalization is an important step in the generalization process for maps at medium scales (until 1:100K). Modifications of shape and modification of location are led by geometric, topological and gestalt constraints as design considerations. Whilst the shapes of buildings are usually the most affected by generalization (changed into rectangle or square in the best case, typified, removed, aggregated or amalgamated otherwise), it appears that for topographic maps (1:25K and 1:50K), discriminative characteristics of each particular buildings are retained as much as possible (Regnauld et al., 1999). Buildings are enlarged according to visual graphic resolution thresholds to be visible at target scale and their relative sizes are tried to be preserved.

* Corresponding author.

Settlement areas (residential area, industrial and commercial area etc.) are formed by close (indiscernible at target scale) buildings in dense areas during data collection or generalization. General character of an area (urban area, suburban area, rural area) is preserved. A long list on generalization constraints for buildings and settlement areas is given in AGENT Cons. (1998). Surrounding roads of settlement areas must be generalized because they create border for building blocks. Geometric accuracy and road characteristics are preserved within scale limits. Besides, internal conflicts must be eliminated, being generated by symbology and important parts of roads should be emphasized especially in sinuous roads. Another critical question is how to generalize road networks since it will be very dense otherwise. Perceptual grouping principles proposed by Thomson and Richardson (1999), and structural representation by graph principles proposed by Jiang and Claramunt (2002) can be considered for this purpose.

Müller (1990) analysed German topographic map series and found some facts about buildings and settlement areas, given in Table 1 and Table 2. His research shows us contextual character of cartographic generalization with the different changes in building quantity in dense and scattered settlement areas and also different size changes of buildings.

Scale	**Roads**	**Buildings**	**Settlement Areas**
1:5K	no change	no change	no change
1:25K	$\times 2$ - $\times 4$	little change	no change
1:50K	$\times 4$ - $\times 8$	$\times 1.5$ - $\times 2$	$\times 1.2$
1:100K	$\times 6$ - $\times 16$	$\times 2$ - $\times 4$	$\times 1.5$
1:200K	$\times 32$	$\times 4$ - $\times 8$	$\times 2$

Table 1. Size changes for roads, buildings and settlement areas (Müller, 1990).

Scale	**Dense Settlement Areas**	**Scattered Settlement Areas**
1:5K	no change	no change
1:25K	% 60-80 preserved	no change
1:50K	% 30-40 preserved	% 80 preserved
1:100K	% 10 amalgamated in blocks	% 30-50 preserved
1:200K	% 0-3 amalgamated in blocks	% 0-10 preserved

Table 2. Changes in building quantities in dense and scattered settlement areas (Müller, 1990).

Ormbsy and Mackaness (1999) propose phenomenological approach for generalization regarding geometry, semantic meaning and interrelationships of objects. Mackaness and Ruas (1997) states that decisions of generalization depend on an understanding of geographical situation (context) and geographical context must be made explicit for successful automated cartography. Brassel and Weibel (1988) mentioned from this in their generalization model as structure recognition.

3. A CASE STUDY FOR CARTOGRAPHIC GENERALIZATION OF BUILDINGS AND SETTLEMENT AREAS

3.1 General Considerations and Approaches for the Generalization

In this case study, LAMPS2 software and its programming language Lull is used. Here generalization of roads and road networks are given in a limited focus while generalization of buildings and settlement areas are dealt with in detail.

Sequence and selection of generalization operations, and parameter selection are important since they can cause different design solutions for target map. Therefore, a logical approach should be used in determining generalization sequence and parameters considering possible effects on each other.

In road generalization, basic operations are simplification, smoothing and selection (of subset of road network) respectively. Besides, displacement and local enlargement can sometimes be necessary.

In building and settlement area generalization, operations are collapse, symbolization, simplification, enlargement, amalgamation, aggregation, typification, elimination, displacement.

To characterise the buildings, some shape measures are generated, which are compactness, rectangularity, convexity, elongation, corner number, granularity, orientation.

In the first approach we tried, settlement areas were collected and stored as a whole and they have no direct interaction with roads. In general roads create boundaries for settlement blocks and give a possibility for controlling the generalization in manageable parts. Independent generalization of roads, buildings and settlement areas can create some problems such as very small parts of settlement area objects falling within a settlement block i.e. the area surrounded by roads, after road generalization and symbology. Besides we did not have the possibility of analysing the areas bounded by roads for the decisions in some building generalization operations such as aggregation, amalgamation, typification and displacement. So, the results were partly satisfactory. To solve these problems, we create settlement blocks using road segments after creating buffers on generalized roads at the symbol sizes giving in the specification (GCM, 2002) by regarding target scale and then partition existing settlement areas according to these blocks. Thus, building and settlement area generalization problem is converted to giving appropriate generalization decisions within each block.

