Analytic Generalization of Topographic and Hydrologic Data and its Cartographic Display - Intermediate Results

Eli Itzhak[(1)], Pinhas Yoeli[(1)], Yerahmiel Doytsher[(2)]

(1) Geography Dept., Tel Aviv University
(2) Geodetic Engineering, Technion Institute, Haifa

1. Introduction

Map generalization is an important and complex field in classic cartography, but even more so in digital cartography. This field is in the forefront of research and application in the academic and applied cartographic world, and has been so for several years. The main goal of most of the current research is to develop generalization applications, automated as possible, for the creation of thematic and topographic maps from existing spatial databases. This goal has not been fully realized, and the digital solutions offered so far have had limited contributions to this complicated process (Itzhak, et all., 2000).

This paper describes the intermediate results of the research, which was designed to develop analytic generalization procedures for hydrological and topographical data based on the following three major processes:

1. Generalization based on the purpose of the map, its scale and the process of classification, categorization and selection of data from the database, and their graphical representation.
2. Systematic generalization – the systematic selection of information based on pre-determined parameters and rules.
3. Geometric generalization – simplification of lines and elimination of points along these lines.

2. Fitting the Generalization procedure to the purpose of the Map

Analytic cartographic generalization is the processing of spatial data based on statistical and geometric parameters, such as hierarchy, combination, and simplification. When preparing a map, the cartographer includes information, which he deems important, and leaves out information which has no superfluous, depending upon the goal of the map, its scale, or its potential users. This process is called selection (Robinson, 1984). Statistical generalization is the process by which relevant spatial data are chosen for a specific map from a spatial database. Geometric generalization operates on the cartographic representation of chosen data.

2.1 Hydrologic and Topographic Data, the Map's Purpose and Display

As part of the research, several organizations that deal with different aspects of hydrological and topographic mapping were examined. In the sample, questions were asked about the kinds of maps they use, their reason for using these maps, and their scale. The purpose of the sample was to determine a methodology to choose the relevant hydrological and topographical data based on the users needs. This methodology is the basis for a process which automatically chooses the relevant geographic data layers for the production of topographic and thematic maps. In the field of hydrological mapping, the following uses have been defined: Analysis of the hydrological network, Drainage basin analysis, Water and pollution sources, Soil and Flora analysis.

The graphic display of a map depends upon its aim. In traditional and digital cartography there are various methods to depict topography. For instance, contour lines shaded relief, perspective views, and other dynamic or static methods. The purpose is to discuss graphic display options in relation to the map scale and ultimate use. Therefore, the first stage of the analytic generalization is the determination of which data layers are relevant, and how they should be displayed, based on the target audience, the map scale and the use of the map.

3. Systematic Generalization

At the stage of systematic generalization we choose, based on parameters and rules, which elements out of the relevant information will fit the purpose of the map. The central questions in the selection procedure are what do we include and what do we leave out? Did we make the correct choice? How can we measure our success? The mathematical aspects of the selection are given attention. The methods that have been developed give us objective tools to determine the number of objects to be included in a specific map. Even so, these methods do not indicate which specific objects should be included on a map (Malling, 1963). One of the favor methods is Pillewizer and Topfer's square root law (Topfer &Pillewizer,1966).

3.1 The Selection methodology for Hydrological and Topographical Data

Within this research, a methodology is developed for the selection of hydrological and topographical data and quality control methods are defined for the selection process. This research illustrates how Characteristic lines segments and hydrological networks are selected out of the detailed and accurate spatial database. The methodology is divided into two procedures: selection parameters determination; and, procedures for determining rules selection.
Characteristic lines were grouped into three, based on their contribution to the portrayal of the land.

1. Structure lines – describing the prominent structure lines. For example, ridge lines delineating the highest points and ravines, delineating the lowest points. The contribution of these lines among all of the structure lines is cardinal.
2. Break lines – describing prominent changes in the topography. Break lines include cliffs and steep inclines (which can not be illustrated by using contour lines), crevasses, dunes, etc. The contribution of these lines among all of the structure lines is important.
3. Man made features – describing the effect mankind has had shaping the landscape. Features such as roads, landfills, mines, ditches, dikes, terraces, etc. The contribution of these features among all of the structure lines can be important.

