

# UNIT 48 - LINE GENERALIZATION

## UNIT 48 - LINE GENERALIZATION

Compiled with assistance from Robert McMaster, Syracuse University

- [A. INTRODUCTION](#)
- [B. ELEMENTS OF LINE GENERALIZATION](#)
  - [Simplification](#)
  - [Smoothing](#)
  - [Feature Displacement](#)
  - [Enhancement/Texturing](#)
  - [Merging](#)
- [C. JUSTIFICATIONS FOR SIMPLIFYING LINEAR DATA](#)
  - [Reduced plotting time](#)
  - [Reduced storage](#)
  - [Problems with plotter resolution when scale is reduced](#)
  - [Processing](#)
- [D. LINEAR SIMPLIFICATION ALGORITHMS](#)
  - [Independent Point Routines](#)
  - [Local processing routines](#)
  - [Unconstrained extended local processing routines](#)
  - [Constrained extended local processing routines](#)
  - [Global routines](#)
- [E. MATHEMATICAL EVALUATION OF SIMPLIFICATION](#)
- [F. LINEAR SMOOTHING](#)
- [REFERENCES](#)
- [DISCUSSION/EXAMINATION QUESTIONS](#)
- NOTES

## UNIT 48 - LINE GENERALIZATION

Compiled with assistance from Robert McMaster, Syracuse University

### [A. INTRODUCTION](#)

- generalization is a group of techniques that allow the amount of information to be retained even when the amount of data is reduced
  - e.g. when the number of points on a line are reduced, the points to be retained are

chosen so that the line does not change its appearance

- in some cases generalization actually causes an increase in the amount of information
  - e.g. generalization of a line representing a coastline is done best when knowledge of what a coastline should look like is used
- this unit looks at line generalization
  - line generalization is only a small part of the problem of generalization in cartography - the larger problem includes e.g. generalization of areas to points
- the focus of the unit is on line simplification
  - simplification is only one approach to generalization (see below)

## B. ELEMENTS OF LINE GENERALIZATION

- generalization operators geometrically manipulate the strings of x-y coordinate pairs

### Simplification

- simplification algorithms weed from the line redundant or unnecessary coordinate pairs based on some geometric criterion, such as distance between points or displacement from a centerline

### Smoothing

- smoothing routines relocate or shift coordinate pairs in an attempt to "plane" away small perturbations and capture only the more significant trends of the line

### Feature Displacement

- displacement involves the shifting of two features at a reduced scale to prevent coalescence or overlap
- most computer algorithms for feature displacement in vector mode concentrate on an interactive approach where the cartographer positions displacement vectors in order to initialize the direction for shifting
- another method uses a smaller-scale version of the feature to drive the displacement process

### Enhancement/Texturing

- enhancement allows detail to be regenerated into an already simplified data set
  - e.g. a smooth curve may not look like a coastline so the line will be randomly textured to improve its appearance
- one technique is to fractalize a line by adding points and maintaining the self-similarity of the original version
  - this produces fake (random) detail

## Merging

- merging blends two parallel features at a reduced scale
  - e.g. the two banks of a river or edges of a highway will merge at small scales, an island becomes a dot
  - algorithms for merging fuse the two linear features together

## C. JUSTIFICATIONS FOR SIMPLIFYING LINEAR DATA

### Reduced plotting time

- plotting time is often a bottleneck in many GISs
- as the number of coordinate pairs is reduced through the simplification process, the plotting speed is increased

### Reduced storage

- coordinate pairs are the bulk of data in many GISs
- simplification may reduce a data set by 70% without changing the perceptual characteristics of the line
  - this results in significant savings in memory

### Problems with plotter resolution when scale is reduced

- as the scale of a digital map is reduced, the coordinate pairs are shifted closer together
  - with significant scale reduction, the computed resolution could easily exceed the graphic resolution of the output device
  - e.g. a coordinate pair (0.1, 6.3) reduced by 50% to (0.05, 3.15) could not be accurately displayed on a device having an accuracy of 0.1. Simplification would weed out such coordinate pairs before reduction

## Processing

- faster vector-to-raster conversion
- faster vector processing
  - the time needed for many types of vector processing including translation, rotation, rescaling, cartometric analysis will be greatly reduced with a simplified data set
- many types of symbol-generation techniques will also be speeded up
  - e.g. many shading algorithms calculate intersections between shade lines and polygonal boundaries
    - a simplified polygonal boundary will reduce both the number of boundary segments and also the number of intersection calculations required

## D. LINEAR SIMPLIFICATION ALGORITHMS

## overhead - Linear Simplification Algorithms

### Independent Point Routines

- these routines are very simple in nature and do not, in any way, account for the topological relationship with the neighboring coordinate pairs
  1. nth point routine
    - every nth coordinate pair (i.e, 3rd, 10th) is retained
  2. randomly select 1/nth of the coordinate set

### Local processing routines

- these utilize the characteristics of the immediate neighboring points in deciding whether to retain coordinate pairs
  1. Euclidean distance between points
  2. Angular change between points overhead - Perpendicular distance and angular change
  3. Jenks's simplification algorithm overhead - Jenk's simplification algorithm diagram

- three input parameters:

MIN1 = minimum allowable distance from PT 1 to PT 2

MIN2 = minimum allowable distance from PT 1 to PT 3  
ANG = maximum allowable angle of change between two vectors connecting the three points

