A METHOD FOR CONSTRUCTION OF 2D HULL FOR GENERALIZED CARTOGRAPHIC REPRESENTATION

Jacqueleen JOUBRAN , Yair GABAY

Technion - Israel Institute of Technology, Haifa, Israel
Department of Civil Engineering
Division of Geodetic Engineering
jacquele@tx.technion.ac.il , cvryair@tx.technion.ac.il

KEY WORDS: Generalization, Data aggregation, 2D hull, Spatial data mining, Triangulated data structure.

ABSTRACT

Aggregating a group of features into one entity is a basic operation of the generalization process in Cartography. In the manual process, the operation is performed by drawing a circumscribing outline around a group of features sharing similar cartographic properties. This outline becomes the new feature itself, when presenting the data in a smaller-scale map. For instance, A group of point features such as trees, will be transformed into an entity of a planted area or woods, whereas, a group of features such as buildings will be transformed into an entity of a built area, etc. This paper presents a method of generating a 2D hull for circumscribing cartographic data, thus rendering the generalization process efficient. Generating the hull is performed via a computerized automated process based on self investigation of the data features. The goal is achieved by means of geometric analysis of the results obtained by applying constrained triangulation on the original cartographic data. The triangulation enables a geometric and topological investigation of the data features and assists in generating their outline, compatible with the requirements of the cartographic generalization. The process results can be adjusted according to the user's needs, by changing the parameters affecting the outline generation.

1 INTRODUCTION

The extensive use of computers significantly affects the materialization of cartographic tasks. The transition of presenting geographic data on a computer as a soft copy (digital data), as opposed to the hard copy which, used to be directly printed on paper, brought up the need to present them in different scales (Doytsher 1998). One approach, for example, regards the map objects as dynamic objects moving on the screen. The behavior of the objects is run by operators and computer applications, not necessarily having full control of their shape, as well as changes occurring in them due to change in the scale, for example (Timpf 1997). An important goal for the future is to attain performance of dynamically scale change processes of geographic data in different levels of abstraction. Achieving this goal involves operations of data generalization. For this reason the generalization, which has been an important and central issue in the domain of cartography, continues to maintain its prominent status at the era of computerized cartography.

The traditional cartography was based on the experience and expertise of the human cartographer in creating maps, as far as the various aspects of graphic presentation of information is concerned. The problem of shifting to computerized cartography is that it is difficult to make the computer replace the human visual and cognitive potential. Furthermore, it is necessary to adjust the existing applications to the computerized graphic presentation, so that they can serve for cartographic purposes, since most of them were developed by computer experts who lack the know-how and expertise of the cartographer (Peled 1998).

This paper deals with a method of finding a 2D hull to circumscribe geographic data, which represents general layout shape and defines the simple boundaries. The hull is a simple polygon, circumscribing the area containing all the original data. In some cases it may be advisable to perform the circumscription by applying a number of simple polygons, according to the feature scattering nature in the given information. The contained features can be of the type of points, lines or polygons. Constructing the hull reflects the geometric properties of the features and their interrelationships.

2 BACKGROUND

A widespread tool in computerized graphics for data circumscription in two-dimensional space is the "convex hull", which is a convex polygon with the minimum circumscribing area of a given group of points (Preparata & Shamos). Figure 1 demonstrates a convex hull of a group of points. In the course of time, a large number of algorithms were developed to define its shape, which is not affected by the method of definition. This hull has served for the purpose of a variety of cartographic applications, especially for generalization processes (McMaster 1996). The convex hull is not the optimal tool for circumscribing an area for cartographic purposes. The area defined by it is not sufficiently minimal, and includes many areas free of data. In many cases, the shape obtained does not adequately represent the geometric layout of the circumscribed information.