3.2 Choice of parameters

In this research eight groups of parameters were defined which include scores of attributes of topographical and hydrological elements. A distinction is made between primary and secondary parameters. The primary parameters are used as major filters in order to reduce the amount of data. The secondary filters are used to refine the results of the major filter. Primary parameters for the selection of characteristic lines of the relief would be length. A primary parameter for the selection of streams would be stream length, its hierarchy, and the hydrological network pattern. Examples of secondary parameters are density, closeness, longest flow route, etc.
The groups of parameters are:

1. Geometry – refers to geometric attributes of characteristic lines and streams. Such as, stream length, how winding it is, the aerial distance between its ends, the relation between its length and the aerial distance, etc.
2. Topology – refers to the spatial relationship between the line and neighboring lines, such as density, nearness, inclusiveness, etc.
3. Structure – refers to the form of a line as an isolated object or as a part of a more complex structure. Such as, stream pattern, their hierarchy, the continuity of the hydrologic network, etc.
4. Altimeter – refers to the elevation of points defining the characteristic lines. Such as, slopes, elevation differences.
5. Geography – refers to the geographic characteristics of rivers and streams. Such as, the type of line (perennial, intermittent, etc), geographic region (desert, humid), additional significance (is the river a national or administrative boundary ?).
6. Thematic – refers to the attributes of elements dealing with the spatial organization of lines. Such as, the area and shape of a drainage basin.
7. Statistic/mathematics – refers to the quantitative aspects of lines. Such as, various distributions, variances or means.
8. Symbology – refers to descriptive aspects of line depiction. Such as, cliffs or slopes portrayed as conventional symbols.

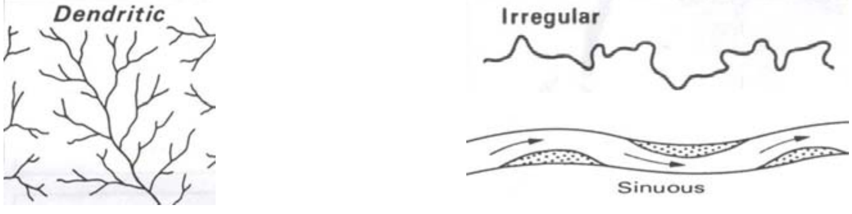


Figure 1 - Structure parameters – hydrologic network and streams pattern

3.3 The determination of selection rules

The second stage of the methodology is the process for determining the selection rules. Some 120 rules were defined to select specific elements from the data layers chosen. As mention before the methodology consist of there stages: selection rules by primary parameters; selection rules by secondary parameters; and, selection by correlated rules.

Selection rules by primary parameters

First, a layer is chosen for the selective processing. For instance, ridge lines, streams, etc. Next, secondary groups are identified based on an analysis of morphological-geographical characteristics of the elements. And then, a statistical analysis and selection of the data is performed based on the length parameter. Finally, the layer is printed at a specific map scale and the value of the primary parameter is determined.

Selection rules by secondary parameters

During the second stage, rules defined, in order to improve the selection process of elements from each of the groups (those that will be chosen and those that will be discarded), and to coordinate the choices among the various data layers. For example, the proximity parameter:

Drop all cliff symbols that are within 100 meters of a longer cliff of at least 200 meters, or choose a cliff, which is less than 600 meters from a longer cliff whose length is less than 200 meters.

At this stage, there are still "life savers" for the replacement of elements that were omitted or included. At this stage, the selection rules can be divided into three types: omission rules – elements that fit the primary rule, but are filtered out by the secondary rule. The selection rules – rules used for the repeat selection of elements filtered out by the primary rule, but still fit the secondary rule. Modify rules – rules that were designed to alter the structure of an object, such as shortening, disconnecting, connecting, etc. so that they will meet the selection criteria.

Selection by correlated rules

During the third stage, a correlated selection is performed during which the final data model is reached. The selection rules at this stage take into account relationships among elements from geographically different layers. For instance:

Drop a cliff symbol if the cliff is shorter than 200 meters and less than 100 meters from a ridge at least 200 meters long.

