SCALE, SINUOSITY AND POINT SELECTION IN DIGITAL LINE GENERALIZATION

Geoffrey Dutton

ABSTRACT. This paper examines some assumptions and results of cartographic line simplification in the digital realm, focusing upon two major aspects of map generalization , *scale-specificity* and the concept of *characteristic points*. These are widely regarded as critical controls to generalization, but in our estimation are rarely well-considered or properly applied. First, a look at how scale and shape are treated in various research papers identifies some important conceptual and methodological issues that either have been misconstrued or inadequately treated. We then conduct an empirical analysis with a set of line generalization experiments that control for resolution, detail and sinuosity using four source datasets as examples (two island coastlines digitized at two scales), resulting in about 100 different generalized versions of these figures. These tests systematically explore consequences of linking scale with spatial resolution and a variety of point selection strategies. Graphic results of this exercise are displayed (at scale and enlarged) along with some basic statistics. We feel that most of the common measures of generalization performance rely on dubious assumptions about the value of fidelity of output to source, but at this point we have no specific alternatives to propose in their stead. Rather than proof, then, we offer evidence that whether our methods are used or others are, sensitivity to scale and sinuosity can and must be built in to all digital map generalization toolkits.

keywords: digital cartography, generalization, line simplification, map scale, hierarchical coordinates, sinuosity, cartometry, characteristic points

Introduction

On paper maps, lines usually have beginnings, ends, widths and flow from place to place as continua, turning occasional corners. In digital maps, lines have beginnings and ends but not widths, and are sequences of dimensionless points, each of which has little intrinsic importance, possibly representing nothing in particular. Digital cartography has, however, instilled in its practitioners a sense that such points really do exist, and that their importance can be quantified and classified. Perhaps that is why — seen from other perspectives — the literature on map generalization often seems to make much ado about nothing, or almost nothing. This paper attempts to sort through some assumptions found in the literature on line generalization, and then to empirically test them to see how useful they are.

We begin with some observations regarding the primary impetus for generalization, changing map scale. Given the pervasiveness of this activity, there is surprisingly little reference to scale in the automated generalization literature. Most of what is found derives from a misinterpretation of Töpfer's work, still the standard in the field. We then examine the concept of *characteristic points*, seen in the literature as describing major inflections along a feature that contribute more overall shape information than most locations do. The notion comes from the work of psychologist Attneave, and into cartography via Douglas, Poiker, Buttenfield and others, including computer vision researchers. This line of thought makes the assumption that the identities of such points are knowable only after analyzing an entire feature as a whole. Furthermore it has been assumed — based on relatively slim evidence — that having identified such points on a feature, they should be preferred to other points in the process of simplifying its shape for generalization purposes. To test these notions, we present evidence that selecting characteristic points chosen for their global importance may produce less than optimal results in simplifying map features, compared to other point selection strategies. We then show how adding local intelligence to cartographic data and processing it using local, contextual generalization methods can be helpful. While such techniques impose added burdens of complexity for computation and control, this price is probably inevitable if automated generalization is to become truly feasible for interactive and hardcopy mapping.

Map Generalization and Map Scale

Generalization is not simply making little things look like big things. Children are proportioned differently than adults, and map features at small scales should not slavishly mimic their

shapes at large scales. In addition, contextual factors may also influence shape representation, such that simplifying features can require more than merely removing vertices; sometimes entire shapes (such as a meander or a switch back) should be deleted at a certain scale (Plazenet et al. 1995; Wang and Muller 1998). In other instances entire features (points, polylines, polygons or sets of them) will need to transformed or eliminated. But long before it vanishes, a feature will tend to lose much of its character as a consequence of being reduced. This needs to occur in a sensible way; as the following sections describe, commonly-used numerical criteria for evaluating solutions do not necessarily provide useful guidance, in part because they do not reflect the imperatives of map scale, in part because they are too global, and because the geometric properties they preserve may be undesirable.

Articles describing techniques for map generalization frequently observe that their application should always be tempered by the scale and purpose of the map being produced (McMaster and Shea 1992). To these considerations some authors also add constraints based on the display medium and its resolution (McMaster 1987). Thus, maps intended for interactive data exploration may be generalized to different standards or in different ways than maps destined for atlases or published articles. Having made such disclaimers, many technical studies proceed to ignore them, or at least do not outline control strategies for handling constraints of map purpose, scale or medium. Somewhat surprisingly, even map generalization studies directly concerned with scale-changing operations often fail to relate either techniques or results to specific map scales (Cromley and Campbell 1990, 1992; Zhan and Buttenfield 1996). Others do describe specific numerical invariants of scale change (Töpfer and Pillewizer 1966; Töpfer 1972; Muller 1987; Li and Openshaw 1992, 1993). Some such studies restrict themselves to relative scale change, but their graphics are not always sized appropriately. Some fail to specify the scale of source data, and may not even describe the scale reduction ratios arising from the application of techniques; often results are presented in terms of (input) tolerance value parameters and (output) absolute or percentage changes in the numbers of points, line lengths, etc. (McMaster 1986). The otherwise well-designed study by Jasinski (1990) is representative of a tendency to use test data with undocumented sources, projections, coordinate units and particularly scales. Failure to provide such metadata impedes evaluation and implementation of techniques.

When interacting with a geographic information system (GIS), a user may not have even a vague idea of the scale at which maps are being drawn unless the system computes and presents this information continuously (ESRI's ArcView™, for example, does this when projection parameters are provided to it). When printing maps from a GIS, users may designate a scale (as when a standard map product is being generated), but frequently map scale is defaulted to fit maps to the available display space. This avoids burdening users with operational details, but it requires that the mapping system manage them properly. As GIS map generalization tools rarely treat output scale as an inherent constraint, default displays are often unacceptable. A recent monograph by João (1998) effectively describes and discusses critical characteristics and operational problems of map generalization as implemented in various manual and automated environments.

Applying the Radical Law

Töpfer's *Radical Law* (Töpfer and Pillewizer 1966; Töpfer 1974), mentioned above, has been a quantitative touchstone since its introduction to the cartographic literature over 30 years ago. It was originally expressed not as a single equation but as a family of them, parameterized according to the nature of the phenomena being generalized. The most oftencited variant is given as:

$$n_f = n_a \sqrt{(M_a/M_f)},$$

"where n_f is the number of objects which can be shown at the derived scale, n_a is the number of objects shown on the source material, M_a is the scale denominator of the source map, M_f is the scale denominator of the derived map" (Töpfer and Pillewizer 1966, p 11, eq. 1). Töpfer generalized this rule to include a multiplicative constants and an exponent, such that the general form of the law became:

$$n_f = n_a C \sqrt{(M_a/M_f)^x}.$$

These parameters vared according to the type of objects being mapped, as well as the purpose of the generalized map. Some researchers seem unaware of these variations, and apply the principle as if the exponent *x* is always unity. Töpfer specified that a value of 1 applies to point symbols, whereas for areal symbols — for example islands — *x* should be 3. And for linear symbols — including points along digitized boundaries— the exponent should be 2, leading to:

$$n_f = n_a M_a/M_f.$$

In other words, according to this formulation of the Radical Law, *the number of points required to represent a digitized line should decrease linearly to the ratio of scale denominators*. The study by Barber et al. (1995) adopted the value 1 for *x*, which according to Töpfer and Pillewizer is appropriate not for linear symbols but for

point symbols. As a result, the generalizations they conducted all have more coordinates than Töpfer's rule, correctly applied, would yield, leaving considerable imperceptible detail at the intended scales. In our study we did not consciously attempt to make simplifications consistent to the Radical Law; in some instances the results were consistent, but in others they diverged from it. Muller (1987) noted similar empirical discrepancies, and suggested that fractal dimensionality be included in a revised formulation of the Radical Law.

The Radical Law, an empirical heuristic that reflects basic physical realities[1], quantifies how graphic scale constrains the number of map features that can occupy a given space. No useful alternatives to it seem to exist. Given appropriate parameter values, the Radical Law can provide estimates of map feature density that generalization operators can target. These can serve as performance criteria which should help in choosing parameter values for algorithms. Because the Radical Law is global and constant for any given feature type and reduction ratio, it offers a way to tune generalization operations to enforce an underlying Gestalt (graphic unity), maintaining the visual balance of a map displayed at smaller scales in a consistent way. More research should be devoted to exploring these numerical possibilities.