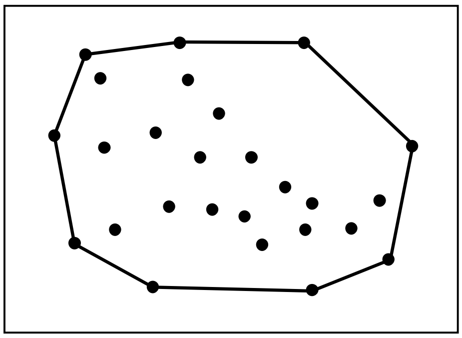

Figure 1. The convex hull of the group of points

2.1 Spatial data mining

Spatial data mining is a field which deals with producing new spatial data from existing spatial data. Data mining is facilitated by utilizing shapes or properties which are not explicitly expressed in the original data bases, and is performed without changing them (Kang 1997). Databases kept in GIS systems constitute a valid potential for data mining due to the vast and diversified data stored in them. The task dealt with in this paper, "automated definition of feature circumscription", is a good example of data mining. In order to attain this goal, ***Delaunay Triangulation*** was employed as an effective means of expressing the geometric properties of the objects. The triangulation allows production of a large amount of information about the data scattering, density and shape, although no explicit previous knowledge exists regarding these properties.

2.2 Delaunay Triangulation

Generally, triangulation of a planimetric surface is the method of dividing the surface into a finite number of triangles. This method is the key tool for handling problems whose solution is based on area division according to the principles of the finite element theory. Triangulation of 3D surfaces is an old work method in the topographic mapping domain. However, with the development of the digital cartography, the use of the 2D information triangulation has grown, among other things, for the purpose of investigating topological relations within that information.

There are a number of methods for performing triangulation, of which the Delaunay Triangulation is the preferred one for cartographic purposes, since it supplies triangles with especially short edges. This type of triangulation, according to Scewchuk (1996), "is based on the shortest distance criterion, defined by the condition claiming that the circumscribing circle for each triangle does not contain any point from the set of data". Saafeld (1993) mentions some properties that made the Delaunay Triangulation especially compatible with cartographic applications, which are:

- The principle of the circumscribing circle produces a situation in which each edge of the triangulation triangle edges joins between the pairs of points most adjacent to each other.
- Delaunay Triangulation is preserved after performing any transformation on a set of points.
- Delaunay Triangulation produces the maximum value of the smallest angle of all the angles in the triangulation. In cartographic problems, one usually strives to produce triangles closest to the shape of an equilateral triangle, since smaller angles lead to narrow triangles, which could pose an obstacle in the calculations.
- Delaunay Triangulation requires only local updating when adding or deleting a point, since only the local circumscribing circle, to which the point was added or deleted, must be handled.

2.3 Triangulated data structure

The triangulated data structure represents the distribution of the cartographic information after the triangulation by means of the smallest geometric unit in the two-dimensional space: triangles. A complete version of the triangulated data structure is presented in Jones (1995), where it is defined as SDS - Simplicial Data Structure. Such a data structure is designed to be applicable, containing rich information and efficient for cartographic operations, and it includes the following properties:

- Clear presentation in the two-dimensional space of the map.

- Accurate presentation of the object boundaries in the vector data structure.
- Rendering measurements easy.
- Maintaining topological links between points, lines and polygons.
- Facilitating easier local proximity relations between the objects.
- Easy to design and manage as well as flexible.
- Dynamic data structure.

3 METHOD OF PERFORMANCE

In the first stage a constrained Delaunay Triangulation is performed to force the polygon or line edges, describing the objects, to constitute part of the edges of the triangles. In order to utilize the triangulation results more efficiently, the data was arranged in a condensed version of the triangulated data structure, which integrates the data of the original objects together with the new triangle edges. The data structure applied in this study is demonstrated in figure 2. It is designed in such a way that it contains the following information:

- A list of objects.
- A list of points depicting the objects.
- Each point linked to the object, if it concerns lines or polygons.
- A list of triangles, which is a product of the triangulation.
- Each triangle is linked to three points, which are its three vertices.
- Each triangle is linked to its three adjacent triangles, which have one edge in common.
- A list of edges of the triangles.
- Each edge is linked to the two points which it joins.