To characterise the blocks, density, number of buildings, number of dominant buildings, biggest building, average building, smallest building, common building type, common building total area, total settlement area, number of settlement area object, black and white ratio etc. are computed.

Another question rising is how we will give these decisions optimally. As stated before, geographical context must be made explicit for successful generalization decisions. Among solutions to this problem are minimum spanning tree (Regnauld, 1996), Delaunay triangulation (Jones, 1997;

Ruas, 1998b), multi-variate clustering (Ormbsy and Mackanness, 1999).

For this purpose, we decided to use buffering technique and voronoi diagrams (polygons). Using the semi-size of minimum distance between two building symbols (10 m - 1:50K) according to visual graphic resolution (cartographic minimum sizes) as buffer size, we can find buildings in conflict for target scale. As can be guessed, combined buffers (building clusters) are created after individual building generalization otherwise no conflict will occur. Later, vertices of blocks and buildings are derived and using them, voronoi polygons are created and partitioned according to blocks, and combined according to the clusters (Figure 1).

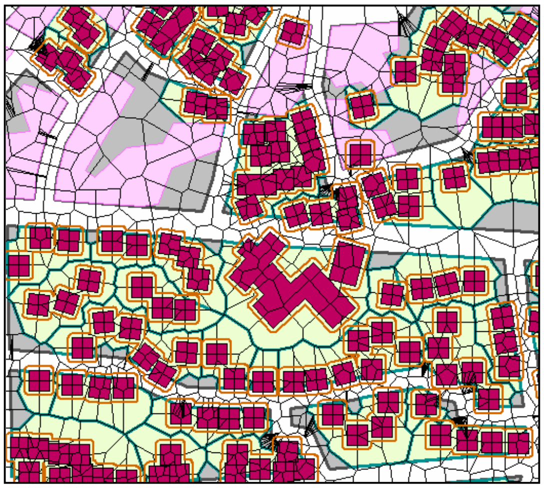

Figure 1. Combined voronoi polygons of building clusters (pink: settlement areas, light green: combined voronoi polygons, gray: empty areas)

3.2 Road Generalization

Road generalization mainly consists of three steps: simplification, smoothing and selection. While simplifying the roads in order to maintain geometric accuracy, parameters must be selected carefully. For this purpose, Douglas-Peucker algorithm is used with 0.2 mm (10 m – 1:50K) tolerance value (band with) according to visual graphic resolution. Thus, road geometry will be within a 10 meter-tolerance band, namely accurate within the scale limits. Besides, angle tolerance and vertex separation is controlled. Later, smoothing is applied but this can create a deviation more than the tolerance. This can sometimes be useful to prevent road symbols from self- or inter-overlapping however this will not work in every case. Sinuosity of a line, the ratio of the distance between first and last points of the line to the line length, can be used here to decide smoothing parameter. But this will also not give good results every time. Segmentation strategy and more advanced measures and algorithms are needed here. Besides, local enlargement or caricature (Plazanet et al., 1998) for emphasizing shape characteristics can be necessary. These are beyond scope of this paper. Consequently, some small corrections are made interactively.

Generalization, i.e. selection of subset, of road networks is another important issue. Due to the increase in symbol sizes of roads and buildings, namely in density, while scale decreasing, it will not possible to show every road at target scale. So, we have to generalize road networks. According to Thomson and Richardson (1999) "good continuation" perceptual grouping principle can serve as the basis for analysing a road network into a set of linear elements, i.e. 'strokes'. Further analysis allows the strokes to be ordered, to reflect their relative importance in the network. The deletion of the elements according to this sequence provides a simple and effective method of generalizing (attenuating) the network. Jiang and Claramunt (2002) proposes a novel generalization model which retains the central structure of a street network, it relies on a structural representation of a street network using graph principles where vertices represent named streets and links represent street intersections. In our study, only computer-assisted techniques are used.

To tackle with this problem, four criteria are used in removing the road segments interactively: road type, connectivity to main roads, continuity with same orientation and the area of blocks created using the buffers of surrounding roads at their symbol sizes. Important roads are always retained. At this scale range (from 1:25K to 1:50K), inner-city roads usually needs generalizing because of their density. By means of a simple code, we select a few blocks interactively until their total area is about 1 ha (= 10 000 sq m) regarding first three criteria. When the size criteria are met, blocks are combined and the road segments intersecting these new blocks are removed. After finishing this operation, roads are displaced if they have conflicts with each other. Finally settlement blocks surrounded by roads are created.