Scale and Perceptibility

One reason for a lack of scale-awareness is that digital (vector-encoded) map data are inherently scaleless, even though their source material may have well-defined scales. Also, map digitizing processes can vary greatly in the density and uniformity of detail captured unless they are well-controlled. In the absence of metadata, standards employed for data capture may not be communicated to users. Even if they are, users may not know how much reduction in coordinate volume is necessary to reach a target scale. The simplest strategy for avoiding irrelevant detail is *coordinate culling*, weeding out all vertices that are closer to their neighbors than a certain distance determined by scale, map purpose and output resolution. Tolerances for culling should be just below the threshold of perceptibility (about 0.2 mm segment length at target scale).

However, culling only assures that a minimum separation between points is maintained; it can eliminate most of the imperceptible inflections but does not guarantee that the remaining details are sufficiently simplified or serve a useful purpose (and might even cause perceptual problems by eliminating important points). Usually, culling criteria are enforced as a *post hoc* condition (requiring analysis and post-processing of generalized lines) unless the generalization algorithm itself is able to guarantee a minimum segment length. Note that minimum desired segment size is a graphic specification, typically described as mm on a display surface (i.e., a measurement on a map). Enforcing it requires knowledge of the coordinate ground units (e.g., meters) and of the target scale.

One useful statistic that generalization operators can enforce is *average segment size* (mean segment length, or MSL). Two types of MSL should be distinguished, those in *ground* coordinates (MGSL) and those in *map* coordinates (MMSL). When features are simplified for scale reduction, MGSL will of course increase, and should generally do so monotonically with scale denominator. MMSL, however, should stay roughly constant at all scales. Its magnitude should be above the threshold of legibility, which depends on output medium, line weight and possibly on feature type (MSL can be skewed by the presence of long segments, as might be found in transportation features). In our studies (of coastlines) we tried to maintain MMSL between 0.5 and 1 mm. We consider MMSL to be a useful indicator of the *degree of detail* in generalized features, but to assess the *quality of detail* other measures of their shape complexity are needed.

Assessing Significance of Shape Points

The second major concern of this study is to review the concept of *characteristic points* on cartographic lines. The goal is to better understand whether locations on lines globally-identified as significant are really the most appropriate points to select for generalization purposes. To this end, we have processed several versions of two island coastlines using two algorithms, systematically varying parameter settings. Maps, charts and statistics are provided to help compare and assess the outcomes, but there are also subjective factors that cannot be quantified or precisely evaluated. These in part concern the scale and the purpose of map generalization; while we are able to control scale, the purpose here is to explore map generalization techniques rather than to generate a specific cartographic product. In any case, realizing that readers' aesthetic judgments can diverge from ours, we are

[1] Although empirical, the Radical Law expresses lawful physical constraints of size on shape. The field known as *allometry* organizes these concepts, and has been applied to natural and social sciences and engineering to predict how objects and organisms scale, changing shape with growth. The Radical Law is one of many expressions of allometric growth; for examples, see (Gould 1966).and Dutton (1973).

not attempting here to "prove" ours are correct.

Local v. Global Generalization Strategies

Along with other characteristics, line generalization algorithms are described as local, regional or global in scope (Douglas and Peucker 1973, Zycor 1984; McMaster 1987; João 1998). Although Fourier-based methods for simplifying lines have been proposed (Moellering and Rayner 1982), there is really only one *global* algorithm in use, Ramer-Douglas-Peucker (RDP).[2] It dominates in part because of that property, which everyone seems to assume is helpful. Starting with Attneave (1954) — who studied drawings rather than maps — the idea that certain locations along lines provide more information about line shape than others do became part of cartographic conventional wisdom. Because RDP is superficially good at finding such locations, it was assumed it should offer a sound perceptual basis for line generalization, and this was reinforced by various empirical studies (Marino 1978; White 1985; Buttenfield 1984; McMaster 1986; Jasinski 1990). However, other work has shown that RDP often yields generalizations that are visually unacceptable (Visvalingam and Whyatt 1990; Beard 1991; Zhan and Buttenfield 1996). Jasinski's study might also have shown graphic shortcomings of RDP, but did not present enough plots for readers to determine this). Over the years the algorithm was built into many GIS platforms, and its ubiquity along with findings that it produced recognizable simplifications and good metrics[3] (such as small displacements) seem to have obscured the fact that it can easily produce bad-looking, if not incorrect, lines.

It is also widely acknowledged that RDP is sensitive to initial conditions (anchor lines); in the GIS world these tend to be dictated by arc-node topology and are not easily changed. However, while some work has been done that describes consequences of changing initial RDP anchor points (Buttenfield 1986, Visvalingam and Whyatt 1990) little experience and few practical guidelines or tools exist in this area. New approaches to feature segmentation, such as by Plazanet et al. (1995), Wang and Muller (1998) and Dutton (1998) are confronting such problems. By segmenting line features to be more homogenous, then applying appropriate algorithms and parameters to each regime individually, simplification results can almost always be improved. To do this with RDP would require setting the tolerance parameter rather frequently, which could burden users with difficult-to-make decisions.

Selecting Tolerance Values

Tolerance-setting for line generalization is normally left to users to decide on a case-by-case basis, and very little human guidance or automated help is generally available for this.[4] Some simplification algorithms have tolerances that are fairly intuitive, but this is not the case for RDP. As Cromley and Campbell (1992) discuss, there may be no simple relationship between tolerance values and intended display scales; the effects of changing the parameter may be non-linear, making the process difficult to tune. They and others have experimented with determining the number of points to be retained by applying the Radical Law to the scale reduction factor, using a modified RDP that will reduce lines to just the number of points requested. A study by Barber et al. (1995) comparing hierarchical and non-hierarchical RDP point selection indicated that both approaches yield good statistical results for measures such as vector- or area-displacement.[5] Yet, even well-chosen tolerance values can still result in self-crossings, uneven densities, clutter and other unwanted artifacts. Some of these problems can be due to the given initial end-points, which as mentioned above are rarely modified. Others are simply direct consequences of choosing the most salient points, as we will further discuss.

There is more to good simplification than just cutting down clutter to acceptable levels. Cromley and Campbell (1990, 1992) also demonstrate clearly that results depend on *which*, not just on *how many* points are selected; they then apply linear programming techniques

---

2This is referred to as the "Douglas" or (usually), "Douglas-Peucker" algorithm (developed by David Douglas and presented in Douglas and Peucker 1973). However, essentially the same algorithm was independently proposed slightly earlier by Urs Ramer, in an image processing context (Ramer 1972). To reflect this, we refer to the method by both names.

3Many of these measures derive from the work of McMaster (e.g., McMaster 1986), and compare generalized lines to source lines using various linear and areal measures. In this and later work there seems to be an assumption that good generalizations minimize these differences (that is, cause the least change in shape). It is debatable how useful this criterion is for judging generalization quality.

4Arc/Info™'s GENERALIZE command (which simplifies arcs in a coverages using the RDP algorithm) provides a default tolerance of 0.02% of the map width. Assuming a coverage is intended to be printed on a sheet one-half meter across, the default value is 0.1 mm.

5In hierarchical selection, no points are retained at smaller scales that would have been rejected at larger scales. RDP does not guarantee this property, but can be modified to achieve it.

to enforce certain geometric criteria for point selection. In the 1992 paper they compare this type of optimization (which is too computationally expensive for general use) to RDP using a digitization of the coast of Connecticut, minimizing vector displacement, and both minimizing and maximizing line length. RDP is shown to produce lines about 90% as long as the latter optimization, but of very similar visual quality. The minimum-length optimization solutions are extremely smooth, having lost almost all hint of the bays and estuaries characterizing this coast, and the authors describe them simply as "bad." That, however, is a premature judgement which takes no account of the intended scale and use of the simplified data.

Assessing Generalization Performance

McMaster's cartometric measures of lines such as total length, segment size, vector displacement, areal displacement, angularity and others have objective validity and have often been used to quantitatively compare generalization solutions (McMaster 1983; 1986; Jasinski 1990; Cromley and Campbell 1990, 1992; Barber et al. 1995). Used as guides for selecting procedures or parameters, however, they can be quite misleading. Optimizing for them tends to produce caricatures that closely resemble input data to the exclusion of aesthetic concerns, due to satisfying global criteria rather than local ones. Embracing this approach, Cromley and Campbell (1990: 78) claim that "... line simplification should minimize the number of retained points while ensuring that all points of the original string are within some tolerance of the simplified string." Recognizing that there might be generalization goals other than fidelity to source data, McMaster (1987), following Jenks (1981), classified controls for simplification as a 3-dimensional framework having continuous or quasi-continuous orthogonal axes. This took into account 1) output scale, 2) output resolution and 3) map purpose. But the goal still remained "to select the salient features along a digitized line" (McMaster 1987: 100).

