

MAP GENERALIZATION BY SKELETON RETRACTION

Christopher Gold

Dept. Land Surveying and Geo-Informatics
Hong Kong Polytechnic University
Hung Hom, Kowloon, Hong Kong SAR, PRC
(ChristopherGold@Voronoi.Com)

David Thibault

Centre for Research in Geomatics
Laval University
Quebec City, Quebec, Canada

Abstract

Much work has been done in the last few years on various aspects of automated map generalization. However, many of these efforts have been directed to the generalization of individual objects - houses, roads, etc. - while ignoring the spatial relationships between the objects themselves. This problem is particularly apparent where map generalization includes the simplification of contour lines (which may overlap as a result, if no spatial relationships are preserved), or the question of preserving the relationships between different linear features (roads, rivers, railways) that must be cartographically displayed (even if displaced) in the smaller-scale maps after generalization.

We here propose a new generalization method based on the medial axis transform, or skeleton. This new technique, as opposed to other traditional generalization methods, has the advantage of preserving the topological structure of the object, which eliminates any risk of overlapping of neighbouring objects. It is thus especially appropriate for the generalization of contour lines, among other things, where there may be high density and proximity of linear elements.

Introduction

Cartographic generalization techniques have been discussed for many years, often under the headings of simplification, exaggeration, elimination and displacement. The early work of Douglas and Peucker (1973) on line generalization was very successful, but as stated by Weibel (1995) it can not be considered to be true "Model generalization". He suggested that a valid method should produce predictable and repeatable results; should minimize deviations of the resulting model from the original model; should maximize data volume reduction; should preserve topological consistency; should use as few control parameters as possible; and should be efficient. He also noted that model generalization is traditionally applied to discrete objects but is just as relevant to continuous surfaces. Other relevant papers on this topic include Weibel (1992), McMaster (1989), Brassel and Weibel (1988) and Bundy et al. (1995). This last is of interest as it uses a topological structure to manage aspects of generalization - in particular the displacement and merging of cartographic objects, and some forms of skeletonization. The approach developed in the current paper is addressed initially towards line and object generalization (simplification), but it is clear that further work could lead to methods for displacement and merging, among others. Work on these is ongoing.

The problem of line generalization is especially interesting for workers in GIS and automated cartography, as there is a frequent need to reduce the complexity of linear elements in order to represent them at a smaller scale with reduced detail. Weibel (1995) stated that "Another reason to reduce the accuracy and resolution of a data set is the homogenization of different data sets in the process of data integration". Various line generalization techniques have been developed, such as that of Douglas and Peucker (1973), but these often cause problems when the generalized elements are very close, as they may then overlap or intersect, due to the absence of any topological structure linking adjacent objects. Our recent work has used another approach, based on the skeletons of map objects - both the interior endoskeleton, and the exterior exoskeleton.

The objectives of this work are to develop techniques for detail reduction without permitting object overlap - this presupposes a topological structure. In addition, we assume that the input data points may have any

position in the Euclidean plane - while input may be from raster images in some cases, there is no supposition that points occupy pixel locations. In addition, we wish no tolerance value or other parameter involved in the process: it is completely scale free. We would like to show that we achieve some of the objectives listed in Weibel (1995) cited above.

The key idea applied here to generalization processes is the concept of *skeleton retraction*. The idea is quite simple - simpler objects have simpler skeletons which means simpler shapes (Blum 1967). (This is also true for the space between objects - retraction here simplifies the common Voronoi boundary between them.) This approach is only made possible because we can preserve the link between the skeleton and the boundary (or "crust"). This paper only describes some of the initial steps in this controlled retraction process - that of retracting each leaf of the skeleton tree to the same location as its parent. This provides a skeleton with fewer branches, and a crust/boundary with fewer minor perturbations. More advanced retraction procedures are possible given certain knowledge rules about the desired shapes of the objects concerned. Some of these issues are discussed in Ogniewicz (1994) and Ogniewicz and Ilg (1990).