3.4 Parameters for the Selection of Hydrologic Data

Here, a methodology is developed for the selection of rivers based on their order. Several methods were examined : Horton (1945); Sherve (1966) and Strahler's model (Strahler, 1957) as the chosen one. Three primary parameters were defined for the selection of streams in hydrological networks. They are: stream length, hierarchy order of the stream, and the hydrological network pattern. In addition, numerous secondary parameters were defined based on the following groups:

(One) Hydrological parameters: stream type, diameter, magnitude, capacity, bank width, geometric pattern, hydrological pattern, distance between bifurcation, the name of a stream, longest flow course, stream's father and sons, ratio between stream's amount, length ratio, regularity, etc.

(Two) Topographic parameters: area and circumference of the drainage basin, shape, hierarchy and basin pattern, density ratio, elevation mass, hypsometric profile, stream slopes etc.

(Three) Importance: historical, national, symbolic, ecological, planning, tourism.

(Four) Relation to other elements:

- Hydrological elements – spring, well, pit, dam, water channel, pool, lake, etc.
- Topographical elements – cliff, ravine, break line, slope, dune, ridge, and depression.
- Land cover – sand, salt, basalt, agricultural fields, forests, flooded areas, wetlands.
- Cultural elements – settlements, archaeological digs, terraces, nature trails, etc.

3.5 Automated Identification of Hydrological Network Elements

A methodology was developed for the automatic identification of hydrologic network patterns and of stream channels. The identification of hydrologic network model components is cardinal for the application of selection rules according to the network model, and thereby keeping the uniqueness of each network in the generalization of streams.

1) Identification of hydrologic network patterns

Hydrologic networks do not develop chaotically, but adhere to one of several models. The drainage pattern attests to the character, of the stream system. It determined by the local lithology, by tectonics and by slope. The hydrologic literature describes 6-8 major drainage models, and tens of minor drainage patterns. Major patterns are: dendritic, trellised, radial, parallel, annular and rectangular (see figure 2) (Zernits, 1932; Howard, 1964; Shatner, 1971).

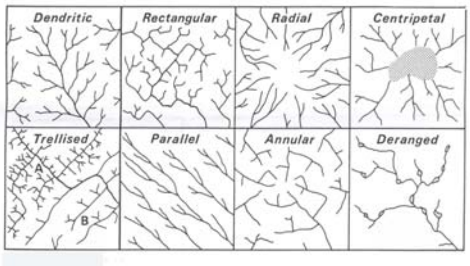


figure 2 - Stream network patterns from Morisawa, 1985.

The automated identification of a drainage system depends on the topological and statistical fit among the various elements in the hydrological system (drainage basin, hydrological network and stream channels) to the prototype of the defined models (Itzhak, et al., 2000). This research will inspect the automated network identification procedure of two typical models: "Nahal Me'arot", on the west Carmel Mountain is to be used as an example of a dendritic pattern in a humid area, and "Nahal Ofakim" as an example of a pinnate pattern ('feather like') in a desert

A 13 quantitative, topological and thematic characteristics were defined for pattern identification. The most important are : Bifurcation – how many bifurcation and endpoints are there in the network; Drainage basin – area, circumference, shape; Texture – the mean distance between bifurcation in each hierarchy; Density – the total length of streams per unit area and etc.
The data shown in table 1 and in figure 3 (a – "Nahal Me'arot" ; b- "Nahal Ofakim") clearly indicate the variance in quantitative values of the chosen characteristics of the two patterns.