However, there is some controversy — if not confusion — about how salient features along a digitized line might be identified. The consensus seems to be that these tend to be the same critical or characteristic points identified by RDP. Nevertheless, there is disagreement, and some evidence to the contrary. Visvalingam and Whyatt (1990) conducted an evaluation of RDP's performance, comparing it to manually generalized versions of the source data (British coastlines). Various differences were noted, almost all favoring manual generalizations: "Points selected by the Douglas-Peucker algorithm are not always critical. Manual generalizations take into account the relative importance of features. This is partly dependent upon the purpose of the map. Even if we ignore such variable factors, manual generalisations tend to preserve the shape of geometrically large features at the expense of smaller ones." (Visvalingam and Whyatt 1990: 224).

We concur that selection of characteristic points, in being so blindly emphasized, often causes simplifications to be sub-optimal because the most descriptive shape points at one scale may be less so at another — and may in fact be bad choices — precisely because they create such strong caricatures. To test this assertion, however, a more controllable approach to point selection than RDP permits is required.

Linear Simplification via Mesh Filtering

Exploiting techniques first described in Dutton and Buttenfield (1993), updated in Dutton (1997a) and documented in Dutton (1998), a planetary triangular hierarchical tessellation called quaternary triangular mesh (QTM) was used as a framework for linear simplification across a range of map scales. Unlike many hierarchical space partitioning schemes such as most rectangular quadtrees (see, for example, van Oosterom and Vijlbrief 1997), QTM is defined in geographic coordinate space and covers an entire planet, pinning each areal unit at all levels of detail to a specific triangular region of the earth's surface that never changes. Several levels of detail of this structure are illustrated in figure 1. It provides both upper and lower limits to encodable detail that are roughly constant across the planet (the triangular facets vary somewhat in size and shape within each level). The lower limit is given by partitioning a spherical octahedron into 32 facets (dividing edges in half and connecting their midpoints to produce a fourfold triangular decomposition). This process continues recursively, doubling linear resolution at each level of detail. Thirty such levels yields triangular facets about one centimeter across, which provides sufficient detail for any cartographic application. Based on its known or presumed positional accuracy, each geo-coordinate is replaced by the QTM leaf node ID that it occupies. Medium-scale maps do not require much more than 20 QTM encoding levels to represent their coordinates. The examples given below all have 18 levels of detail or fewer (ca. 40 m resolution or larger).

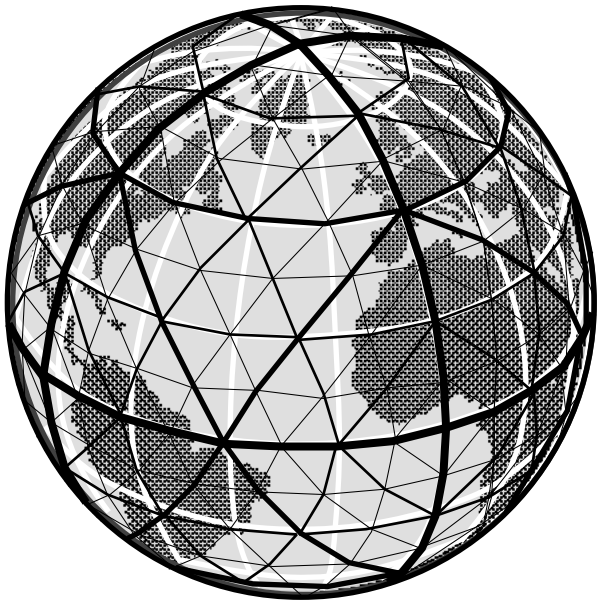


figure 1: Quaternary Triangular Mesh to three levels of detail referenced to the globe

Generalization via QTM is primarily a spatial filtering process. Starting at a dataset's highest encoded level of QTM detail, this filtering works by collecting line vertices that occupy facets at the (coarser) target level of resolution, then sampling coordinates within each such facet, similar in spirit to methods developed by Van Horn (1985), Li and Openshaw (1992, 1993) and Zhan and Buttenfield (1996), but tied concretely to geographic space, rather than filtering projected coordinates through a floating grid. The fixed nature of the sampling grid is illustrated in figure 1. The basic sampling method is shown in figure 2, and some specific effects in figure 3. The process is efficient, as it only compares QTM identifiers and does not require geometric computations. Conversion of coordinates to and from QTM is also quite efficient and can be done without involving users, except to obtain positional accuracy parameters (Goodchild and Yang Shirin 1992; Dutton 1998: app. A).

The essential aspects of this simplification approach are:

1. Geographic coordinates (latitudes and longitudes) are directly handled, after conversion to QTM location identifiers.
2. A fixed, global, hierarchical triangular grid is used for sampling points.
3. Edges of the hierarchy bifurcate, halving linear resolution at each level of detail.
4. Ten to twenty levels of detail are adequate to capture and generalize medium-scale map data, twenty-five levels will suffice for most large-scale maps.
5. A map scale can be associated with each level, based on certain assumptions.
6. The degree of simplification is based on QTM level plus several other parameters.
7. The quality of simplification can be manipulated by utilizing point selection strategies.

The last point relates to the fact that within QTM sampling units (called *mesh elements* or *mels*, as shown in figure 2), applications are free to decide both which and how many points are selected. Unlike RDP — which only *accepts* points rather than *rejecting* them — our method allows either approach (even both at the same time), and uses additional parameters for increased (but also more complex) control over its operation.

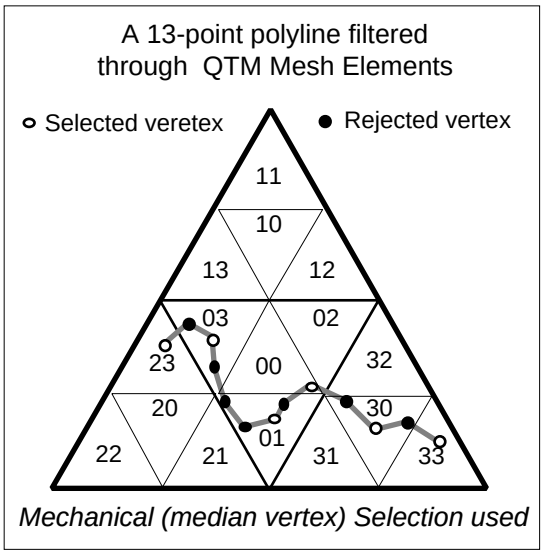


figure 2: Basic approach to filtering vector map data through QTM mesh elements; one line vertex per *mel* is retained at each level of detail (2-level QTM IDs are shown; *mels* are the small triangles).

Due to all its parameters and control strategies, a complete description of QTM-based line simplification would take more space than is available, (but see Dutton 1998, ch 4). Its flexibility also makes it complex to apply, and the results presented here are but a small (hopefully representative) sample of possible outcomes. Still, as they demonstrate effects of controlling for both scale and shape, we believe the ensemble of results has heuristic value, even if some geometric differences are difficult to measure or see. In quantifying results, we employ only a few simple statistics, and do not use standard cartometric generalization measures such as total vector displacement, total areal displacement, total curvilinearity, etc.; in our view, these global measures do not express generalization quality well enough to rely on them to automate quality control. So, mostly we depend on visual comparisons. This is clearly a limitation, and one that needs to be addressed in future studies.

QTM Resolution and Map Scale

How is it possible to associate a QTM level of detail with a map scale, as we do in our examples? Despite minor variations, there is a typical size for facet edges at each level of detail. We thus define *mean ground QTM resolution*, MGQR = 1.0e4 / 2^L, where L is a QTM level of detail and 1.0e4 is one-fourth of the circumference of the earth in km. For example, at level 20, MGQR = 1.0e4 / 2^{20} = 0.00953 km, ca. 10 m resolution. This derives from the fact that each QTM initial octant is a spherical triangle 90 degrees across, and each level of subdivision creates triangles half as wide. Calculation of a map scale uses this relation, and also requires making an assumption in order to relate ground distance to map distance. We use Tobler's observation that "... there is a relation between map scale and geographic resolution. This is readily obtained by observing that the smallest physical mark the cartographer can make is about one half millimeter in size." (Tobler 1988: 131). Thus we can define a constant, *mean map QTM resolution*:

MMQR = 0.5 mm,

and use it to compute scales at which MGQR maps to this value. The scale denominator, SD, at any level is therefore:

SD = MGQR/MMQR,

so that at level 20, (using units of mm),

SD = 9,530 / 0.5, or about 1:20,000.