- algorithm:

IF distance from PT 1 to PT 2 &LT MIN1, OR distance from PT 1 to PT 3 &LT MIN2

THEN PT 2 is removed

ELSE IF angle 123 &LT ANG THEN PT 2 is removed

### Unconstrained extended local processing routines

- these algorithms search beyond the immediate neighboring coordinate pairs and evaluate sections of the line
  - the extent of the search depends on a variety of criteria, including:
    - the complexity of the line

- the density of the coordinate set
- the beginning point for the sectional search

Reumann-Witkam simplification algorithm overhead - Reumann-Witkam simplification algorithm

- the algorithm uses two parallel lines to define a search region
- after calculating the initial slope of the search region, the line is processed sequentially until one of the edges of the search corridor intersects the line

### Constrained extended local processing routines

- these algorithms are similar to those in the previous category, however, they are restricted in their search by: 1. coordinate search regions and 2. distance search regions

Opheim simplification algorithm overhead - Opheim simplification algorithm

- same as the Reumann-Witkam routine, except the algorithm is constrained by a minimum and maximum distance check, much like the Jenks's routine
- after the initial search region is set up which is similar to that of Reumann-Witkam, any points within DMIN are eliminated
- as soon as the line escapes from the search region on any side, including DMAX at the end, a new search

corridor is established and the last point within the region is saved

- the behavior of this routine around a curve is represented in C and D

Lang simplification algorithm

Johannsen simplification algorithm

### Global routines

- consider the line in its entirety while processing

Douglas simplification algorithm overhead - Douglas simplification algorithm I and II

- select a tolerance band or corridor (shaded area on slide)
  - this corridor is computed as a distance,  $t_1$  in length, on either side of a line constructed between the first and last coordinate pairs, in this example 1 and 40
  - point 1 is the anchor point and point 40 is the floater
- after the establishment of a corridor, perpendicular distances between the line connecting points 1 and 40 to all intermediate points (coordinate pairs 2- 39)

are calculated to determine which point is farthest from the line

- this maximum distance is to point 32, which is positioned well outside the corridor

the position of this coordinate pair (pair 32) is now saved in the first position of a stack

- next, a new corridor is calculated between points 1 and 32 and point 23 is found as the farthest from the centerline
  - here, point 32 is the floater
- this process continues until all points are within the corridor
- after the search has backed up to point 4, a new anchor and floater are established between points 4 and 23--the last position saved within the stack
- in this fashion the Douglas algorithm processes the entire line, backing up when necessary until all intermediate points are within the corridor and then selecting from the stack the position of the next saved coordinate pair
  - thus eventually the segment of the line between coordinate pairs 23 to 32 will be evaluated and the corridor from coordinate pair 32 to the end of the line will be the final computed segment

### E. MATHEMATICAL EVALUATION OF SIMPLIFICATION

- many different types of measures may be used to evaluate the simplification process
  - one type is simple attribute measures
  - another type are displacement measures
- simple attribute measurements are those which may be applied to a single line, such as line length, angularity, and curvilinearity.
  - these apply to either the base line or a simplification
- displacement or comparative measurements, on the other hand, evaluate differences between the base line and simplification

overhead - Measures for linear simplification

overhead - Areal displacement

- it appears that some of the algorithms are much better than others in maintaining the critical geometric characteristics of the data
  - Douglas, Lang, Reumann-Witkam, and Opheim all appear to be reasonable choices
  - the two best are Douglas and Lang

### F. LINEAR SMOOTHING

- smoothing is applied to digital line data in order to improve the aesthetical qualities of the line and to eliminate the effects of the digitizing device
  - in general, it is felt that smoothing improves the quality of these data
- smoothing increases the number of coordinates needed, so is normally used only for output

### REFERENCES

Buttenfield, B.P., 1985. "Treatment of the Cartographic Line," *Cartographica* 22(2):1-26.

Douglas, D.H. and T.K. Peucker, 1973. "Algorithms for the Reduction of the Number of Points Required to Represent a Line or Its Character," *The American Cartographer* 10(2):112-123.

McMaster, R.B., 1987, "Automated Line Generalization," *Cartographica* 24(2):74-111.

McMaster, R.B., 1987. "The Geometric Properties of Numerical Generalization," *Geographical Analysis* 19(4):330-346.

McMaster, R.B., 1989. "The Integration of Simplification and Smoothing Algorithms," *Cartographica* 26(1).

Peucker, T.K., 1975. "A Theory of the Cartographic Line," *Proceedings, Second International Symposium on Computer- Assisted Cartography, AUTO-CARTO-II, September 21-25, 1975* U.S. Dept. of Commerce, Bureau of Census and ACSM, pp. 508-518.

White, E., 1985. "Assessment of Line-Generalization Algorithms Using Characteristic Points," *The American Cartographer* 12(1):17-27.

#### DISCUSSION/EXAMINATION QUESTIONS

1. Discuss the differences between sequential and global approaches to line simplification.
2. What are the five generalization operators for digital line data? Discuss each one of these and give examples.
3. Using a series of diagrams, discuss the procedure used by the Douglas algorithm.
4. Discuss the different approaches you might use to evaluate the effectiveness of line simplification procedures and the advantages and disadvantages in each case.

---

*Last Updated: August 30, 1997.*