Previous work

The work described here is a continuation of our previous efforts (Gold et al., 1996; Gold, 1997; Gold, 1998) on data input. Briefly, in the first paper we developed a digitizing technique for forest polygons based on digitizing around the interiors of each polygon, giving the points the polygon label, creating the Voronoi diagram, and extracting the Voronoi boundaries between points with different labels. (Okabe et al., 1992, give an excellent summary of Voronoi algorithms and applications.) This was rapid, and guaranteed to be topologically correct. In the second paper we extended this to scanned maps, by deriving "fringe" points on the interior of each polygon by image filtering, labelling these points with a floodfill algorithm, and extracting the polygon boundaries as just mentioned. In the third paper we show that the work of Guibas and Stolfi (1985), in developing the Quad-Edge structure, may be extended to connected maps by permitting the simple edge between two vertices to become an arc (the "Quad-Arc"), with multiple points along it.

The properties of the Quad-Edge/Quad-Arc led us to examine our previous techniques. The Quad-Arc allows the management of any connected graph on the plane or orientable manifold. However, our earlier algorithms required closed polygons, with different labels, in order to identify the edges we wished to preserve. Thus unclosed line work would be lost, as the polygon label would be the same on both sides of the line. However, we could not find a satisfactory algorithm based on point distribution alone, without some labelling function.

In 1998 Amenta et al. published a seminal article where they showed that the connected set of boundary points (the "crust") could be extracted using the Voronoi diagram. Their objective was to eliminate all Delaunay edges that crossed the skeleton of the polygonal object under consideration, with the remaining Delaunay edges forming the crust. They achieved this by first generating the Voronoi diagram of the boundary points, and exporting the Voronoi vertices. In a second step, they constructed the Delaunay triangulation of both the original boundary points and the new Voronoi vertices. Any Delaunay edge connecting two boundary points must be part of the desired crust, as Delaunay edges that originally crossed from one side of the polygon to the other would now be cut by the newly-inserted "skeleton" points formed from the original Voronoi vertices. They showed that if the samples along the boundary were not more than (approximately) 0.25 times the distance to the skeleton, it was guaranteed that the complete crust (and no other edges) would be preserved.

However, our previous work had used the *skeleton* (a subset of the Voronoi edges) to define our topologically complete polygon boundaries, and not the *crust*. In Gold (1999) and Gold and Snoeyink (2001) we showed that the two steps were unnecessary - a local test of associated Voronoi/Delaunay edge pairs sufficed. Our previously-used Quad-Edge structure contained this information directly. The work of Amenta et al. reduced to a simple test on each Delaunay edge of the original Voronoi/Delaunay diagram: was there an empty circle touching the two Delaunay vertices that did not contain any Voronoi vertices? If so, then this was a crust segment. Otherwise it was a skeleton segment. Thus it was possible to generate both the crust and the skeleton in one step, as a simple test on each Voronoi/Delaunay edge pair. Essentially, each edge pair (stored as the Quad-Edge structure of Guibas and Stolfi (1985)) would be assigned either to the crust or the skeleton.

The use of this skeleton-based structure has been suggested in Ogniewicz (1994), and Ogniewicz and Ilg (1990), for simplifying individual objects, using fairly elaborate tree-pruning techniques. The difference between our work and theirs is that they were working with raster data, and were only concerned with the skeleton. In addition, they used a tolerance figure to determine how much to prune the skeleton. Our work does not have these constraints. Figure 1 shows the outline of a maple leaf - the crust and the skeleton are clearly distinguishable. However, this is only a graphical display, where each Voronoi/Delaunay edge pair is drawn as either the crust or the skeleton.

Advantages of the Quad-Edge structure

In order to use our skeleton-based generalization algorithm, we needed a structure that was capable of managing both the primal and the dual graph with equal felicity, and that could apply to simple Delaunay/Voronoi structures as well as to polygon arcs. Since we were often working interchangeably with Delaunay/Voronoi representations, the Quad-Edge data structure of Guibas and Stolfi (1985) became an obvious candidate. It has various attractions: it is a method for representing the edge connectedness of any connected graph on a manifold; it is symmetric in its storage of both the primal and the dual edge; it has a complete algebra associated with it, requiring no searches around polygons, nodes, etc.; and it is admirably adapted to an object-oriented programming environment. There are only two operations: Make-Edge to create a new edge, and Splice to connect or disconnect two ends. Gold (1998) shows a half-page of code that implements the basic operations. To us it appears the most elegant structure.