We express such scales in round numbers because QTM resolution does vary spatially, and Tobler's constant is approximate and can vary circumstantially. Also, one might define scale fractions as averages between successive QTM levels; this would shift their magnitude by 25%.

Given this approach, we can derive the approximate scale at which source data were digitized. One way to do this is to encode the latitude and longitude coordinates of features into QTM IDs at a resolution known to be higher than the data actually possess, then to filter these identifiers at successively coarser resolutions (using the type of mechanical selection described below) until less than 98% of points remain. We then compute a scale factor from the QTM level prior to the one at which this process halted, and call this the source scale. There are more robust ways to make such estimates as well (see Dutton 1998, ch 3 and 5) that space does not permit discussing here. Note, however, that any QTM-based scale specification can only be accurate by a factor of around two.

Filtering line detail of QTM-encoded points can be done in various ways, some of which are simple and mechanical, while others involve more parameters and additional data. The basic method is shown in figure 2. There, the triangles represent the mesh elements that exist at the target (output) scale, which never move or change; they are fixed on the earth. A linear feature may enter and exit a *mel* one or more times; each such transit is called a *run* of vertices. Runs are identified by inspecting the sequence of QTM IDs along a feature at the target resolution; when a QTM ID occurs that contains the initial substring of specific quaternary digits identifying some *mel*, and this identifier occurs in two successive vertices, a run has commenced. A run lasts until some subsequent vertex no longer contains the *mel*'s identifier. Table 1 illustrates how runs are formed at different resolutions. If a feature leaves, then re-enters a *mel*, a new run is begun; however, such nearly adjacent runs in a given *mel* can be merged by setting a *lookahead* parameter that scans a specified number of vertices ahead to determine if the current *mel* is re-entered and combining the runs into one. As each *mel* will normally be scaled to the limit of resolution on the output map, we usually select only one vertex to represent the data within it (but more may be, if desired). Consequences of choosing more vertices depend on both local geometry and the value used for MMQR.

In its simplest form, our algorithm finds *the median point of each run* and marks all other points within the run for deletion. Middle points are chosen for consistency's sake and in order to space selected points relatively evenly. This default action, which we term *mechanical selection*, can be overridden in several ways, all of which depend on local information about line shape, as well as on parameters that express the user's selection criteria, discussed below.

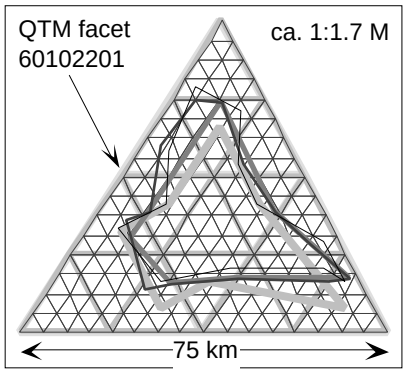


figure 3: Melville Island, Northern Australia, from WDB1 shown at1:1.17M; four mechanically-filtered versions (QTM levels 8, 9, 10 and 11) are displayed over the QTM level 11 grid.

A simple example of this process is shown in table 1 and figure 3, using data for Melville Island, Northern Australia from World Data Bank 1 (WDB1), compiled by the US CIA from source maps at 1:2M to 1:5M. The island is initially represented as a 15-point polygon; its

vertices were encoded into QTM IDs at level 13 (ca. 1:2.5M), and iteratively filtered, stopping when only one point remained. In table 1, QTM IDs of vertices rejected each level from 10 down to 7 are shaded, and the ones selected to represent runs are underscored. For this simple polygon, no filtering occurs until level 10 is reached, corresponding to about 1:20M scale. The island then reduces to 13 vertices at level 9 (ca. 1:40M), to 7 vertices at level 8 (ca. 1:80M), and contracts to a point at level 7 (ca. 1:160M). The first and last vertices of each polyline or polygon are always retained whether they are part of runs or not. Figure 3 shows the result of this mechanical selection from level 11 (thinnest line) down to level 8 (thickest line), registered to the local level 11 QTM graticule. At level 11 facet edges span about 5 km (the island is about 50 km across).

VTX	LONG	LAT	QTM ID[11]	QTM ID[10]	QTM ID[9]	QTM ID[8]	QTM ID[7]
1	130.06	-11.84	601022012123	60102201212	6010220121	601022012	60102201
2	130.00	-11.85	601022012121	60102201212	6010220121	601022012	60102201
3	129.97	-11.74	601022012323	60102201232	6010220123	601022012	60102201
4	129.97	-11.70	601022012301	60102201230	6010220123	601022012	60102201
5	130.01	-11.68	601022010301	60102201030	6010220103	601022010	60102201
6	130.03	-11.53	601022011301	60102201130	6010220113	601022011	60102201
7	130.11	-11.42	601022011122	60102201112	6010220111	601022011	60102201
8	130.18	-11.43	601022011031	60102201103	6010220110	601022011	60102201
9	130.22	-11.57	601022011210	60102201121	6010220112	601022011	60102201
10	130.29	-11.68	601022010201	60102201020	6010220102	601022010	60102201
11	130.36	-11.75	601022013212	60102201321	6010220132	601022013	60102201
12	130.45	-11.83	601022013310	60102201331	6010220133	601022013	60102201
13	130.41	-11.84	601022013013	60102201301	6010220130	601022013	60102201
14	130.21	-11.83	601022010132	60102201013	6010220101	601022010	60102201
15	130.06	-11.84	601022012123	60102201212	6010220121	601022012	60102201

table 1: Geographic and QTM encoding for vertices defining Melville Island, showing effects of QTM simplification via mechanical selection. Underlined vertices are those selected to represent runs (shaded ones sharing the same QTM ID).

Note that some retained vertices in figure 3 drift from their original locations; this was a deliberate choice, to illustrate that vertex locations can be controlled by decoding them at their filtered resolution (i.e., truncating precision), placing them at the center of the QTM facets comprising that level of detail. In the maps shown in figures 8 – 13, this was not done, however. Instead, the full QTM precision used when encoding QTM IDs from geographic coordinates was maintained, preventing any vertex drift. Melville Island's vertex movement is a form of *typification* that regularizes point placement to conform to a triangular grid having a certain resolution. This is controlled by one of the parameters that can be set when filtering QTM-encoded line detail; its default value is the QTM encoding level (full spatial precision), as in figures 8–13.

Sinuosity-guided Point Selection

In addition to point placement, QTM filtering can give users contol over point selection. The type of information we employ to guide point selection consists of point-by-point estimates of local *line sinuosity* (a statistic similar to a *route factor*, also akin to fractal dimension). This statistic is estimated for each point of a linear feature according to the method described in figure 4, scaled to an integer range (seven classes were used in this study), and stored in unused space in each point's QTM ID.

This measure of sinuosity is simply the ratio of distance along a digitized line to the length of the trend line connecting its endpoints (also called a *route factor*), but may be averaged across various ranges. By varying the parameters *m* and *n*, one can specify the locality within which sinuosity is estimated; *n* is the topological distance from a given vertex to the closest neighbors between which a trend line is constructed, and *m* gives the farthest neighbors between which to analyze. At a minimum, ($n = m = 1$), only the given point and its immediate neighbor on either side will be used. As *m* increases, more points are drawn into this collection, and the computed sinuosity represents a larger linear region. As *n* grows, the influence of the point and then its neighbors is removed from the calculation, an option which may not seem to make sense, but can help to determine if the point's sinuosity is typical of its surroundings.

The sinuosity values (*SV*) calculated by this method are based on Cartesian or great-circle distances, and are dimensionless real numbers which range from 1 to no fixed upper limit, but will rarely exceed 2. We then classify them into discrete classes using a non-linear transformation, the top curve in figure 5. The horizontal bands in figure 5 show how *SV* would map to three sinuosity classes, using four related classifiers. While no theoretical justification can be given for using it, the transform $1–(1/SV^2)^{1/2}$ (the top curve) is able to discriminate low sinuosities better than others. As the representative configurations at the bottom of figure 5 show, most local values of SV are likely to fall below 1.5, which calls for greater discrimination below that value than above it. The examples of shoreline classification shown in figure 6 indicate that this classifier seems to be reasonable.

