

SHAPE DEFORMATION IN CONTINUOUS MAP GENERALIZATION

JEFF DANCIGER, SATYAN L. DEVADOSS, JOHN MUGNO, DON SHEEHY, AND RACHEL WARD

ABSTRACT. Given a collection of regions on a map, we seek a method of continuously altering the regions as the scale is varied. This is formalized and brought to rigor as well-defined problems in homotopic deformation. We ask the regions to preserve topology, area-ratios, and relative position as they change over time. A solution is presented using differential methods and computational geometric techniques. Most notably, an application of this method is used to provide an algorithm to obtain cartograms.

1. INTRODUCTION

Cartographic generalization strives to represent geographical information on a map whose specifications are different from those of the original dataset. The ubiquitous use of computers to handle geographic information has led to the call for automation in map generalization. This demand has been further strengthened by the prevalence of the internet, with its access to map-based websites. Based on pre-generalized datasets, such sites allow users to zoom in and out of regions of maps at certain predetermined levels. Thus, most approaches to this problem have been *discrete* in nature, with several transitional jumps from one scaling to the next, with operators performing local changes to accomplish the generalization task. It is natural to demand the ability to change the level of detail on such systems in a smooth fashion, rather than the discrete jumps which often portray the current approach. This introduces the notion of *continuous* cartographic generalization.

The interest in this field has started to grow over the past decade; Jones and Ware [7] furnish a nice overview. Yang and Gold [16] provide a systems approach to automated map generalization. Issues and algorithms with regards to area partitioning during continuous generalization using hierarchical schemes is studied by van Oosterom [14]; van Kreveld [13] presents concepts in visualization which minimize the discrepancy that occurs when discrete changes are made. Recently, Sester and Brenner [11] offer a possible solution to continuous generalization when restricted to small mobile devices.

The ideas behind generalization, without even introducing a continuous parameter, are quite complex. They involve numerous operations such as simplification, exaggeration, elimination, and displacement, along with several issues such as features, label placements, and line thickness. A considerable amount of effort has been exerted in the GIS community in this area, making it overly ambitious to discuss the full scope of this multi-faceted problem. Indeed, the problems surrounding the subject of scale change itself is quite difficult and involved; a pleasant survey is supplied by Lam et al. [8]. Our interest is restricted to one particular aspect of this, notably shape.

This paper focuses on the deformation of the shapes of regions in a map during the process of continuous scale change. The scale change transformation is given in the language of continuous, rather than discrete mathematics. There are three key objectives for this work:

- (1) We model certain issues concerning shape in generalization from a mathematical viewpoint. In particular, we extract properties of the map that are valuable such as area-ratios, topology and adjacency, and recast it in a rigorous setting.
- (2) We extend these ideas to address certain problems of shape deformation in *continuous* scale change using homotopic tools. In particular, a treatment of continuous, multi-scale representation of spatial data is given.
- (3) We supply solutions to the problems, giving theoretical proofs of the results, as well as providing some algorithmic solutions.

One can argue that discrete models which are classically used in GIS should also be sufficient for continuous scale change. We do not oppose such a view; indeed, most of spatial data is stored in discrete databases. However, one purpose of this work is to introduce new approaches and new tools which could possibly pave the way for novel approaches to this difficult topic. The strength of theoretical mathematics lies in the continuous language, and we present a means of harnessing part of it in this context.

2. DEFORMATIONS USING FUNCTIONS

2.1. As our starting point, we choose and fix one largest-scale map drawing Ω at some scale s . This map has several regions drawn on it. During the deformation which takes place as scale continuously changes, the regions which are within Ω will expand and take available map space while satisfying certain requirements; Saalfeld [10] presents a nice discussion of issues in such map space management. For scaling values near s , there is a large amount of map space available for the regions to occupy; as the scaling decreases, the regions grow, using up the surrounding map space. It is clear the *shape* of the regions will radically deform during scale change; our goal is to understand *how* these regions within Ω will deform. Before proceeding further, we need to properly define the objects of study. We note that almost all of the mathematics background needed can be found in a standard topology text such as Hatcher [5].

