STRUCTURAL KNOWLEDGE TO SUPPORT THE GENERALIZATION OF A COASTLINE

Sven Arve Saga
Statens kartverk (Norwegian Mapping Authority)
N-3500 Hønefoss, NORWAY
e-mail: sven-arve.saga@skripost.md.telemax.no

Abstract

Extensive knowledge of the phenomenon in question is essential in the generalization of geographic data. Using a coastline as a reference, the concepts of geometrical, structural and procedural knowledge for cartographic generalization are discussed. It is shown how structural information may be derived from the geometrical information. A brief discussion follows of how different structures demand different generalization solutions.

1 Introduction

The area of cartographic generalization has been one of the most focused within cartographic research for decades. The process of cartographic generalization has proven extremely difficult to formalize and implement in a computer environment. Mapping organizations concerned with geographic information at different levels of details are in great demand of automated routines conveying more 'cartographic insight' than those available today. The lack of adequate automated routines for cartographic generalization force organizations to establish and maintain several databases at different levels of details representing the same information. Such a strategy violates one of the basic principles of database theory; namely that information should be stored only once in the database to prevent inconsistent data in the organization.

Armstrong (1991) emphasized that knowledge of the geometry alone, as represented by coordinates, is not sufficient to perform cartographic generalization. From the manual generalization process, we know that cartographic generalization is often characterized as an 'intuitive' process. We know however that an appropriate generalization presuppose knowledge of the features in question. Hence, there is a need to collect and store structural information in the geographical databases to support generalization.

2 Geometrical, structural and procedural knowledge

Armstrong (1991) categorizes the knowledge needed for generalization in

- Geometrical knowledge
- Structural knowledge
- Procedural knowledge

<u>Geometrical knowledge</u> is represented by the coordinates *defining* the relative and absolute positions of the features. Also topological information, whether derived in each instance or explicitly stored in the database, is part of the geometrical knowledge.

Structural knowledge 'brings expertise that ordinarily resides with the cartographer into the automated generalization process' (Armstrong 1991). Such knowledge is based on knowledge about the nature of the features in question and their generating processes.

Procedural knowledge is the knowledge of when and how to apply the appropriate generalization operators and algorithms for a specified task.

Armstrong (1991) pointed out that important areas of future research will centre on how existing information stored in GIS databases can be transformed into knowledge to support generalization decisions. What structural information is relevant to the generalization processes and how to establish and represent this information are important questions.

3 Coastal forms and structure signatures

One of the primary goals of cartographic generalization is to preserve important characteristics of the features treated while at the same time simplifying the geometry. The following is cited from Imhof (1982) from a chapter discussing the generalization of contour lines:

> To generalize well, to enable the simplified, reduced form to present the correct character, requires frequent and careful observations of the landscape and intensive study of exact contour maps rich in detail. As already emphasized, a geomorphological training will help the cartographer to develop an eye for characteristic form. The relief of the earth's surface is composed of innumerable combinations of geomorphological features.

A digital representation of 2300 km of coastline from the western part of Norway is used here to explore the concepts of geometrical, structural and procedural knowledge in the context of generalization. The coastline is digitized from the main map series in scale 1:50 000 from the Norwegian Mapping Authority by scanning at 50 μm and subsequent vectorization. How should structural knowledge be expressed in this case ? In Websters (1989), structure is defined as 'The aggregate of elements of an entity in their relationships to each other'. It seems logical to include a geomorphological classification of the feature in the structural knowledge. Two distinct geomorphological classes are represented in this coastline, namely *Strandflatcoast* and *Fjordcoast*. A description of the two coastal forms is found in Klemsdal (1982):

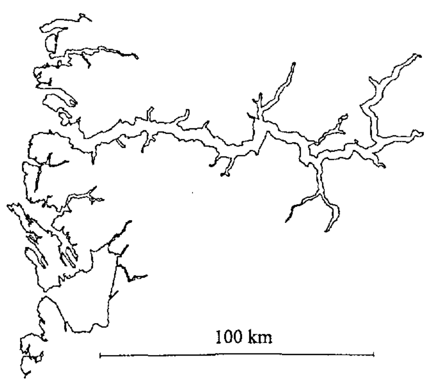

Figure 1. The coastline referred to in the paper. From western part of Norway.

Strandflatcoast. The strandflat is a brim of a gentle sloping bedrock plain in front of a higher land or coastal mountains. The plain has a very rugged terrain with small differences in height. Most of the bedrock plain is covered with a thin cover of loose material which only locally have forms of its own. Meeting the sea, the gentle sloping flat produces an uneven coastline with numbers of bays, coves, inlets, islands, islets and skerries.

Fjordcoast. A fjord is an arm of the sea stretching inland between distinct fjord-sides which either plunge steeply into the sea, or via a more gently sloping valley bench reach the sea. The fjord-sides continue down to the floor of the fjord, giving the general U-profile in cross-section. Along the fjords, the shore-zone is mainly made up of steeply sloping ice-smoothed rocky shores or stony beaches.

The concept of structure signatures is used by Pike (1988) and Buttenfield (1991) in different contexts. Pike (1988) defined a set of parameters for terrain classification, capable of discriminating between different terrain types. Buttenfield (1991) used the concept of structure signatures as a tool for describing changes in geometry during change of scale. As used by Pike (1988), structure signatures can be viewed as multi parameter "fingerprints" capable of distinguishing the different terrain type from each other.