For our experiments we computed two different sinuosity estimates for each vertex and stored them separately in its respective QTM ID. The first value is a local estimate, with $m = 1$ and $n = 2$ (averaging sinuosities across 3 and 5 points each time). The second estimate used $m = 2$ and $n = 7$, which averaged 6 sinuosities around each point but excluded its immediate two neighbors. We identify the first set of sinuosity values as *CSL* (classified sinuosities, local) and the second set as *CSR* (classified sinuosities, regional), and used seven classes for each, so that their values range from 1 to 7. We then examined the effects of using *CSL* and *CSR* to select points for generalization, exploring three basic strategies. For each run within a *mel*:

1a. Select points having the *minimum* absolute difference between *CSL* and *CSR or*

1b. Select points having the *maximum* absolute difference between *CSL* and *CSR*

2. Select points having a *CSL closest* to a user-specified *CSL*, either 1 or 7

3. Select from whatever candidate points remain the one nearest the median point.

Either step 1a or 1b or neither is taken, followed optionally by step 2, and ending with step 3 whenever more than one point remains to be eliminated from a run. Should only one candidate point per run remain after any step, subsequent steps will not be executed. The difference between 1a and 1b is that in 1a points having similarities typical of their regions are selected, while in 1b, points most different in sinuosity from their regions are chosen. In step 2, users specify a *CSL* value that selection should prefer, should choices remain (we term this parameter *preferred sinuosity value*, or *PSV*); any value within range can be given, but we always

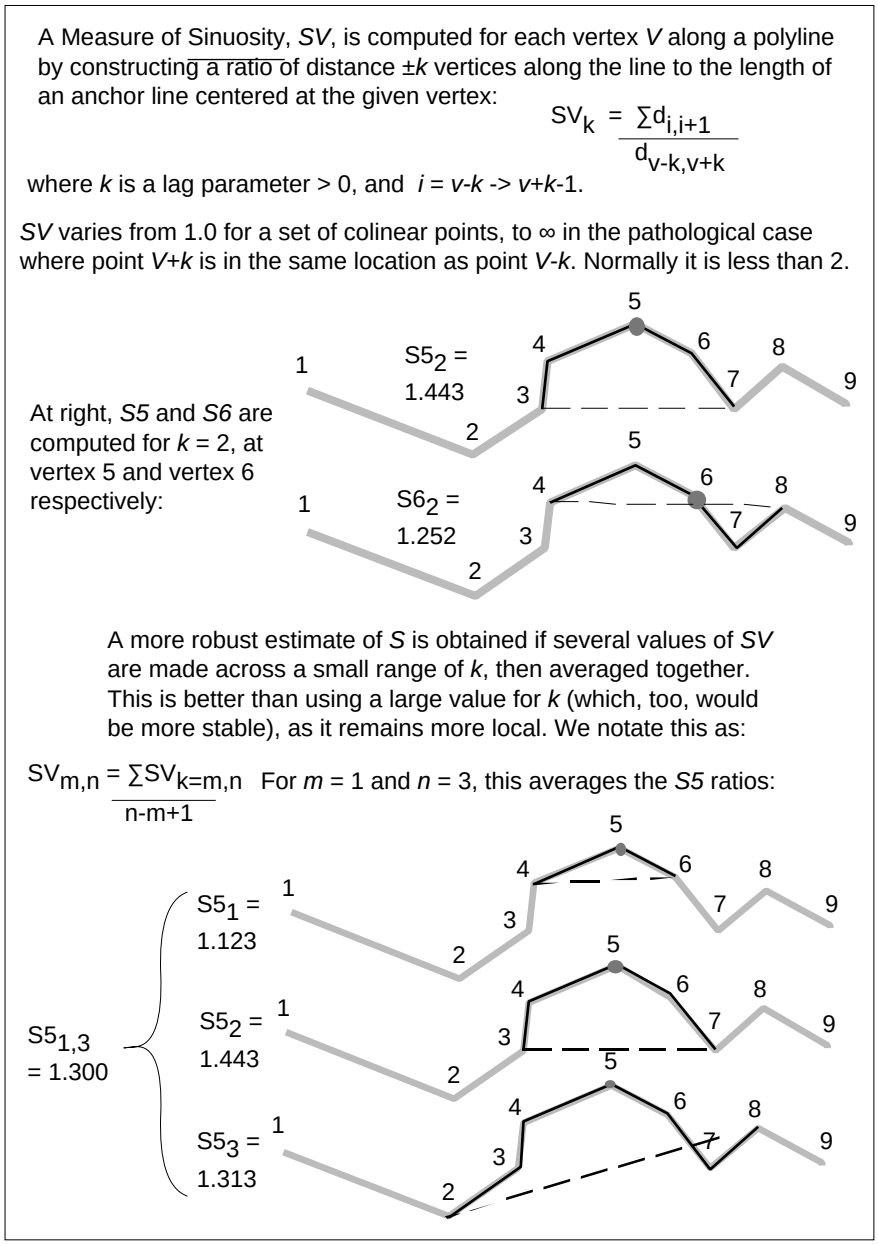
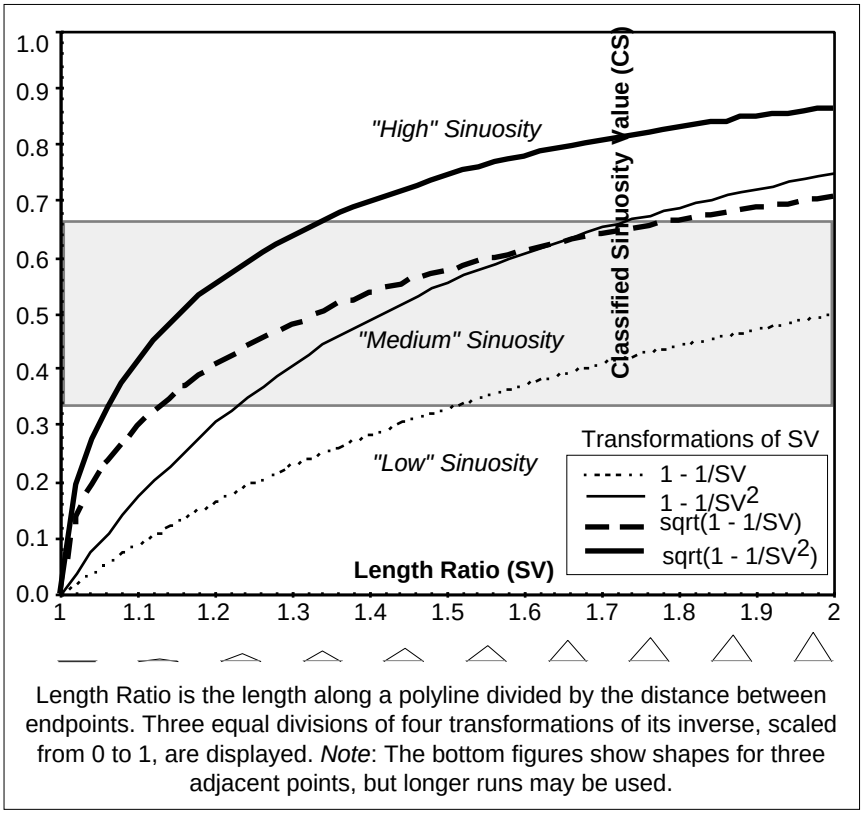


figure 4: Method for computing local and regional sinuosity statistics for polyline vertices; values of SV are then scaled to an integer range

used the extremes in our tests. The algorithm requires from one to five passes through each run, depending on which steps are executed; steps 1 and 2 each need two passes; first to find the value to prefer, next to select the point(s) possessing it. However, all computations are simple integer or logical operations.