The original motivation for this work was to take advantage of our scanned mapping technique. This started with the complete Voronoi structure, which guarantees topological completeness of the processed map, but uses large amounts of storage. Using the Quad-Edge representation of the Voronoi structure permits the elimination of unwanted edges, and the Quad-Edge to Quad-Arc simplification provides even greater storage efficiency. The resulting structure is then equivalent to other polygon-management data structures, in that if there are many intermediate points along the arcs, then the number of pointers contributes little to the storage costs. The resulting data structure cleanly separates the topology from the attributes, in that the pointer structure is entirely within the topological elements: four pointers between the branches of each Quad-Arc in the object-oriented version, four pointers to adjacent Quad-Arcs, and four pointers to attribute information. These are traditionally the two faces and the two nodes associated with the Quad-Edge, but in the Quad-Arc the face pointers point directly to polygon attribute information, perhaps in a database, and the two node pointers may be used to reference the associated arc (chain of x-y coordinates) or other arc information as desired.

The Quad-Edge or Quad-Arc data structure permits topology maintenance in a mapping system with a minimum of complexity. It fits very easily with our current scanned map software using the Voronoi model for data input and topological structuring, and permits straightforward topology maintenance and spatial queries after simplification. Most of all, it is simple to implement in an object-oriented environment, reducing some of the complexity of cartographic topology.

Skeleton-based generalization algorithm

One issue that needs improvement is the existence of spurious "hairs" on the skeletons generated. This is a well-known artifact of skeleton generation, where any irregularities in the boundary generate unwanted skeleton branches. Ogniewicz attempted to reduce skeletons formed from raster boundary points to a simple form by pruning the leaf nodes of the skeleton until a specified minimum circumcircle was achieved, but with the development of the one-step crust and skeleton algorithm this process may be greatly simplified.

Alt and Schwartzkopf (1995), as well as Blum (1967) showed that leaf nodes of a skeleton correspond to locations of minimum curvature on the boundary. For a sampled boundary curve this means that three adjacent sample points are cocircular, with their centre at the skeleton leaf. If we wish to simplify the skeleton we should retract leaf nodes to their parent node location. This means that we now have four cocircular points instead of three. The retraction is performed by taking the central point of the three defining the leaf node, and moving it towards the parent node of the skeleton until it meets the parent node

circumcircle. This smooths outward-pointing salients in the boundary of the object. The same should be done from the other side of the boundary, retracting those salients also. This may displace some of the points involved in the first smoothing step, but as the process is convergent a small number of iterations suffices to produce a smoothed curve having the same number of points as the original, but with a simplified skeleton. Figure 2 shows the result of smoothing the maple leaf of Figure 1.

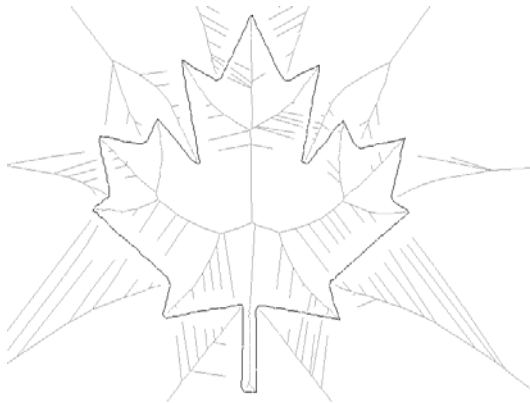


Figure 1: Crust and skeleton of a maple leaf.



Figure 2: Smoothed version of Figure 1.

This process is interesting for several reasons. Firstly, we have retracted the leaf nodes of the skeleton, greatly reducing its complexity. This then gives us a resulting skeleton closer to Blum's idea of the Medial Axis Transform, but without the artifacts due to perturbation of the samples along the "ideal" boundary. Blum's motivation was to provide stable descriptors of shape for "biological" objects, and the skeletons we have derived are well suited to that. However, retraction of skeleton leaves may be appropriate in the middle of a curve, but inappropriate at, for example, the corner of a building. Such boundary points would need to be flagged prior to the smoothing operation. Note that this smoothing process has no metric tolerance associated with it - so there is no problem with scale, etc. Convergence is based on cocircularity. (This is particularly useful in the case of contour maps, as slope is perpendicular to the contour segments, so drainage is always towards skeleton nodes.)