Definition. A *region* is a connected subset of the plane bounded by a closed curve which does not intersect itself.

In other words, a region is a generalization of the notion of a polygon. We choose and fix some region Ω at some scale which contains a collection of pairwise disjoint regions $\{R_1, R_2, \dots, R_n\}$ within it. FIGURE 1(a) shows an example, where the base region Ω is the square, with the three regions A , B , and C contained within Ω . In order to present a rigorous, unified approach, a novel method is introduced to study the regions.

Let \mathcal{R} denote the union of the regions $R_1 \cup \dots \cup R_n$. Then, the set \mathcal{R} in Ω can be considered as the *image* of a continuous function $\phi : \mathcal{R} \rightarrow \Omega$ which maps each point x in \mathcal{R} to its corresponding point $\phi(x)$ in Ω . The function ϕ is called the *inclusion function* (it is sometimes called the *identity* function since it does not alter the image of \mathcal{R} .) FIGURE 1(b) shows the mapping of the example given in part (a). Under this language, we claim that scale changes correspond to appropriate choices of continuous functions $f : \mathcal{R} \rightarrow \Omega$. FIGURE 1(c) shows another continuous function which has deformed the images $f(\mathcal{R})$ of the regions \mathcal{R} within Ω . To do this correctly, however, we must have our functions $f : \mathcal{R} \rightarrow \Omega$ satisfy certain properties.

Remark. The key advantage to this set-up is that the original regions \mathcal{R} are *not* altered during the deformation, only their *images* under the functions. Thus, one can keep track of any collection of points in \mathcal{R} throughout the deformation by simply tracking their images.

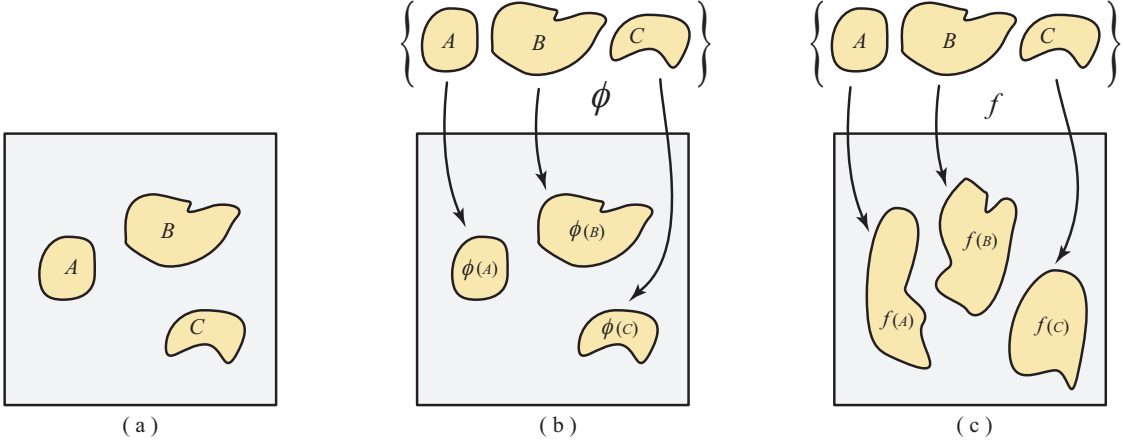


FIGURE 1. (a) The base region Ω is the square containing the three regions $\mathcal{R} = A \cup B \cup C$. (b) The inclusion function ϕ taking \mathcal{R} to their corresponding images. (c) A continuous function deforming the regions \mathcal{R} within Ω .

Notation. We use the symbol ∂ to denote the boundary of an object. Let $S_i = \partial R_i$ be the boundary curve of region R_i , and let \mathcal{S} denotes the union of all the curves $\{S_i\}$. Moreover, let $\mathcal{S}_\partial = \mathcal{S} \cup \partial\Omega$.

Definition. The function $f : \mathcal{R} \rightarrow \Omega$ *preserves topology* if $f(R_i \setminus S_i)$ is disjoint from $f(R_j \setminus S_j)$ for all i, j . In other words, regions are allowed to meet along their boundary but not within their interior.