Trying to find similar structure signatures to distinguish between fjordcoast and strandflatcoast, 6 parameters were defined to characterize the coastline. These are *angmean* (for each vertex-point a break angle between 0 - 400^g is calculated. Angmean is the mean value of all break angles in the line); *angdev* (the standard deviation of angmean); *lenmean* (the mean length of all the sides constituting the line); *lendev* (the standard deviation of lenmean); *entropy* (an expression of the information content of the line using an information theoretic approach, see Bjørke (1992)); *curvilinearity* (the presence of low frequent details along the line, see McMaster (1985)).

The line shown in figure 1 consists of 51974 points and was broken up in smaller pieces of 250 points, giving an average length of 11 km per line. The assumption is made that these pieces may be seen as homogeneous with respect to geomorphological classification. The 6 parameters described above were then computed for each of the 209 lines. A principal component analysis of the 6 parameters of the 209 lines gave two major components, explaining a total of 86 percent of the variance in the data. The first component had loading from entropy, lenmean, lendev, angmean and angdev and is interpreted to express the presence of high frequent details along the coastline (ruggedness). This interpretation seems reasonable after a visual inspection. The second component was loaded mainly from the parameter curvilinearity, but after a visual inspection it seems difficult to give this parameter an adequate geometrical interpretation.

A classification of the coastline is given by Klemsdal and Sjulsen (1992) and is considered 'true' when tuning the values for automated classifications. An automated classification based on the single parameter *angmean* is chosen, since this single-parameter 'fingerprint' served better for the purpose to distinguish the two coastal forms than multi-parameter 'fingerprints' based on the principal components. Using a value of 37^g (low values for fjordcoast, high values for strandflatcoast) to classify the two coastal forms, only 45 out of the 209 lines (21.5%) were incorrectly classified. As all the incorrectly classified lines are located in the transition zone between the two coastal forms, we conclude that the parameter *angmean* may serve as a tool for an automated classification of this particular coastline. Hence structural knowledge represented by a geomorphological classification may be established from the coordinate data and stored explicitly in the database to support generalization operations.

4 Implications for generalization

The reason for including information about coastal forms in the database is an underlying assumption that the two coastal forms should be treated differently in the generalization process. Each of the two forms are shown below and illustrate the difference in structure. For each form both the original line and 3 simplified versions are shown. The fjordcoast and the strandflatcoast consists of 2250 and 2500 points respectively.

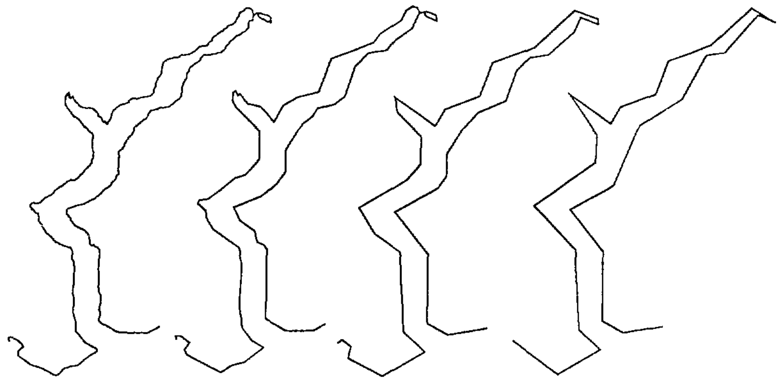

Figure 2. Generalization of fjordcoast. Douglas-Peucker algorithm, tolerances 0, 200, 400 and 800 meters.

Figure 3. Generalization of strandflatcoast. Douglas-Peucker algorithm, tolerances 0, 200, 400 and 800 meters.

Figures 2 and 3 show that the coastal form fjordcoast is more robust than the strandflatcoast. Not surprising, the character of the smooth outline of the fjordcoast better survives the increasing simplification than the more rugged outline of the strandflatcoast. This implies that the parameters should be defined separately for the two coastal forms.

5 Discussion

One of the primary goals of cartographic generalization is to preserve the character of the features involved. In this example the coastal form were chosen to represent structural knowledge. The next step is to establish the generalization rules for the two coastal forms. More research is needed to explore the nature of the structural knowledge needed to support the generalization processes and to define guidelines of how to establish the procedural knowledge (i.e. algorithm and parameter selection). To make further progress in the area of automated cartographic generalization, an interaction between theoretical work and empirical experiments is necessary to develop a better understanding of the data and processes involved. Formal tools to evaluate the information content and the visual effectiveness of a cartographic line would enable an objective evaluation of different generalization alternatives. The application of information theory for the purpose of map evaluation is a very interesting approach (Bjørke 1992)

References:

Armstong, M.P. (1991)
Knowledge classification and organization
In Buttenfield, B.P & McMaster, R.B (ed): Map generalization, Longman 1991

Bjørke, J.T. (1992)
Towards a formal basis for cartographic generalization
In J.T.Bjørke (ed.): Proceedings Neste Generasjons GIS
The Norwegian Institute of Technology, 1992

Buttenfield, B.P. (1991)
A rule for describing line feature geometry
In Buttenfield, B.P & McMaster, R.B ed: Map generalization, Longman 1991

Imhof, E. (1982)
Cartographic Relief Presentation.
de Gruyter, 1982

Klemsdal, T. (1982)
Coastal classification and the coast of Norway
Norsk Geografisk Tidsskrift 3/82.

Klemsdal, T. and Sjulsen O.E. (1992)
Landforms
National Atlas of Norway, Norwegian Mapping Authority, 1992

Pike, R.J. (1988)
Toward geometric signatures for geographic information systems
IGIS, 1988.

Webster's Ninth New Collegiate Dictionary (1989)
Merriam-Webster Inc. 1989