Length Ratio is the length along a polyline divided by the distance between endpoints. Three equal divisions of four transformations of its inverse, scaled from 0 to 1, are displayed. *Note*: The bottom figures show shapes for three adjacent points, but longer runs may be used.

figure 5: Four classifiers for raw length ratios (SVs), mapped to three categories (CVs)

Empirical Generalization Experiments

Tests were performed on two island shorelines, each of which had been digitized at two different scales. For each of the four datasets at each of three derived scales, seven different combinations of the above parameters were tested, in addition to the RDP algorithm, for a total of eight generalized maps per set. Including the depictions of source data, the resulting figures contain a total of 108 maps (not counting scaled insets) arrayed in a series of 3x3 grids. Even these experiments do not begin to exhaust the combinations, but they do communicate a sense of possibilities.

One of the two test areas chosen is an island off the east coast of the U.S., the other from its west coast. The East coast island is Nantucket, Massachusetts, the largest of a group of three islands about 40 km south-east of Cape Cod in the Atlantic. These islands have no bedrock, being stabilized sand, sculpted by waves and tides. The latter is Bainbridge Island, 5,000 km away, in Puget Sound about 10 km west of Seattle, Washington. It is in an area of more complex geology, and has rocky undulating shorelines with deep tidal inlets. Basic statistics for the four datasets are presented in table 2 and the more detailed (1:80K) datasets are shown in figure 6 at 1:250,000. Except for a small cape on its north-east extreme, Bainbridge does not have the prominent spits and bars that characterize Nantucket. Neither does Nantucket have the kind of deep-cut estuaries found along the Bainbridge shore.

All data is from sources prepared by the U.S. National Ocean and Atmospheric Administration (NOAA), and obtained through a USGS facility in Woods Hole, MA.[6]

The 1:80K shoreline were manually digitized from nautical charts. However, some areas, particularly near harbors, were digitized from insets (up to 1:10,000, but not documented). We do not know if insets were traced for our examples, hence assume their scale is 1:80,000. The 1:250,000 data is from the World Vector Shoreline (WVS) file, vectorized from radar imagery. This is apparent from the somewhat aliased appearance of the WVS linework. In the NOAA source files, the islands were represented as collections of arcs. Our software linked these into complete polygons prior to generalization, in order to avoid having to retain arc endpoints (spurious nodes).

Organization of Experiments

Using the QTM scale relations described previously, three target scales were selected for simplifying files. These were at the lower end of the useful range of the data, in order to be able to discriminate between methods; at larger scales method results tend to look more alike, especially when maps are reduced to scale. Coordinates of the two 1:80K datasets were converted to QTM identifiers at QTM level 20 (ca. 10 m resolution), and the 1:250K datasets were converted to QTM level 18 (ca. 40 m resolution) to ensure that point displacement would not be introduced. Generalization was then performed to QTM levels 14, 13 and 12 (1:1.2M, 1:2.4M and 1:4.8M, respectively), varying several other

[6] *Coastline Extractor* (http://www.crusty.er.usgs.gov/coast/getcoast.html).

parameters. These other parameters did not affect the number of retained vertices, only which ones were selected. Then the RDP algorithm was applied to the source data, using tolerances that yielded the same number of points as the QTM generalizations. These tolerances are shown at the upper right of the RDP-generalized features in figures 8–13. QTM filtering facet size is also shown, as a set of six triangles in the upper left corner of each figure. The 1:250K data usually required different RDP tolerances than the 1:80K data to match the QTM point counts, even though the total number of selected points were fairly similar for the two source scales at each level of detail. This illustrates the general difficulty of choosing RDP tolerance values to achieve a particular degree of reduction.

Discussion of Results

Figures 8-13 are all arranged similarly, with 1:80K source data generalizations on the left and 1:250K source data on the right. Each figure shows generalizations to a certain target scale, either 1:1.2M, 1:2.4M, or 1:4.8M, with eight variations. Within each group, the source version is shown in the upper left, and next to it is the default QTM generalization (obtained by sampling median points of runs without consulting computed point sinuosity attributes). To the right of that is the RDP generalization and its tolerance, producing the same number of points as the QTM filtering yielded. Note that the number of generalized points for 1:250K sources is fairly close to the number of points derived for the 1:80K data. This is to be expected, due to selection through a fixed grid, assuming the two source versions are properly geo-positioned and are not too dissimilar to start with.

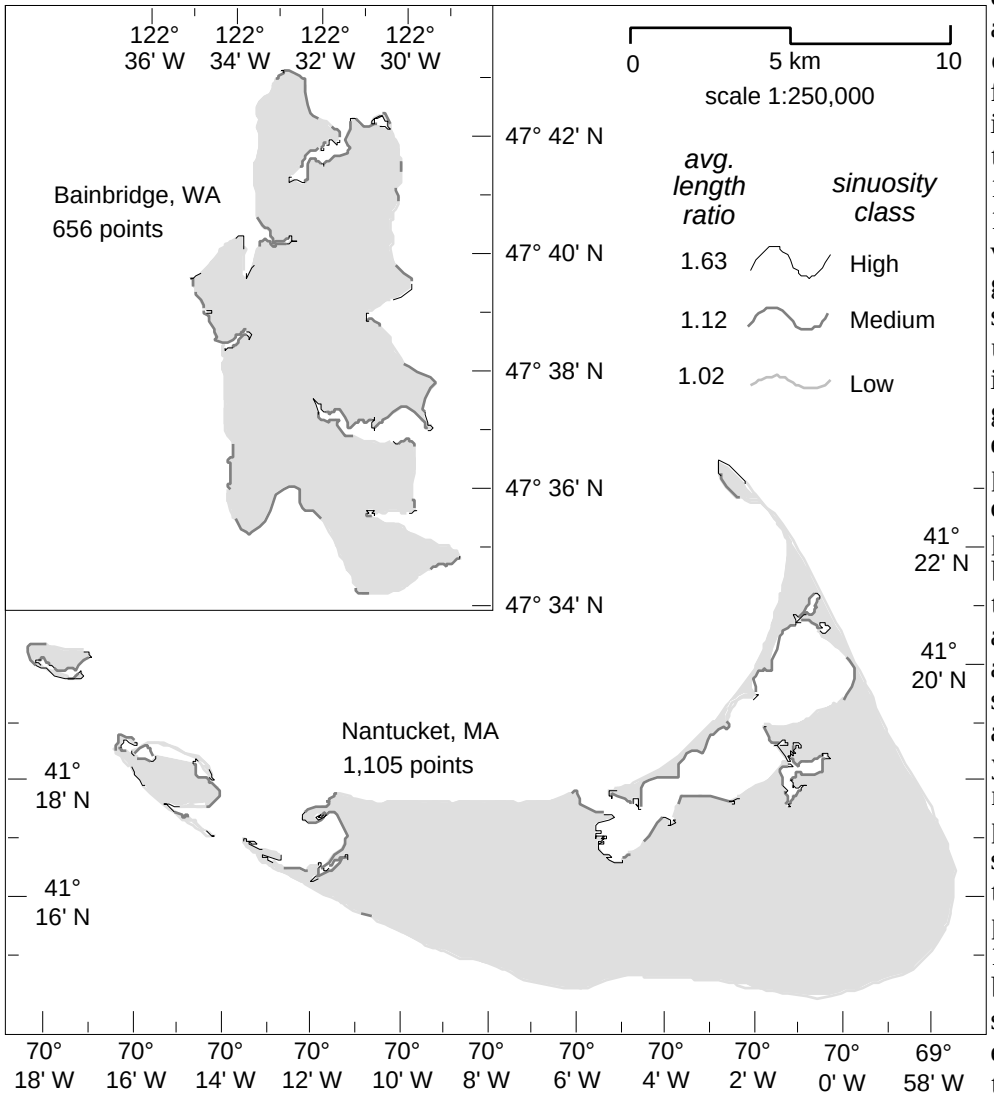


figure 6: Nantucket and Bainbridge Islands, 1:80K source, mapped at 1:250K; shorelines are symbolizes to illustrate results of a 3-classs sinuosity analysis

	Nantucket MA		Bain
Digitizing Scale	1:80,000 or >	1:250,000	1:80
Area (km2)	119.90	131.68	70.0
Perimeter (km)	124.97	102.54	77.4
N/S Length (km)	17.06	17.50	16.4
E/W Length (km)	22.67	22.00	8.59
Avg. Seg. (km)	0.14	0.20	0.12
No. of Vertices	867	525	656
Mean Sinuosity	3.70 (out of 7)	3.01 (out of 7)	3.29

table 2: Statistics for Nantucket MA and Bainbridge WA test data

Each of the nine versions has its average sinuosity printed below it (on a scale of 1 to 7), and each generalized version is reproduced as an inset at target scale to the lower left of the enlarged feature. These insets are essential for interpreting results of methods, as reduction of scale is usually the principle cause for generalization, and results need to be legible at target scales. To aid interpretation, boundaries of the enlarged versions are thickened to reflect the line weight of the target scale figures (0.4 mm) as if photographically enlarged to presentation scales (1:350K for Bainbridge, 1:500K for Nantucket).