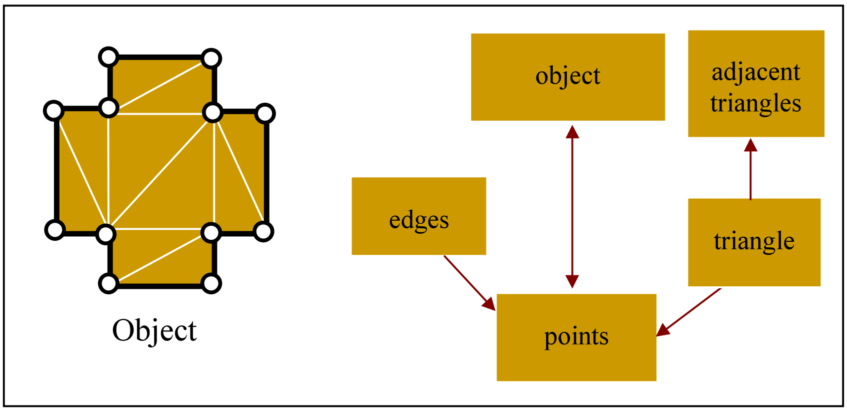

Figure 2. Object presentation in the triangulated data structure

3.1 Analysis of the data properties

The edges in the data structure are divided into three groups: (1) edges composing the original objects, (2) triangulation edges contained in the objects, and (3) triangulation edges stretched between the objects. A map consisting solely of polygon type objects, is composed of all the three groups of edges. A map consisting solely of line type objects, is composed of groups of edges (1) and (3). A map consisting solely of point type objects, is composed of a group of edges (3) only. The triangulation edge lengths of group (3) constitute an indicator of the data density in a certain region. These edges express the proximity relations between the objects. In a line map or a polygon map, the objects' density is also dependent of the objects' dimensions themselves. Therefore, in such cases, the ratio between the lengths of the object edges (groups 1 and 2) and the triangle edges in the intermediate space (group 3) is tested. This ratio allows an estimation of the data scattering in the various areas in the convex hull containing all the triangles formed by the triangulation. By analyzing this ratio, the length of the edges, depicting the low density or even empty space, will be determined. This value will define the tolerance for removing areas free of objects from the convex hull area, thus reducing its size until finding the desired hull.

3.2 Reduction of the convex hull

After determining the tolerance degree for the edge length between the remote objects which demonstrate areas free of data, the main stage of performance begins. At this stage the boundary line circumscribing the data is determined. After implementing the triangulation, as a first stage, the convex hull will be obtained, functioning as an external hull from which the empty “redundant” areas will have to be removed. This stage of removal reduces the convex hull to attain the desired one.

The reduction operation, which entails removal of empty areas from the convex hull, may create internal space in the hull, unless ensured otherwise. Since we are not interested in internal spaces, the process will be performed by traveling over the external edge only. This edge is defined by the external triangles. Such triangles have less than three adjacent triangles, since at least one of their neighbors is the external space. This property will assist in locating those triangles which lean on the external edge. Traveling over the external triangles to execute the reduction operation will result in removing redundant triangles from the external edge. A redundant triangle is a external triangle existing in the intermediate space, and contains at least one edge longer than the tolerance.

The process will be performed iteratively, until a state is attained when no triangle needs to be removed from the most updated external edge. Upon completing the process of removing the longer triangles with relatively empty areas of data, only a group of triangles, which cover the relevant area containing the data, will remain.

3.3 Constructing the desired hull

In order to obtain the desired hull a simple polygon circumscribing the group of triangles, left after the reduction operation, should be constructed. The process is performed by connecting the external edges of the external triangles sequentially.

3.4 The effect of the tolerance size on the obtained hull

The shape of the obtained hull is dependent on the tolerance that was determined by the properties of the geometric data. Following the statistical analysis a unified tolerance was set for the whole hull, although the density of the data along the hull is variable. Too large tolerance will leave free areas in the hull. Too small tolerance will cause the hull to assume an extremely complex shape. In the event that the original data are concentrated in a number of separate clusters, the right choice of tolerance will generate a number of separate hulls, as demonstrated in figure 3.

When the described method is implemented on a map of lines or points, too small tolerance will leave some of the objects outside the obtained hull. If these objects are at a larger distance, in relation to the other data, then this is the desirable result. Thus, if there is no cartographic justification to remove objects from the hull, the tolerance should be reduced so that the obtained hull will contain all the relevant objects.

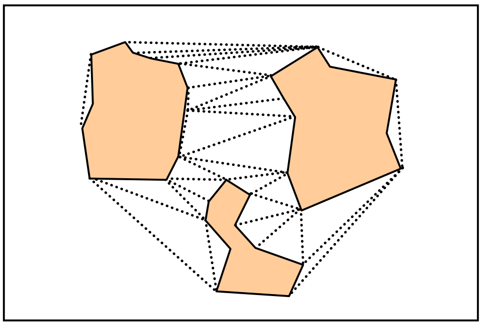

Figure 3. demonstrates the option for obtaining a number of hulls.