3.3 Building and Settlement Area Generalization

Steps for building and settlement area generalization are given below:

- Select complex shaped buildings and enlarge 50%. The criteria are corner_number >= 6, compactness > 1.65, rectangularity < 0.75, convexity < 0.9 and 625 <= area <= 2 000 sq m. Only these buildings are enlarged before simplification to increase the possibility of preserve their shape characteristics.

- Square (the edges of) buildings if rectangularity <>1.

- Simplify buildings if area > 416 sq m.

- Collapse and symbolise buildings (as minimum sized square polygon – 625 sq m at 1:50K) if area <= 416 sq m, enlarge 50% if 416 <= area <= 2 000 sq m and granularity >= 10 m and shape is not complex (see first step).

- Enlarge or diminish the size of building if the ratio of its first and last areas is different from 1.5 and not square (compactness <> 1.27) (and granularity >= 22.5 m – in case of diminishing).

- Change the elongation of buildings by preserving their area if rectangularity = 1 and compactness > 1.27.

- Create single and combined buffers (clusters) of buildings with 0.5*minimum separation value. The buildings in clusters are in conflict at target scale.

- Create settlement blocks among surrounded roads (see previous section).

- Create voronoi diagrams (polygons) using the vertices of buildings and blocks, and combine them if they intersect with the buildings in same cluster. In case a voronoi polygon contains two buildings in different clusters, then divide the polygon.

- Compute voronoi density, i.e. the ratio of total area of single buffers of buildings at same cluster and area of voronoi polygon.

- Check block density, if very dense(>%90), copy block as settlement area.

- Find close buildings to settlement area, check voronoi density of the polygon which the clusters of buildings are within and if the density >= 55% and the number of remaining buildings in the cluster <= 2 then aggregate (re-classify and amalgamate) all buildings in the cluster with the settlement area, otherwise aggregate only close buildings with it. Remaining buildings will be typified if their shapes are similar otherwise amalgamated or displaced (see next step).

- Simplify settlement areas. Aggregate holes with settlement area.

- Check voronoi density of each polygon in the blocks (Figure 2) and displace if the density < 55%, otherwise typify or amalgamate. Typification_distance = 35 m (minimum symbol granularity / 2 + minimum separation distance – 1:50K) if 55% <= density < 90%, typification_distance = 35 * $\sqrt{voronoi\_density/90}$ if density >= %90. What will happen in neighbouring polygons may sometimes have effect on decisions but not considered here.

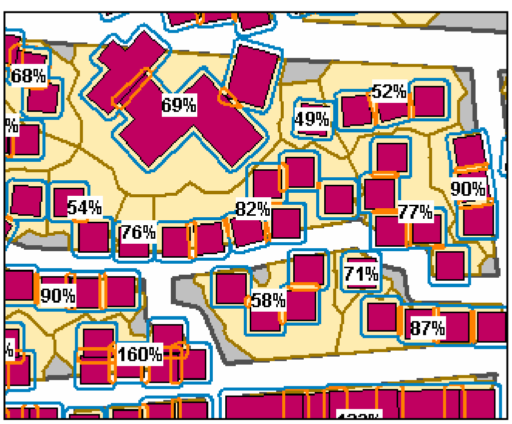

Figure 2. Density of voronoi polygons (voronoi_density)

- Apply amalgamation if the clusters have buildings with different from square or rectangle. Before amalgamation, orientations of buildings can be checked and small ones may be perpendicularly or parallelly rotated from its nearest point to the other building if their orientations are rather different or if it does not intersect another building after rotation. This is another point that will be considered later.

- Before typification, check the cluster homogeneity according to shape and size. Buildings are mostly typified if their shapes are square and sometimes rectangle. Latter was not considered in the study, but average size may be determined using area and elongation, and typification distance can be computed. If the shapes are different from square or rectangle, re-create the clusters including similar shapes. We do not change parameter or voronoi polygon here. This might be considered before clustering we can then not have direct information about conflicts if they are in a separate polygon. In the situation we preferred, the buildings important from semantic and/or geometric aspects will be given priority. In case of over density, unimportant buildings with different semantic meaning may be eliminated. This was also not considered in the study.