The lower two rows of generalized features are variations of QTM filtering that vary several sinuosity-related parameters. The middle row contains features for which points were selected that had local sinuosity values (CSL) *closest* to their regional ones (CSR), *so that selected points are representative of their localities*. Conversely, the bottom row contains features having selected points with CSL *farthest* from CSR, meaning *that the points tend to be unrepresentative of their localities*. It is not always easy to see the differences, especially at larger scales, where there are fewer choices available, because *mels* (hence runs) are smaller.

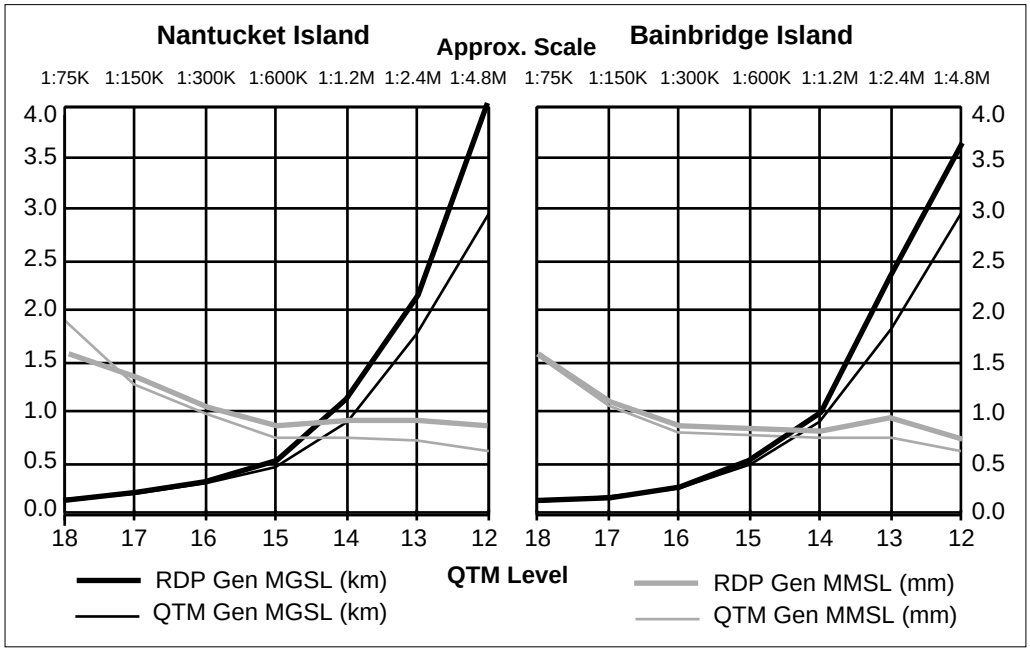


figure 7: Nantucket and Bainbridge Islands: Mean Ground and Map Segment Lengths (MGSL and MMSL) by QTM level of detail for QTM and RDP generalization methods

The three columns of the two bottom rows show the effects of specifying a sinuosity value for breaking ties between points once the minimum/maximum similarity criterion had been applied. That is, having identified all the points in a run whose CSL was most dis/similar to their CSR, the user then can select those which have a CSL closest to a particular preferred value (PSV). If a PSV is not specified (column 1), median selection mechanically picks the point nearest the middle of the run. If a low value of PSV is specified (column 2), the resulting outline ought to be smoother, at least in the middle row. In the lower row, this logic tends to reverse, as the set of remaining points may tend to mitigate the effects of PSV. In the third column, specifying a large PSV usually results in higher angularity, and the accompanying sinuosity statistics tend to bear this out. Note that the average sinuosity is computed using grouped data (with seven bins), and that it represents the sinuosities of the points selected from the source data, rather than a re-computation for the generalized versions; were sinuosities to be re-computed from only the selected points, they would tend to increase, but still maintain their relative magnitudes. Such increases are due to the local averaging process, the range of which extends farther as the number of coordinates decreases, and (in the case of RDP) because higher-sinuosity points tend to be retained. QTM filtering, on the other hand, may result in lowering sinuosity. The graphic results indicate this can be desirable, as it tends to produce simplifications that work better at the target scale, even though some shape distortions are evident.

Quantitative results from the tests do not truly capture the aesthetic qualities of the graphical results. Some of the representations found in figures 8-13 look inadequate at enlarged scale, but at their intended scale are much more satisfactory. By the same token, some that seem to work at enlarged scale are too detailed when reduced to target scale. This is in spite of the fact that our primary scale-related metric, Mean Map Segment Length (MMSL) indicates that in principle, legibility is roughly constant from about 1:150K to 1:4.8M (it is, after all, only an average). Figure 7 is a plot of MMSL and MGSL for both examples, and shows results of both RDP and default QTM simplification (mechanical selection). While RDP results in

slightly longer segments at smaller scales, the difference is marginal and probably does not affect the perception of maps (recall that RDP tolerances were set to yield the same number of points as QTM generalizations did). What seems to matter more is RDP's attraction to more sinuous points, as would be expected from its algorithmic description.

At the two largest scales shown in figure 7, 1:75K and 1:150K, MMSL is more than 1 mm, indicating that these versions could have additional detail without compromising legibility. Note that both methods exhibit the same behavior in this range, and for both examples. To flatten out the MMSL curves, one could change generalization parameters for these scales: lower the RDP tolerance or select more than one point per QTM *mel*. But as the degree of flattening required seems data-dependent, it might be difficult to design a robust heuristic for this. As maps at these larger scales would be too large to display here, we did not pursue such strategies for this report.

Independent of this, there is also the possibility that the QTM scale relations are not calibrated correctly in several respects: (1) all scale denominators may be too high by some fixed ratio; (2) a non-linear adjustment to them is required. The first is quite possible, as it depends on the value used for MMQR (0.5 mm in this study), and on whether scale fractions are assigned to QTM integer levels or half way between them. Regarding the second point, the fact that data volumes resulting from QTM filtering do not always double at each level is not in strict agreement with the linear version of the Radical Law. We feel that this probably relates to the geometric complexity of lines, and points to a possible re-definition of the Radical Law adjusting for fractal dimensionality, along the lines proposed by Muller (1987). As estimation of fractal dimensionality of QTM-encoded features is quite straightforward (Dutton 1998, ch. 3, but note not all features are self-similar enough to reliably estimate a dimension for them) it would be possible to test Muller's hypothesis that for linear feature reduction the exponent x of the Radical Law should be the fractal dimension of the line rather than unity. We suspect that such modeling will largely explain rates of coordinate volume decrease.

Conclusions

We admit that our experimental findings are not easy to interpret. In figures 8 – 13, some graphic qualities (such as simplicity) may occur at the expense of another (such as shape fidelity). The more accurate versions of a map can have some excessive detail (indicated by length and average sinuosity). Some otherwise convincing ones may have bays and promontories filled in, suggesting the need for internal point displacement. QTM filtering doesn't do it all, nor is it intended to. Although interpretations can vary, our 108 experiments suggest that points which would not be regarded as globally salient may be better candidates for selection. Even though such points may not be visually prominent ones, they still can be more typical of local line configurations; the nature of their localities rather than their global gestalt determines their cartographic utility. To exploit this, places where shape complexity changes need to be identified in databases by applying feature segmentation techniques. Measures of shape computed at and stored with vertex data seem to be effective way to achieve this. Hierarchical location codes such as QTM IDs provide an efficient way to organize locational data and metadata for map generalization.