Secondly, the skeleton may act as a link between adjacent objects - for example between adjacent curves of a contour map. The retraction process can not cause one curve to cross another (as can happen with simple curve smoothing), because the two crusts must have a separating skeleton. We thus get simplified curves that fit together. This is completely general, and provides a new model for map generalization.

Thirdly, because the skeleton is a linking structure between the curves, we may use this property to help us when we are doing cartographic generalization involving the exaggeration of some feature and the displacement of others. Displacement of one boundary of a feature means that the next feature may also need to be displaced, and the crust/skeleton test indicates the minimum necessary distance. We have not yet pursued this idea in detail.

Applications of the skeleton-based generalization algorithm

The algorithm has been tested on a variety of data sets: a contour map of Mont Albert, (Québec, Canada), a polygonal map, a river network and scanned maps. Originally in vector format, the data were rasterized to give a single line of pixels, except for the case of the scanned urban map, where an edge-detection filter was used. The Voronoi diagram/Delaunay triangulation of these points was then constructed using the Quad-Edge data structure of Guibas and Stolfi (1985), in order to extract the crust and skeleton with our algorithm.

Contour map generalization

The results obtained for the contour map were very satisfactory (Figures 3 and 4). For sufficiently well-sampled lines the results were excellent: almost all the secondary branches (the “hairs”) were eliminated, giving smoother curves without the previous minor perturbations. This gives both a cleaner-looking contour map and a smoother terrain model derived from these contours.



Figure 3: Small contour map and skeleton.

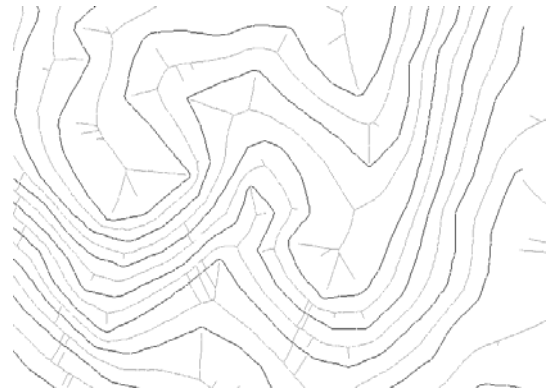


Figure 4: Smoothed version of Figure 3.

The remaining secondary branches of the skeleton are valuable as they indicate the probable intermediate form of the surface between contours - such as minor ridges and valleys. Derived where there are re-entrants in a single contour, they approximate the human interpretation of the surface form, and can serve to improve runoff modelling. Problems may exist where the skeleton breaks through the crust, due to the sampling conditions of Amenta et al. (1998) not being maintained, especially at sharp corners. Additional filtering procedures are being developed to improve this situation. To get the best generalization, the algorithm must make several iterations, until no further points are displaced. In many cases a single iteration is sufficient, but up to five have been needed on occasion.

Polygon generalization

Polygon generalization also gives excellent results, as shown in Figures 5 to 8. Our experiments show that the form of the polygon has been preserved, and the skeleton simplified - but both are recognizably similar to the ungeneralized version. This conforms to the original concept of the Medial Axis Transform as a shape descriptor, as originally defined by Blum in 1967.

The current algorithm does not reduce the number of points - it merely simplifies the form of the curves. However, point reduction is made much simpler with the generalized curves. Research is ongoing on this topic. While excellent for individual polygons, generalization of connected polygon sets can rupture



Figure 5: Simple polygon and skeleton.



Figure 6: One smoothing iteration.



Figure 7: Two smoothing iterations.

the crust if the node at the junction of three arcs is displaced as required by the algorithm. A method needs to be introduced to flag such nodes as immobile. A polygon map is shown in Figure 5. Figure 6 shows the results of one pass of the retraction algorithm, and Figure 7 shows the results of a second pass. No significant changes were observed after two passes. It can be seen how the retraction of the skeleton leaves retracts the associated crust segments. Figure 8 shows Figures 5 and 7 superimposed. It can be clearly seen that the form of the polygon is preserved. While we have not yet attempted it, further generalization beyond the removal of “first-order” skeleton branches is clearly possible. Indeed, it appears that the only “parameter” required by the skeleton retraction method would be how many orders of the skeleton to retract.