4 RESULTS

In this chapter the method of performance is demonstrated on a group of entities of polygon type. Following is a presentation of the results for group of entities of line and point types.

4.1 Polygon data

A map of buildings in a certain area is given, as depicted in figure 4. Each building is described as a closed polygon composed of a number of points. A constrained Delaunay Triangulation is implemented by forcing the building edges to constitute part of the triangle edges formed by the triangulation, as demonstrated in figure 5. The building edge lengths are analyzed against the triangle edges stretched through the intermediate area between the buildings, and the results are presented in figure 6.

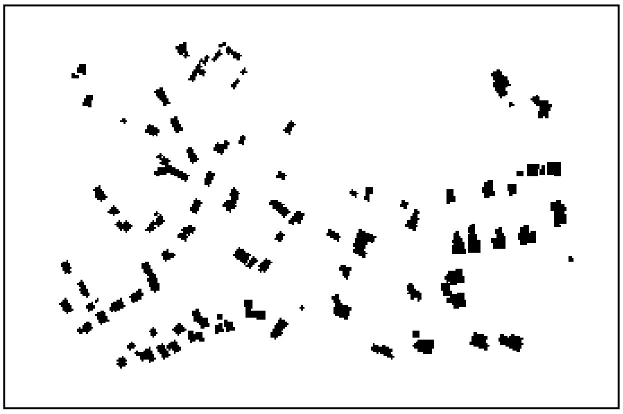

Figure 4. A layer of building data

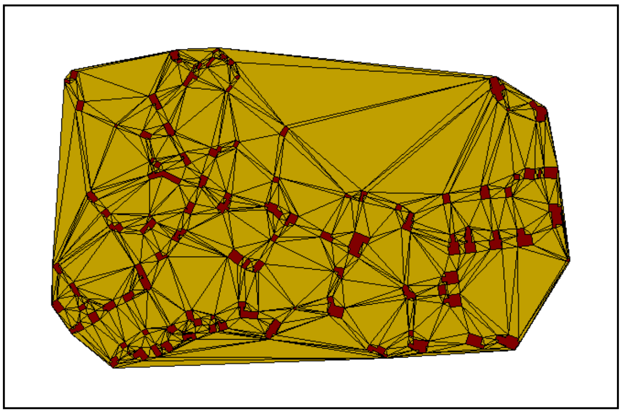

Figure 5. A constrained Delaunay Triangulation of the data

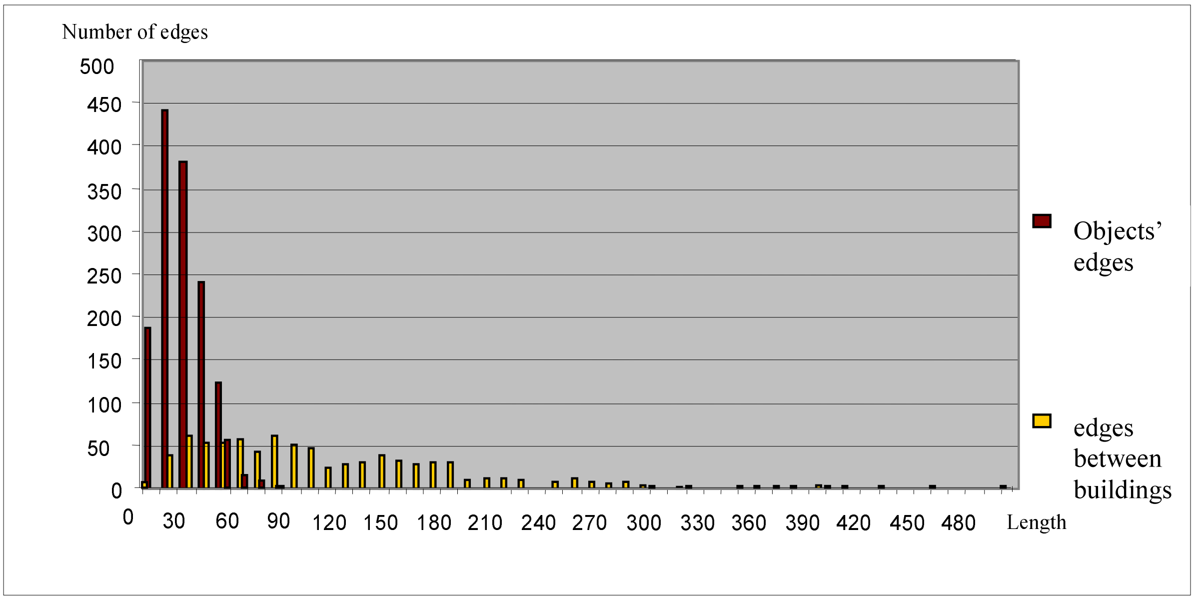

Figure 6. The common histogram for the object edge lengths against the intermediate space edges