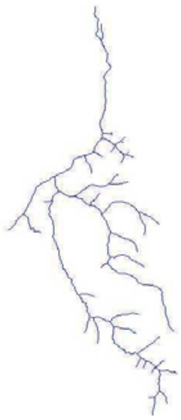


Figure 3a

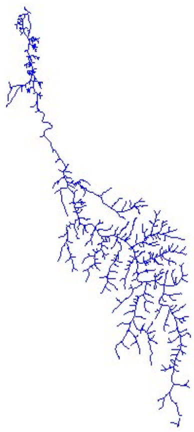


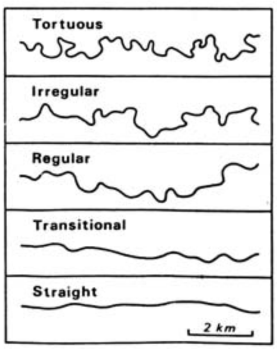
Figure 3b

Attribute	Nahal Me'arot Dendritic	Nahal Ofakim Pinnate
streams	35-40	350-400
total length	35-40 km	75-80 km
Texture	1000 m	200 m
Bifurcation ratio	1:2	1:10
Density	70-80 m/km^2	180-200 m/km^2
Frequency	0.7-1.0 segments/km^2	2-2.5 segments/km^2

Table 1 – Quantitative characteristics of hydrological Network models.

2) Stream Pattern Identification

The geometry of streams is affected by hydrological factors such as flow and load (amount, type, frequency); stream characteristics (width, slope, depth, model) and the structure of the valley and the river banks (Morisawa, 1985). The hydrological literature identifies five geometric and hydrologic stream patterns based on their sinuosity and course: straight, transitional, regular, irregular and tortuous (see figure 4)(Schumm,1963).



A methodology was defined to identify stream channel patterns. Seven characteristic groups were identified, and within them, tens of geometric characteristics (see figure 5) Such as : Length and width – attributes related to length, such as length, aerial length, length/width ratio; Area and density – attributes related to the spatial extent of a stream, such as the area of the blocked rectangle, circumference, density; Meandering – attributes related to the sinuosity of the stream such as wavelength, wave height, length/height ratio; Angularity – attributes related to the angles at which stream segments intersect along the channel, such as average, the number of straight segments, smallest and largest angles, variability model along the line; etc.

Figure 4 – Stream patterns

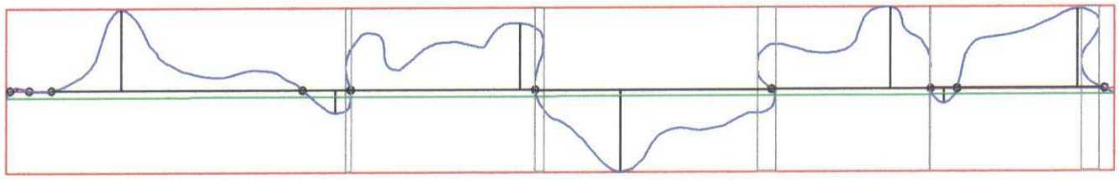
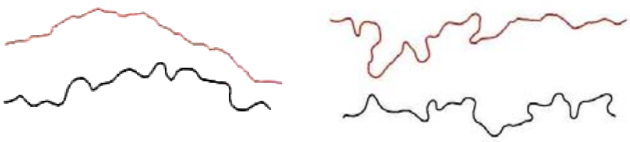


Figure 5. Analysis of geometric attributes in the irregular model.

Included in the research were 18 streams segments from various hydrological networks in the Carmel area. The first stage was the collection of data based on the numerous geometric attributes defined. The values collected were compared to the prototype values. In the second stage, major attributes were identified based on the frequency of instance in the model. In the third stage, the decision making tree was defined in the automatic identification algorithm.

These results show that each of the streams fits one of the hydrological patterns. Even though all streams fit a particular model, on occasion it is difficult to positively identify the correct model (regular versus irregular). The most important attributes when identifying a stream model are: the ratio between the length of a stream to the distance between its start and base, its X momentum attribute, the ratio between the stream's blocked rectangle's sides, the variation of azimuths of various secondary segments along the stream, and the difference between the smallest angle measured between segments to a right angle. An example of the comparison between an actual stream to its prototype is shown in figure 6.