Figure 8: Figures 5 and 7 superimposed.

River Networks

Referring to “first order” and “second order” skeletons clearly shows the analogy with Strahler numbers for stream networks. Examining the maple leaf in Figures 1 and 2, the first-order skeleton branches are associated with minor perturbations in the boundary data points. Removal of the first-order “hairs” results in a skeleton that corresponds to the expected vein structure. If, by a stretch of the imagination, we considered our maple leaf boundary to be a watershed, then the skeleton shows the expected drainage pattern that would develop in homogeneous terrain. This is based on the fact that the line of maximum slope is presumed to be perpendicular to the boundary - and a Voronoi edge is always perpendicular to its associated Delaunay edge. (The same holds true for the contour map - in the crust identification algorithm, the dual Voronoi edges are suppressed. These, however, are perpendicular to the crust, and thus give the line of steepest descent.)

Figure 9 illustrates the case where the input data form a hydrological network. The Voronoi edges perpendicular to crust segments may thus be imagined to give the “uphill” direction. (This conforms to Blum’s (1967) concept of associating a height proportional to the distance from the boundary.) The skeleton thus gives the ridge - or valley - at maximum distance/height from two or more boundaries. We thus have a good estimate of the watersheds associated with each river segment. In Figure 10 we have retracted the first-order ridges/skeletons, and in so doing have simplified both the ridge and the valley forms. (Individual points have been added manually to overcome the sampling problems at junctions that were described previously.) Clearly we have produced a good generalization of Figure 9, suitable for use at a smaller mapping scale.



Figure 9: Hydrological network and skeleton.



Figure 10: Smoothed map of Figure 9.

Scanned maps

Our original interest was in the automated extraction of topologically complete maps from manual or scanned input. In this case we presume that we are generating thick line work, and we use an edge detection filter to separate the black pixels from the white ones. Our primary interest is thus more in the extraction of the skeleton of the line work than in the skeletons of the polygon interiors. Gold (1997) described an approach for processing scanned maps, using the algorithm of Gold et al. (1996). The adoption of the crust concept of Amenta et al. (1998) and the one-step algorithm of Gold (1999) has greatly simplified the extraction of topology from scanned images of line work, especially where this is not in the form of closed polygons. We use an edge detection filter to obtain the boundaries between the black space and the white space, and then extract the crusts and skeletons of the various map components. Figure 11 shows a scanned cadastral map, and Figure 12 shows the generalization. Note that there are two types of skeleton here: the “black” one, forming the centreline of the line work, and generating the topological network that may be used in GIS applications, or with our Quad-Arc approach; and the “white” ones, giving the form of the interiors of the individual polygons. The generalization process has mostly affected these white skeletons, giving a simpler structure to the shape descriptors envisaged by Blum. The text has also been processed in the same way, and may be extracted using the ideas of Burge and Monagan (1995). This would permit the one-step extraction of the text and its association with a particular polygon. Since the primary objective in this application is to extract the appropriate topology, the Quad-Arc structure is used to collect this and export it.



Figure 11: A small cadastral map.



Figure 12: A smoothed map of Figure 11.

Conclusion and future work

This article presents a simple and effective method for generalizing cartographic lines, for many types of features. It is of interest because the topological relationships between the various elements are preserved throughout - although attention must be paid to the required sampling criterion. The Quad-Edge structure was used because it is explicitly defined to include primal and dual relationships simultaneously, and it may readily be extended to produce the Quad-Arc as an output data structure. It provides a simple data structure that suits the needs of the data input process, preserves the map topology, and may be implemented within small-scale PC-based software for topological map querying. The combination of the crust criterion, the one-step algorithm, the Quad-Arc data structure and skeleton retraction generalization are sufficient to put topological cartographic generalization within the range of simple mapping programs. Future work includes flagging immovable points, such as building corners and nodes at the junctions of several arcs, second-order reduction, data point reduction after smoothing, and application of the generalized contour maps to terrain modelling.