Now, it can be observed that the intermediate space edges may be divided into three groups: (1) edge lengths of the same order of the building edge lengths, (2) slightly longer edges still depicting the acceptable relation between two adjacent buildings, and (3) lengths that are too long, thus demonstrating the relations between the buildings passing through more or less empty spaces.

According to this classification the edges are divided into three categories (figure 7), and respectively divide the triangles into three groups (figure 8). The first category for group (1) of edges whose lengths are between the values 0-220 m., the second category for group (2) of edges whose lengths are between the values of 220-320 m., and the third category for group (3) of edges whose values are larger than 320 m.

After determining the desirable tolerance value to be 320 meter, it is possible to start removing the external triangles which contain edges of that length order, or more. At the end of the process a 2D hull, which depicts the data layout, is obtained by applying the tolerance, as demonstrated in figure 9.

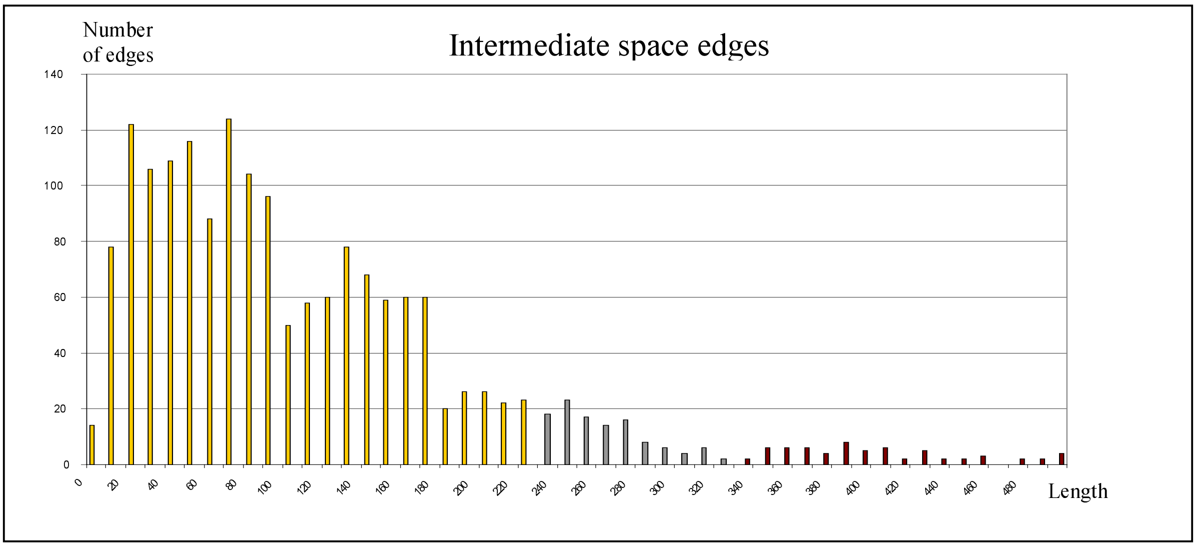

Figure 7. The histogram for the intermediate space edges divided according to the length category