Figure 6 – A comparison between true mesurments (red lines) and prototype (black lines). Regular pattern (left), Irregular (right)



3.6 Selection Rules for Hydrological Data

The parameters and selection rules for streams are determined based on hydrological network patterns. Each pattern has unique rules based on its geometric, topological and thematic characteristics. Table 2 presents the selection rules for dendritic streams :

Selection Stage	Parameter	Rules
Early Selection	Length/hierarchy level	Drop streams shorter than 300m if in hierarchy level 1
Advanced Selection	Neighbors	1. Drop level 2 streams shorter than 300m if closer than 200m to the next downstream intersection. 2. Drop level 2 streams shorter than 300m if closer than 150m to the next upstream intersection.
	Intersection angle	Drop level 1 or 2 streams shorter than 300m whose intersection angle with a higher level stream is $<35^o$
	Azimuth	Drop streams shorter than 300m that run parallel to another stream less than 100m away.
	Diameter	Choose all streams from each branch that have the longest channels.

Table 2. Selection rules for dendritic streams at a map scale of 1:100,000.

3.8 Parameters and Selection Rules for Break lines

The primary parameter of break lines, is length. Major secondary parameters that have been defined include: density, neighbors, continuity, parallelism and overlap. Other secondary parameters are: shape, structure, absolute elevation, direction, meander density, ground slope, etc. Table 3 presents the selection rules for ridgelines:

Selection Stage	Parameter	Rules
Early Selection	Length	Drop ridgelines shorter than 100m.
Advanced Selection	Nearness/Neighbors	Drop ridgelines shorter than 200m if closer than 60m to a longer one.
	Continuity	Connect ridge lines closer than 60m apart if their total length is >200m.
	Shape	Drop ridgelines shorter than 200m whose shape is arched if the arch is <60m diameter.
	Highest Elevation Point	Choose all ridge lines shorter than 100m if its' maximum elevation is the highest within a 200m radius.

Table 3. Selection rules for ridgelines at a map scale of 1:100,000.

4. Geometric Generalization

Analysis of current geometric algorithms indicates that they operate on the x-y position of the lines, and does not take into consideration the z values (elevation). In the current research, there is an attempt to develop a methodology for geometric simplification in conjunction with parameters from the line's profile. At the current stage of the research, the methodology is based on two major stages: two-dimensional geometric simplification, and three-dimensional simplification.

4.1 Two Dimensional Simplification

The first stage of the development of the algorithm was implemented on lines that were geometrically and hydrological identified, as streams. At this stage, it was determined whether there are different simplification parameters for different stream models. The main principles of Peaucker&Douglas's (1973) algorithm were implemented. First a buffer was drawn along the vectors that depicted streams, at varying widths, based on the map scale (see figure 7). The buffer was calculated so that it would be 0.2mm wide, which is the usual width on a printed map. With traditional cartography, the line width is dependent on the stylus width being used. In digital cartography, minimum line thickness is dependent on the resolution of the output device. It is assumed that there is no geometric significance for any point within the buffer, and therefore does not contribute to the graphic output at a given scale. According to this assumption, buffer values were calculated, and are shown in table 4.

Map Scale	Line Thickness on Map (mm)	Width on Ground (m)	Offset (m)
1:20,000	0.2	4	2
1:50,000	0.2	10	5
1:100,000	0.2	20	10
1:200,000	0.2	40	20

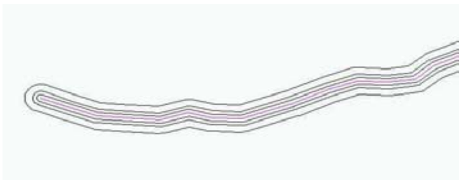


Figure 7- Buffer around a stream

Table 4. Buffer data for two-dimensional simplification, by scale.

The second stage was to sample the vectors from within the buffer at each map scale. Starting at one end, a vector is drawn to each of the center points, in their order of appearance along the line, until the new vector leaves the area of the buffer. In other words, as long as the new vector is within the buffer, it depicts meaningless points, but the point where it crosses the buffer it becomes a critical point. In this case, a point nearest to this critical point is recorded as a turning point in the new vector (see figure 8).


Figure 8 - Two diamentional line Simplification

Analysis of the new vectors shows a considerable reduction in the number of points along the reduced vector in each of the map scales and in each model. Some 12% of the original points were deemed as critical and therefore, appear in each of the models at all of the map scales. Some 15% of the original points are dropped from all of the models at all of the map scales. These points do not contribute to the geometry of the streams, and are therefore, superfluous. Table 5 summarizes the number of points along simplified vectors for each map scale in comparison to the original.