- Collapse relevant buildings, namely change their geometry to point. Hierarchical clustering by dendrogram is done using collapsed (changed to points) buildings before typification (Figure 3). The dendrogram is built by repeatedly finding the two closest points (according to typification_distance) being considered by the process, adding them to the tree, creating a new node to represent the cluster defined by the two new nodes. This new, average node is then added back into the pool being considered for finding the closest pair. In this way, the number of nodes in the pool is always reduced, as two are replaced by one (Laser-Scan, 2001).

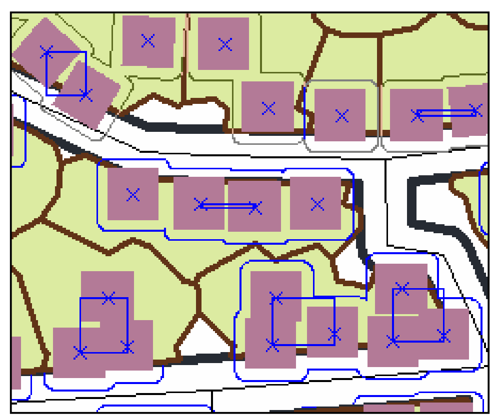

Figure 3. Clustering of buildings using dendrograms

- Then, they are typified using mean points (average coordinates) (Figure 4). After typification, building symbols are rotated parallelly to the nearest road. In this case they may be close to each others. To prevent this, nearest distances among building polygons instead of points need to be considered, however this can possibly get the strategy more difficult.

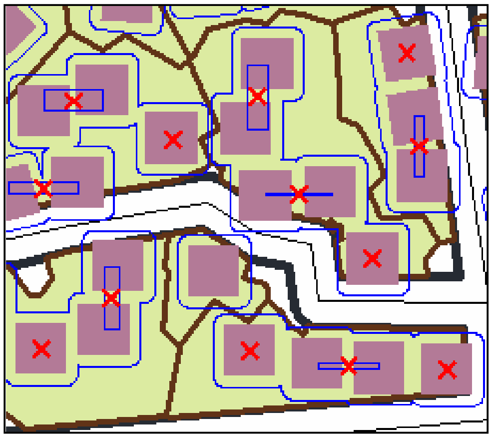

Figure 4. Typified buildings using dendrogram.

- Displacement is first done by computing the perpendicular distance to nearest road from building with point geometry and the orientation using these two points if the building is very close to the nearest road regarding symbology and minimum building-to-road distance threshold (10 m – 1:50K). In case of polygon geometry, nearest vertex of polygon is used and no rotation is applied (Figure 5). Namely, geometrically unimportant buildings (square shapes) are rotated because they are usually derived from a few buildings. If buildings are still very close to roads (e.g. buildings at the corner of blocks) second displacement is done. We can also displace buildings at the same cluster together with the one in highest conflict to preserve relative locations. Besides, we have to displace the buildings in conflict with each other. In this case their intersecting buffers can be used and they can be displaced perpendicular to intersection points with semi-distance of conflict. If buildings have a few conflicts with neighbouring buildings, combined displacement vectors need to be calculated. This strategy is still in development phase.

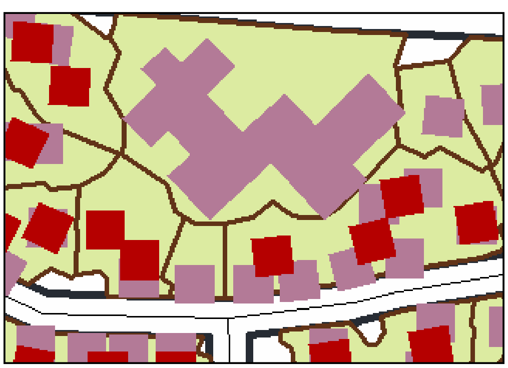

Figure 5. Displacement from roads after first iteration.

4. CONCLUSIONS AND FUTURE WORKS

In this paper, we present a case study for generalization of roads and settlement areas from 1:25K to 1:50K. Generalization is a contextual operation and requires object and inter-objects level characteristics to be made explicit as much as possible. For this purpose, we created blocks among surrounding roads and created voronoi diagrams (polygons) within block. Thus, we caught the chance of local analysis within the block and local decisions within each voronoi polygon to apply generalization operations optimally. Besides, object-oriented GIS gave the chance of dynamically computing the characteristics of buildings, voronoi polygons and blocks. We defined missing rules and the parameters experimentally. First results are close to solution after visual checking although some editing is required. Next stage of the study will be further develop generalization strategy as we mentioned above and also evaluation strategy can be developed using characterisations before and after generalization. Long transaction mechanism of object-oriented database to backtrack in case of bad generalization will be considered as well.