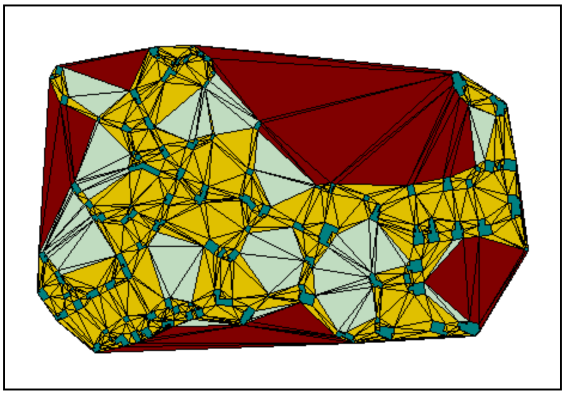

Figure 8. The division of the intermediate triangles according to the edge lengths category

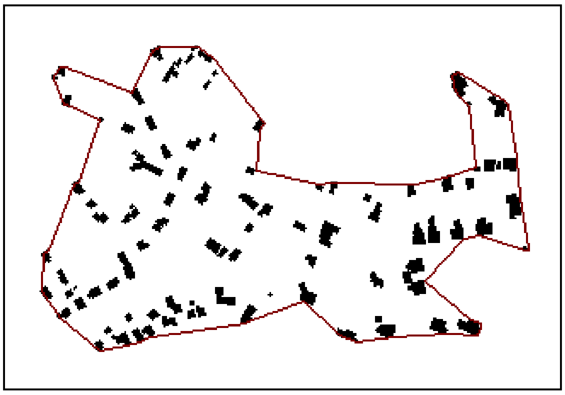

Figure 9. The circumscribing hull obtained from a 320 m. tolerance

4.2 Point and line data

If data of point or line type is handled, removing triangles in the reduction operation, may lead to reducing the hull in such a way, that it will no longer contain some of the original objects. In many cases this is the desired result, because these objects are remote from the rest of the data. In cases that it is preferred to produce a hull that will contain these objects, the solution is to increase the size of the tolerance.

Implementing this method on a map, composed of entities of the centroid point type of land plots in a certain area, will produce the result depicted in figure 10. It can be seen that two separate hulls were obtained with a small number of points outside the hull. If a single hull, which circumscribes all the data, is preferred, then the tolerance is increased, and the hull depicted in figure 11 is produced. Implementing the method on a map composed of entities of street line type in an urban area, produced the result depicted in figure 12.

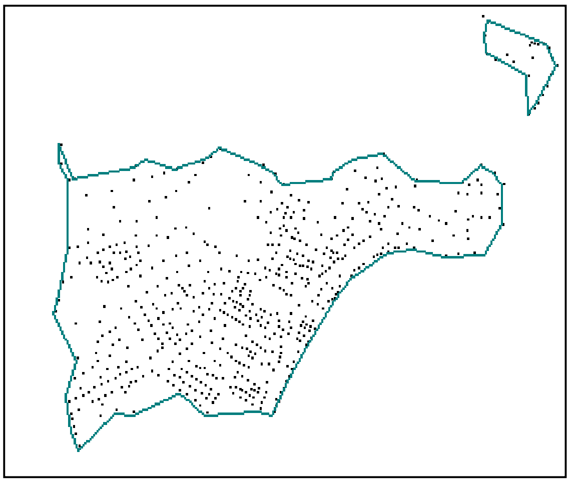

Figure 10. Two circumscribing hulls for the point data

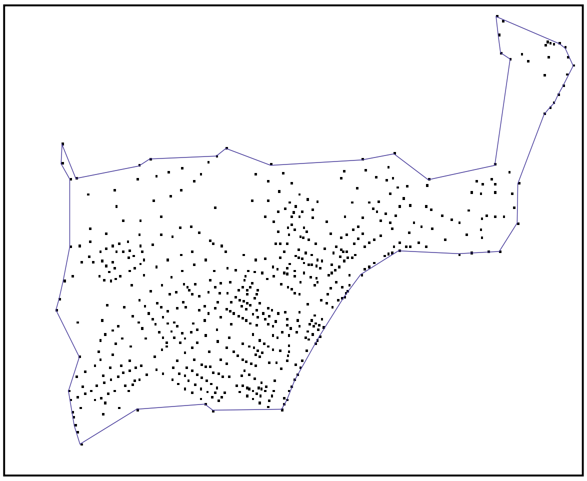

Figure 11. A single hull circumscribing the point data.