To recap our findings, section by section,

1. Influences of map scale and point character on map generalization were studied;
2. Scale-specificity is an important factor that is glossed over in many previous studies;

   2.1 Point densities after generalization can be predicted using the *linear* Radical Law;

   2.2 Mean distance between points is a useful heuristic and a function of scale;
3. Globally salient points may not be the most significant ones for generalization;

   3.1 The Ramer-Douglas-Peucker algorithm operates globally (on entire lines as a whole);

   3.2 Setting RDP's tolerance value to achieve scale-appropriate generalizations is difficult; most other algorithms share this shortcoming

   3.3 Existing measures of generalization performance do not capture its visual quality;
4. QTM filtering, a finite-resolution, local detail reduction technique is described;

   4.1 A practical way to link spatial resolution and map scale is discussed;

   4.2 A classified measure of local line sinuosity is described for qualifying points;
5. Objectives and data for empirical analysis of generalization effects are presented;

   5.1 The experimental design is explained;

   5.2 Results are presented in a canonical way to isolate effects of parameters;
6. We hope readers will take time to examine the results and decide for themselves how helpful it is to use sinuosity information for point selection.

We believe this study underscores the importance of evaluating generalizations of map data at target scale, not just in comparison to source data. Until "objective" measures of generalization performance become more reliable, many of these evaluations will need to be done visually and subjectively. It seems clear that features will change shape as they simplify; this is the inevitable consequences of scale change, unless one is viewing a mathematical abstraction such as a fractal curve. And even when map features happen to be self-similar, this does not in itself require generalizations of them to resemble their original shapes in all respects at all scales. Instead, like melting ice sculptures or images on fragments of holograms, sinuous shapes on maps should gracefully degrade as their medium shrinks, their character and importance draining away into their context.

Acknowledgement

This research was funded by Swiss National Science Foundation grants 2000-037649.93 and 2100-43502.95 to the Dept. of Geography, University of Zürich. Sincere thanks to reviewers for helpful comments and valuable suggestions.

References

Attneave, F., 1954. *Some informational aspects of visual perception*. Psychological Review . 61(3): 183-193.

Barber, C., Cromley, R. and Andrle, R., 1995. *Evaluating alternative line simplification strategies for multiple representations.* Cartography and Geographic Information Systems , 22(4): 276-290.

Beard, M.K., 1991. *Theory of the cartographic line revisited: Implications for automated generalization*. Cartographica . 28(4): 32-58.

Buttenfield B. P., 1986. *Digital definitions of scale-dependent line structure*. Proc. Auto-Carto London. London: Royal Inst. of Chartered Surveyors, 497-506.

Cromley, R.G. and Campbell, G.M., 1990. *A geometrically efficient bandwidth line simplification algorithm*. Proc. 4th Int. Symp. on Spatial Data Handling , Zürich, August 1990, vol. 1, 77-84.

Cromley, R.G. and Campbell, G.M., 1992. *Integrating quantitative and qualitative aspects of digital line simplification*. The Cartographic Journal , 29(1): 25-30.

Douglas D.H., Peucker T.K., 1973. *Algorithms for the reduction of the number of points required to represent a digitized line or its caricature.* The Canadian Cartographer 10(2): 112-22.

Dutton, G. (ed), 1973. *Size and shape in the growth of human communities.* special issue of Ekistics. 36:215, 233-308.

Dutton, G. and Buttenfield, B.P., 1993. *Scale change via hierarchical coarsening: Cartographic properties of Quaternary Triangular Meshes.* Proc. 16th Int. Cartographic Conference. Köln, Germany, May 1993, 847-862.

Dutton, G., 1997a. *Digital map generalization using a hierarchical coordinate system* Proc. Auto Carto 13. (Seattle, WA) Bethesda, MD: ACSM/ASPRS, 367-376.

Dutton, G.,1998. *A hierarchical coordinate system for geoprocessing and cartography.* Lecture Notes in Earth Science . Berlin: Springer. 230 p.

Goodchild, M.F. and Yang Shiren 1992. *A hierarchical data structure for global geographic information systems.* Computer Graphics, Vision and Image Processing , 54(1): 31-44.

Gould, S.J., 1966. *Allometry and size in ontogeny and phylogeny* Biological Reviews 41: 587-640. excerpted in Dutton (1973), 253-262.

Jasinski, M.J.,1990. *The comparison of complexity measures for cartographic lines.* Santa Barbara, CA: NCGIA Tech. Paper 90-1, 73 pp.

Jenks, G.F., 1981. *Lines, computers and human frailties.* Annals of the Association of American Geographers 71(1): 1-10.

João, E.M., 1998. *Causes and consequences of map generalization.* Research Monographs in GIS . London: Taylor & Francis, 266 p.

Li, Z. and Openshaw, S., 1992. *Algorithms for automated line generalization based on a natural principle of objective generalization.* Int. J. of GIS 6(5): 373-390.

Li, Z. and Openshaw, S., 1993. *A natural principle for the objective generalization of digital maps.* Cartography and Geographic Information Systems 20(1): 19-29.

Marino, J.S., 1979. *Identification of characteristic points along naturally occurring lines: An empirical study*. The Canadian Cartographer 16: 70-80.

McMaster, R.B., 1986. *A statistical analysis of mathematical measures for linear simplification.* The American Cartographer 13(2): 103-116.

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

McMaster, R.B. and Shea K.S., 1992. Generalization in Digital Cartography . Washington, DC: Association of American Geographers, 134 p.

Moellering, H. and Rayner, J.N., 1982. *The dual axis Fourier shape analysis of closed cartographic forms*. The Cartographic Journal 19: 53-59.

Muller, J-C., 1987. *Minimum point density and compaction rates for the representation of geographic lines.* Proc. Auto Carto 8 . (Baltimore, MD) Bethesda, MD: ACSM/ASPRS, 221-230.

van Oosterom, P. and Vijlbrief, T., 1997. *The spatial location code*. in Kraak, M.J. and Molenaar, M. (eds) Advances in GIS Research II (Proc. SDH7, Delft, The Netherlands). London: Taylor & Francis, 101-117 .

Plazanet, C., Affholder, J-G., Fritsch, E., 1995. *The importance of geometric modeling in linear feature generalization*. Cartography and Geographic Information Systems 22(4): 291-305.

Ramer, U., 1972. *An iterative procedure for the polygonal approximation of plane curves*. Computer Graphics and Image Processing 1: 244-56.

Tobler, W., 1988. *Resolution, resampling and all that*. Building Databases for Global Science . Mounsey, H and Tomlinson, R. (eds.). London: Taylor & Francis, 129-137.

Töpfer, F. and Pillewizer, W., 1966. *The principle of selection*. The Cartographic Journal 3: 10-16.

Töpfer, F., 1974. Kartographische Generalisierung. Leipzig: VEB Hermann Haack,336 p.

Van Horn, E., 1985. *Generalization of Cartographic Databases*. Proc. Auto Carto 7 . (Baltimore, MD) Bethesda, MD: ACSM/ASPRS, 532-540.

Visvalingam, M. and Whyatt , J.D.,1990. *The Douglas-Peucker algorithm for line simplification: Re-evaluation through visualization*. Computer Graphics Forum 9(3): 213-228.

Wang, Z. and Muller, J-C., 1998. *Line generalization based on analysis of shape characteristics*. Cartography and Geographic Information Systems 25(1): 3-15.

White, E.R., 1985. *Assessment of line-generalization algorithms using characteristic points*. The American Cartographer 12(1): 17-27.

Zhan, F.B. and Buttenfield, B.P., 1996. *Multi-scale representation of a digital line*. Cartography and Geographic Information Systems . 23(4): 206-228.

Zycor, Inc., 1984. *Manual of automated feature displacement*. Report for U.S. Army Engineering Topographic Laboratories , Fort Belvoir, VA, 204 p

Figure Captions

The following figures should be collected at the end of the paper

figure 8: Nantucket Island: 1:80K (left) and 1:250K (right) coastline data simplified to 1:61.2M (QTM level 14) using RDP and 8 variations of QTM filtering

figure 9: Nantucket Island: 1:80K (left) and 1:250K (right) coastline data simplified to 1:2.4M (QTM level 13) using RDP and 8 variations of QTM filtering

figure 10: Nantucket Island: 1:80K (left) and 1:250K (right) coastline data simplified to 1:4.8M (QTM level 12) using RDP and 8 variations of QTM filtering

figure 11: Bainbridge Island: 1:80K (left) and 1:250K (right) coastline data simplified to 1:1.2M (QTM level 14) using RDP and 8 variations of QTM filtering

figure 12: Bainbridge Island: 1:80K (left) and 1:250K (right) coastline data simplified to 1:2.4M (QTM level 13) using RDP and 8 variations of QTM filtering

figure 13: Bainbridge Island: 1:80K (left) and 1:250K (right) coastline data simplified to 1:4.8M (QTM level 12) using RDP and 8 variations of QTM filtering