REFERENCES

AGENT Cons., 1998. Constraint Analysis, Project Report D A2, ESPRIT/LTR/24 939.

Anders, K.-H. and Sester, M., 2000. Parameter-Free Cluster Detection in Spatial Databases and its Application to Typification. In*: IAPRS*, Vol. 33, ISPRS, Amsterdam.

Jiang, B. and Claramunt, C., 2002. A Structural Approach to Model Generalisation of an Urban Street Network. In: *Proceedings of 5th AGILE Conference on Geographic Information Science*, Mallorca, Spain.

Brassel, K.E. and Weibel, R. 1988. A Review and Conceptual Framework of Automated Map Generalization. *International Journal of Geographic Information Systems*, 2(3), pp.229-244.

GCM, 2002, *Specification for 1:25K, 1:50K and 1:100K Topographic Maps*, General Command of Mapping (GCM), Ankara, Turkey.

Jones, B.C., 1997. *Geographic Information Systems and Computer Cartography*. Harlow, UK: Addison Wesley Longman, 311 p.

Laser-Scan, 2001. *The Gothic Module Reference Manual*. Online Version, November 2001, Laser-Scan Ltd., Cambridge, UK.

Mackaness W., and Ruas A., 1997. Strategies for Urban Map Generalization. In: *Proceedings of 18th ICA International Cartographic Conference,* Stockholm, Vol.3, pp. 1387–1394.

Müller, J.-C., 1990. Rule Based Generalization: Potentials and Impediments. In: *Proceedings of 4th International Symposium on Spatial Data Handling*, Vol. 1, pp. 317-334.

Ormsby D. and Mackaness W., 1999. The Development of Phenomenological Generalization Within an Object-oriented Paradigm. *Cartography and Geographic Information Science*, 26(1), pp. 70-80.

Plazanet, C., Bigolin, N.M., Ruas, A., 1998. Experiments with Learning Techniques for Spatial Model Enrichment and Line Generalization. *GeoInformatica*, 2(4), pp. 315–333.

Regnauld, N., 1996. Recognition of Building Clusters for Generalization. *Proceedings of the 7th International*

Symposium on Spatial Data Handling, Vol. 2, Session 4B, Delft, The Netherlands.

Regnauld, N., Edwardes, A., Barrault, M., 1999. Strategies in Building Generalization: Modelling the Sequence, Constraining the Choice. *3rd ICA Workshop on Progress in Automated Map Generalization*, Ottawa.

Ruas, A., 1998a. A Method for building Displacement in Automated Map Generalization. *International Journal of Geographical Information Science*, 12(8), pp. 789-804.

Ruas, A., 1998b. O-O Constraints Modelling to Automate Urban Generalisation Process. In: *Proceedings of 8th International Symposium on Spatial Data Handling*, Vancouver, pp. 225-235.

Sester, 2000. Knowledge Acquisition for the Automatic Interpretation of Spatial Data. *International Journal of Geographical Information Science*, 14 (1), pp. 1-24

Thomson, R. C., and Richardson, D. E. 1999. The 'Good Continuation' Principle of Perceptual Organization Applied to the Generalization of Road Networks. In: *Proceedings of the 19th ICA International Cartographic Conference,* Ottawa, pp. 1215–1223.

Weibel, R, 1995. Three Essential Building Blocks for Automated Generalization. In: Muller, J.-C., Lagrange, J.-P., Weibel, R. (eds.), *GIS and Generalization: Methodology and Practice*, London, UK: Taylor&Francis, pp. 56-69.

Weibel, R., 1997. Generalization of Spatial Data – Principles and Selected Algorithms. In: Van Kreveld, M., Nievergelt, J., Ross, Th., Widmayer, P. (eds.) *Algorithmic Foundations of GIS*. Lecture Notes in Computer Science, Berlin: Springer-Verlag, Vol. 1340, pp. 99-152.

Weibel, R. and Dutton, G., 1999. Generalizing Spatial Data and Dealing with Multiple Representations. In: Longley, P., M.F.Goodchild, D.J.Maguire and D.W.Rhind, (eds.) *Geographical Information Systems: Principles, Techniques, Management and Applications*, 2nd Edition, Vol.1, Cambridge: GeoInformation International, pp. 125-155.

ACKNOWLEDGEMENTS

This research has been supported by Yildiz Technical University Scientific Projects Coordination Department. Project Number: 22-05-03-01. We are grateful to Laser-Scan Ltd., Cambridge, UK for both software supply and technical support, and GCM Turkey for data supply.