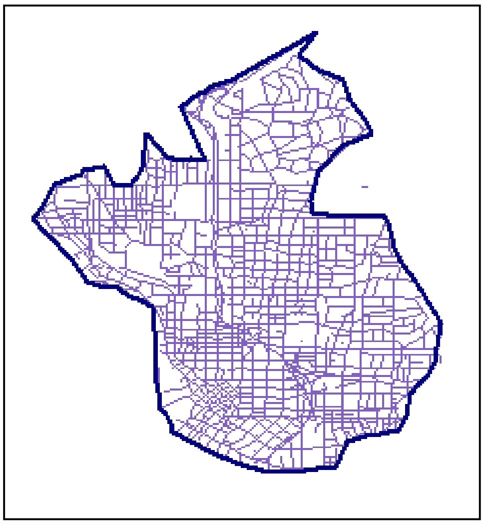

Figure 12. The circumscribing hull of line data

5 SUMMARY AND CONCLUSIONS

The ***method for construction of 2D hull*** in this research, makes use of a convex hull as an initial stage, and reduces it to a hull suitable for the purpose of generalization of cartographic digital data. This process is based on constrained triangulation of the 2D space, for the purpose of triangle division. According to the suitable tolerance, determined by a statistical analysis of the data, the triangles which cover the empty areas of the convex hull are removed. Constructing the circumscribing polygon of a group of the remaining triangles will complete the process and produce the desirable hull. Changing the tolerance enables control of the shape of the obtained hull, so that it will meet the cartographic requirements of the generalization process.

The experiment results demonstrated that the task was performed properly, and that the hulls obtained were similar to the ones that would have been defined by the human cartographer. Since the process is automated, the time invested in the cartographic digital data generalization process can be reduced significantly.

REFERENCES

Bundy, G.L., Jones, C.B., Furse, E., 1995. Holistic generalization of large-scale cartographic data. In Muller. J.C., Lagrange, J.P., Weibel, R., (Eds) GIS and Generalization Methodology and Practice, GISDATA1, pp. 106-119.

Doytsher. Y., 1988. Defining a minimum area rectangle circumscribing given information. The Cartographic Journal, Vol. 25, pp. 97-104.

Doytsher, Y., 1998. Dynamic cartography – zonal focusing. Proceedings of Cartography 98 Workshop, Tel-Aviv, pp. 61-69 (in Hebrew).

Jones, C.B., Bundy, G.L., and Ware, J.M., 1995. Map generalization with a triangulated data structure. Cartography and Geographic Information Systems, Vol. 22, No. 4, pp. 317-331.

Kang, I.S., Kim, T.W., and Li, K.J., 1997. A spatial data mining by Delaunay Triangulation. GIS 97, LasVegas, Nevada, pp. 35-39.

McMaster, R.B., Comenetz, J., 1996. Procedural quality assessment measures for cartographic generalization. GIS/LIS'96, Annual conference and Exposition Proceedings, pp. 775-785.

McMaster, R.B., Veregin, H., 1997. Visualizing cartographic generalization. 1997 ACSM/ASPRS Annual Convention & Exposition, Vol. 5, pp. 174-183.

Peled, A., 1998. Cartography and GIS. Proceedings of Cartography 98 Workshop, Tel-Aviv, pp. 30-35 (in Hebrew).

Preparata, F.P., Shamos, M.I, 1985. Computational Geometry. Springer-Verlag.

Saalfeld, A., 1993. Conflation: automated map compilation. Report of the Computer Vision Laboratory, Center of Automation Research, University of Maryland, College Park, Maryland.

Shewchuk, T.R., 1996. Triangle: engineering a 2D quality mesh generator and Delaunay triangulator. In Ming C. Lin and Dinesh Mancha, editors, Applied Computational Geometry; Towards Geometric Engineering, Vol. 1148 of Lecture Notes in Computer Science, pp. 203-222.

Timpf, S., Frank, A., 1997. Exploring the life of screen object. 1997 ACSM/ASPRS Annual Convention & Exposition Technical Papers, Seattle, Washington.

Zhou, J., Patrikalakis, N.M., Tuohy, S.T., Xiuzi Ye, 1997. Scattered data fitting with simplex splines in two and three-dimensional spaces. The Visual Computer, Vol. 13, pp. 295-315.