Hochschule Karlsruhe (HSKA) - University of Applied Sciences
Geomatics Fakultät für Geomatik

Kandidat

Dimo Dimov

Masterthesis (Jahr: 2012)

Automated Thematic Map Generalization

Referent
Prof. Dr. Heinz Saler (HSKA) - Prof. Dr. Luis Ruiz und Prof. Dr. Palomar Vazquez (UPV)

Keywords
Generalization, Simplification, Aggregation, Skeletonization, Image processing, GIS, ArcObjects, Visual Basic

Zusammenfassung

Cartographic and geodatabase generalization is an essential technique in map production and geovisualization to create new maps with a reduced level of detail. Due to the current development and higher availability of spatial information in geodatabases as well as the increasing demand of data, maps and geoinformation it is of high necessity to develop tools and techniques for automatic generalization. Hence, manual methods are time consuming and not cost-effective due to the high amount of cartographers who are involved in the mapping process.

The Thesis is focused on the automatic generalization of thematic datasets, more precisely: continuous polygonal and gapless mosaics such as landcover or geologic maps. The solution is based on heuristics and mixed-integer programming and the process is more complex than the generalization of networks or topographic maps, due to the adjacency of the objects, topologic consistency and the semantic accuracy.

This work deals with skeletonization, morphological filtering, image and vector processing, polygon aggregation, line simplification and topology. Additionally it will include the aggregation of separate mixed area parcels to determine the major landuse type per parcel and to mosaic mixed parcels with a describing mixed-type attribute.

National land cover datasets like SIOSE (Landcover information system of Spain) or European projects such as CORINE with huge amounts of data and geographic objects in a resolution, demand fast and cost-effective methods of map and database generalization in order to derive new information on smaller scales without or with much less user interaction. Manual generalization would make the production of very large amounts of maps, geodatabases and web services too slow and too expensive. Therefore it is essential to automate the generalization process, to make it stable for any scale step, for any landcover type and geographical region and to provide automated conflict resolving algorithms that always appear, when we try to transform data into a smaller scale, and also develop algorithms that preserve the topology and semantic similarity of the objects.

The goal of this Thesis is to come up with an automated workflow for a landcover dataset in the scale of 1:1.000 that was derived from multispectral airborne imagery, to generalize it to the target scale of 1:25.000 which is the same for the manually produced SIOSE landcover dataset. At the end of this Thesis it is necessary to compare SIOSE with the results from the automated process.

1. **Target scale requirements**

The generalization of objects to any scale requires hard constraints and thresholds that must be fulfilled. The process often requires the elimination of polygons that are below a certain threshold that define the minimum allowed size of a feature or the removal of abundant bends in polylines or the combination of small polygons to a larger one. Each change in the map scenery is leading to a new interpretation of the information. Too many changes may lead to misinterpretation and affect the geographic or topologic accuracy of the data negatively. The question that occurs is the following: "How to reduce the level of complexity and in the same time to avoid wrong map interpretation?"

Geometric requirements of the target scale of 1:25.000:

– min. allowed size: 5000m²

– min. object width: 15m

– min. object length: 60m

– min. distance between two objects: 15m

Therefore most of the secondary and tertiary roads cannot meet the geometric requirements and hence must be eliminated, otherwise the road network will not be legible at all and will disturb the map scenery as a whole. The narrowness of roads is the major factor for elimination.

2. **Skeletonization of road polygons**

This Thesis discusses 4 methods to derive the skeleton of a polygon:

- **Euclidean Distance Skeleton**
- **Euclidean Direction Skeleton**
- **Delaunay Triangulation (TIN)**
- **MATLAB (morphologic filters)**

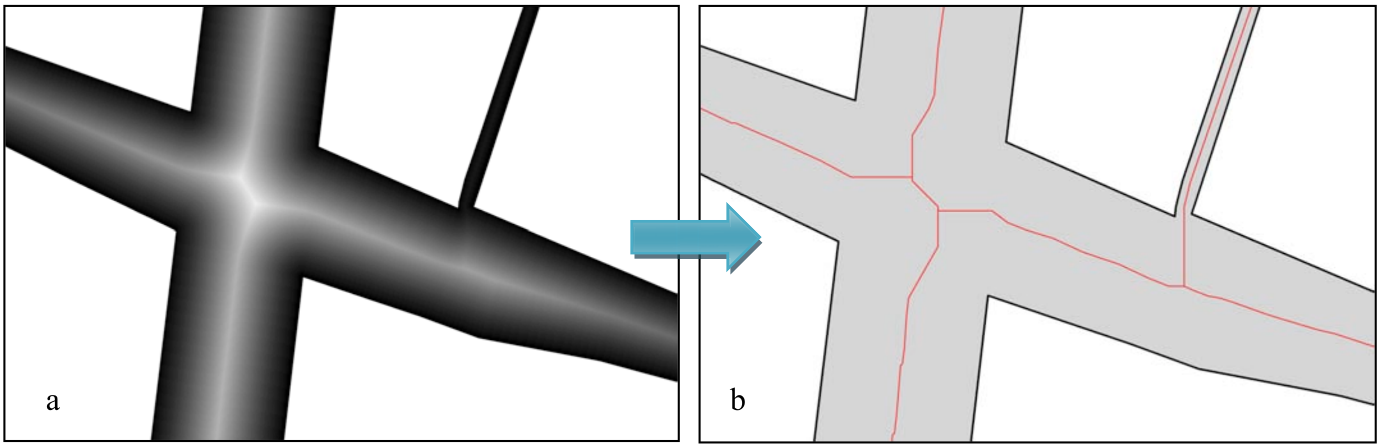

Figure 1: a) Euclidean Distance raster and b) derived skeleton: pixels with the maximum distance to the polygon boundary are extracted. The connected pixels are converted to a polyline that describes the skeleton.

3. Split of road polygons

The collapsed road polygon is represented through its skeleton (polyline). The gaps between the skeleton and the landcover parcels must be filled by splitting the road polygon into segments and aggregating each segment to a corresponding landcover parcel.

The split is performed by connecting the landcover polygon vertices that are incident for adjacent polygons with the nearest skeleton vertice.

Figure 2: a) road polygon (red) as a candidate for collapse b) skeleton c) split into segments d) aggregation e) dissolve

4. **Landcover aggregation**

The aggregation of adjacent polygons of different classes is performed with the implementation of:

Geometric constraints: size
Semantic constraints: semantic similarity of landcover classes
Topologic constraints: shared boundary of adjacent polygons, centroid distances
Compactness: ratio of the polygon´s perimeter and a circle with the same perimeter

The implementation of single constraints causes uncontrolled aggregation: a polygon with a size < 5000m² maybe aggregated with a larger polygon with a size > 5000m², but both polygons may be semantically very different (semantic distance): e.g. the aggregation of an agricultural polygon with an urban polygon has a high semantic distance and has to be avoided, whereas the aggregation of a rice polygon with a fruits polygon is semantically more similar as both are agricultural parcels.

Therefore the constraints must be combined and processed iteratively. Results differ a lot depending on the implementation of hard thresholds. Simulated annealing (iteratively minimization of thresholds, if aggregation is not achievable) has to be performed in some cases.

Figure 3: a) original landcover with roads b) collapsed roads c) geometric aggregation d) semantic aggregation e) topologic aggregation