Map Scale	Straight	Transitional	Regular	Irregular	Tortuous
Source	71	87	149	220	384
1:20,000	76.0	87.5	70.4	80.9	73.4
1:50,000	50.7	63.2	61.0	59.5	52.6
1:100,000	38.0	43.6	37.5	40.9	31.0
1:200,000	9.8	26.4	27.5	25.9	20.0

Table 5 – Point Thinning for the Various models By scale.

The streams that went through the automated identification process based on the algorithm mentioned in paragraph 3.5 were used as "truth" data for the point thinning demonstration in the various models. The data indicates that during the transition from the source data to a scale of 1:50,000, there is an average reduction of some 70% in the number of points. This means that the vector can be represented by 30% of the points without degrading its geometric accuracy. In addition, this research indicates that there is no considerable difference in the number of points dropped among the various stream models. This means that the same algorithm can be used for lines depicting all streams, irrespective of their model.

In the system described above, the major criterion for dropping a point along the line is the length of the perpendicular between the center point along the original line to the thinned out line. The length of the perpendicular is dependent upon the computed buffer at each map scale. The suggested technique allows a significant reduction in the number of center points without altering the graphic appearance of the line, and without reducing its accuracy.

4.2 Three Dimensional Simplification

A 2D geometric simplification algorithm deals with the planimetric representation of lines, not taking into considaration their vertical characteristics. 3D simplification is designed to handle the altimeter of a line. This means that it drops points along the line that do not enhance the vertical representation of the surface. Therefore, three dimensional simplification is significant for the construction and simplification of generalized models of the surface. In this research, a statistical analysis was performed on numerous characteristic line profiles: channel lines ridgelines, break lines, etc. A three dimensional buffer with an elliptical cross section was implemented. The horizontal axis was determined by the perpendicular parameter from the standard two dimensional simplification, while the vertical axis was determined by the perpendicular based the relief characteristics and generalization parameters (see figure 9) :

As a result of the data analysis, several three dimensional simplification parameters were defined (see table 6):

1. Height differences between points.
2. Distance between points by map scale.
3. End points – the start and end points of each profile along the model.
4. Accumulative height is larger then the vertical parameter.
5. Accumulative distance is larger then the distance parameter.
6. The highest and lowest point along the profile.
7. Critical points of the profile depiction
8. Critical points on the two dimensional line.

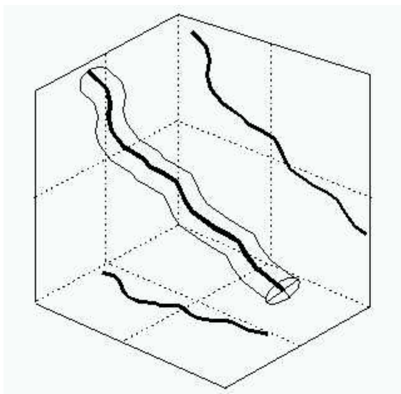


Figure 9 - Elliptical three dimensional buffer.

Scale	Contour Interval (m)	Buffer offset (m)	Distance (m)	Elevation Delta (m)	Slope (m)	Accumulative Height (m)	Accumulative Distance (m)
1:10,000	5	2	8	1.6	11.3	1.6	8
1:20,000	5,10	4,10	16	1.6,3.3	11.6,5.7	1.6,3.3	16
1:50,000	10,20	10,20	40	3.3,6.6	4.7,9.3	3.3,6.6	40
1:100,000	20,25	20	80	6.6,8.3	4.7,9.5	6.6,8.3	80
1:200,000	50	40	160	16.6	5.9	16.6	160

Table 6. Three dimensional simplification parameters of characteristic lines.

5. Surface Models

In the scope of this research, a methodology is developed that discusses the nature of models, which describe the surface in different scales. Areas with varying topographic characteristics, such as flatlands, hilly areas, mountainous, fissured areas, or cliffs, will be examined. The models are constructed in area of 35 km^2 and included a regular grid of 14,000 points every 50m. And based on the following attributes:
Model 1 – DEM grid only
Model 2 – DEM grid and structure lines
Model 3 – DEM grid together with structure lines and break lines
Model 4 –DEM grid together with structure lines and break lines and man-made features.
These four models have been generalized by applying the systematic generalization methodology presented previously.