Acknowledgments

Support for this research was received from the Natural Sciences and Engineering Research Council of Canada and the Networks of Centres of Excellence Program, under the IRIS network, and from the Industrial Research Chair in Geomatics at Laval University.

References

- Alt, H. and Schwarzkopf, O., 1995. The Voronoi diagram of curved objects. Proc. 11th Annual ACM Symposium on Computational Geometry, pp. 89-97.
- Amenta, N., Bern, M. and Eppstein, D., 1998. The crust and the beta-skeleton: combinatorial curve reconstruction. *Graphical Models and Image Processing*, 60/2:2, pp. 125-135.
- Blum, H., 1967. A transformation for extracting new descriptors of shape. In: W. Whaten Dunn (ed.), *Models for the Perception of Speech and Visual Form*. MIT Press, Cambridge, Mass., pp. 153-171.
- Brassel, K.E. and Weibel, R., 1988. A review and conceptual framework of automated map generalization. *International Journal of Geographical Information Systems*, v. 2, pp. 229-244.
- Bundy, G.L., Jones, C.B. and Furse, E., 1995. Holistic generalization of large-scale cartographic data. In: *GIS and Generalization - Methodology and Practice*. Edited by Muller, J.C., Lagrange J.-P. and Weibel, R. (London: Taylor and Francis Ltd.), pp. 106-119.
- Burge, M. and Monagan, G., 1995. Extracting Words and Multi-part Symbols in Graphics Rich Documents. Proc. 8th International Conference on Image Analysis and Processing, ICIAP, (San Remo, Italy), LNCS, Springer Verlag, pp. 533-538.
- Douglas, D.H. and Peucker, T.K., 1973. Algorithms for the reduction of the number of points required to represent a digitized ligne or its caricature. *The Canadian Cartographer*, v. 10 pp. 112-123.
- Gold, C.M., 1997. Simple topology generation from scanned maps. *Proceedings, Auto-Carto 13, ACM/ASPRS, Seattle, April*, pp. 337-346.
- Gold, C.M., 1998. The Quad-Arc data structure. In: Poiker, T.K. and Chrisman, N.R. (eds.). *Proceedings, 8th International Symposium on Spatial Data Handling; Vancouver, BC, 1998*. pp. 713-724.
- Gold, C.M., 1999. Crust and anti-crust: a one-step boundary and skeleton extraction algorithm. In *Proceedings of the ACM Conference on Computational Geometry*, Miami, Florida, (ACM) (in press).
- Gold, C.M., Nantel, J. and Yang, W., 1996. Outside-in: an alternative approach to forest map digitizing. *International Journal of Geographical Information Systems*, v. 10, no. 3, pp. 291-310.
- Gold, C.M. and Snoeyink, J., 2001, A one-step crust and skeleton extraction algorithm. *Algorithmica*, (in press)
- Guibas, L. and Stolfi, J., 1985. Primitives for the manipulation of general subdivisions and the computation of Voronoi diagrams. *Transactions on Graphics*, v. 4, pp. 74-123.
- McMaster, R.B., 1989. The integration of simplification and smoothing algorithms in line generalization. *Cartographica*, v. 26, pp. 101-121.
- Ogniewicz, R.L. 1994. Skeleton-space: a multiscale shape description combining region and boundary information. *Proceedings of Computer Vision and Pattern Recognition, 1994*, pp. 746-751.
- Ogniewicz, R. and Ilg, M., 1990. Proc. 4th International Symposium on Spatial Data Handling, v. 1, pp. 15-24.
- Okabe, A., Boots and K. Sugihara, 1992 *Spatial Tessellations - Concepts and Applications of Voronoi Diagrams*. John Wiley and Sons, Chichester.
- Weibel, R., 1992. Models and experiments for adaptive computer-assisted terrain generalization. *Cartography and GIS*, v. 19, pp. 133-153.
- Weibel, R., 1995. Three essential building blocks for automated generalization. In: *GIS and Generalization - Methodology and Practice*. Edited by Muller, J.C., Lagrange J.-P. and Weibel, R. (London: Taylor and Francis Ltd.), pp. 56-69.