5.1 Generalization of a Grid

In the research a distinction of a grid density based on morphology and slope of the area was made. Thus two slopes groups were defined: flat (0-2 degrees) and hilly/mountainous (over 2 degrees). Grid density parameters were defined based on scale, interval and slope. These parameters are shown in table 7:

Scale	Interval	Grid Density	
		Flat	Mountainous
1:10,000	5	25,50	10
1:20,000	10	50,75	15
1:50,000	10	100	25
1:100,000	20	100	50
1:200,000	50	100	100

Table 7. Parameters for grid reduction.

5.2 Development of Quantitative evaluation methods of Generalization

In order to evaluate the quality of selection and generalization of the models that describe the surface attributes. The research defines quantitative methods such as : Statistical analysis of a selected attributes; Analysis base on a methodology of “error strips” described by Imhof (Imhof, 1982); Analysis and comparison of profiles and cross sections; and etc.

6. Summary

In this paper intermediate results of a research for developing a methodology for analytic generalization of hydrological and topographical data were presented. The analytic model is divided into three main stages: generalization based on the maps’ intended use; selection of data based on specific parameters and rules; and a 3D simplification of characteristic lines.

The authors consider the following improvements as the main achievements of the research:
1. A comprehensive analytic generalization model, which includes selection and simplification Processes.
2. An algorithm for A three dimensional simplification of characteristic lines.
3. An algorithm that will automatically identify networks and stream channel patterns.

The upcoming stages of the research will concentrate on application of subjects discussed above. We are expecting to have a system that will automatically classify and select data upon the ultimate map; apply selection rules on hydrographic and topographic data; and, use s 3D simplification method on characteristics lines in order to generalize a surface model

Bibliography

Douglas D.H., Peucker T.K., 1973. “Algorithms for the Reduction of the Number of Points Required to Represent a Digitized Line or its Character”. **The Canadian Cartographer**, Vol. 10(2), pp. 112-123.

Horton R.A., 1945. “Erosional development of Streams and their Drainage Basins:Hydrophisical approach to Quantitative Morphology” .**Geo. soc. America Bull**, Vol. 56, pp. 275-370.

Howard A.D.,1964. “Drainage analysis in Geologicall interpretation : a summation”.**Bulletin of the American Association of Petroleum Geologists.** Vol. 56 , pp. 275-370.

Imhof E., 1982. **Cartographic Relief Presentation**, ed. Harry J. Stewart. Berlin: Walter de Gruyter&co.

Itzhak E.,Yoeli P., Doytsher Y., 2000. “Analytic Generalization of Topographic and Hydrologic Data and its Cartographic Display – Prelimenary Results”. Israeli Cartographic Association. Israel.

Maling D.H., 1963.” Some Quantitative Ideas about Generalization”. **Nachrichten aus dem Karten- und Vermessungswesen** ,(ICA Bulletin no. 4).

Morisawa M., 1985. “River morphology: channel pattern”. in **Rivers form and process**, edited by Clayton K.m., Longman, London.

Toepfer F., Pillewizer W., 1966. “The Principles of Selection, A Means of Cartographic Generalization”. **The Cartographic Journal**, Vol. 3 (1), pp.10-16.

Robinson A.H., Sale R.D., Morrison J.L., Muehrcke P.C., 1984. **Elements of Cartography**. Wiley &Sons , New-York.

Shatner Y. 1971. Introduction to Geomorphology, Hebrew University, Jerusalem.

Sherve R.L.,1966. “Statistical low of stream numbers”. **Journal of Geology** , Vol. 74, pp. 17-37.

Strahler A.N., 1957. “ Quantitative Analysis of Watershed Geomorphology**”**.
American Geophys. Union Trans. Vol. 38 (6), pp. 913-920

Zernits E.R., 1932. “ Drainage patterns and their significance”.**Journal Geology**. Vol. 40, pp. 498-521.