Definition. The function $f : \mathcal{R} \rightarrow \Omega$ is *area-ratio preserving* if for all i, j ,

$$\mu(f(R_i)) \cdot \mu(R_j) = \mu(f(R_j)) \cdot \mu(R_i),$$

where $\mu(\cdot)$ denotes the area of the region.

2.2. The properties above each have the aim of preserving certain features of the original set of regions during scale changes. A final property we impose is to maintain the *relative position* of regions. Here, regions that are “close” to begin with are expected to remain close after the transformation. A natural tool to encapsulate this is the Voronoi diagram; the reader can turn to [9] for details. The *Voronoi region* V_i of R_i is the set of all points which are closer (or equally close) to S_i than to $\mathcal{S}_\partial \setminus S_i$; in other words,

$$(2.1) \quad V_i = \{x \in \Omega \mid \text{dist}(x, S_i) \leq \text{dist}(x, \mathcal{S}_\partial \setminus S_i)\}.$$

The boundaries of the Voronoi regions consist of *Voronoi edges* (equidistant from two regions) and *Voronoi vertices* (equidistant from three or more regions). The *Voronoi diagram* of \mathcal{R} in Ω is the union of the boundaries of the Voronoi regions. The Voronoi diagram partitions Ω into connected regions, one for each region R_i , and one associated to the boundary curve $\partial\Omega$.

Definition. The *adjacency graph* G_f of the function $f : \mathcal{R} \rightarrow \Omega$ is the dual graph to the Voronoi diagram of $f(\mathcal{R})$ in Ω : Each Voronoi region is associated to a vertex of G_f , where two such vertices share an edge in the adjacency graph if and only if they share a Voronoi edge.

FIGURE 2 shows (a) a map with five regions, along with (b) its six Voronoi regions and (c) the corresponding adjacency graph.

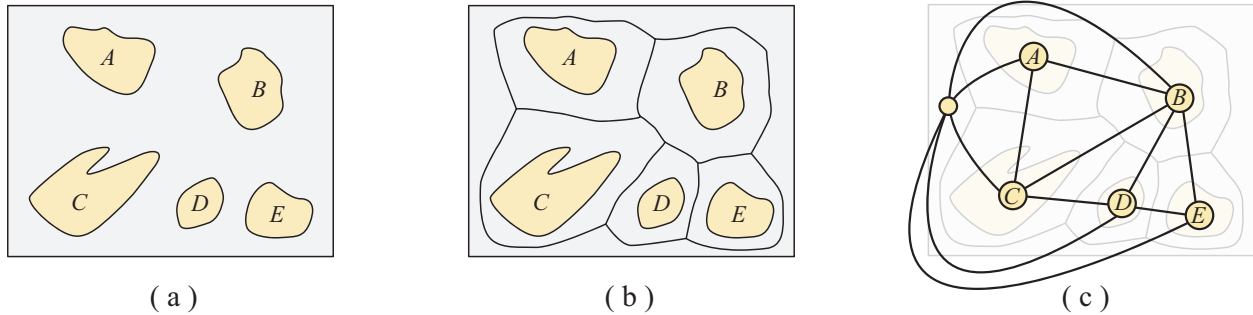


FIGURE 2. Voronoi regions and relative positions.

Definition. The function $f : \mathcal{R} \rightarrow \Omega$ preserves relative position if the canonical function $G_\phi \rightarrow G_f$ taking the vertex $\phi(R_i)$ to the vertex $f(R_i)$ has the property that $\{\phi(R_i), \phi(R_j)\}$ is an edge of G_ϕ if and only if $\{f(R_i), f(R_j)\}$ is an edge of G_f .

Consequently, if two regions are close to each other under the inclusion function ϕ , then they must be close to each other under any scale change defined by some function f .

Definition. A continuous function $f : \mathcal{R} \rightarrow \Omega$ which preserves topology, area-ratios and relative position is called a *scaling function* of ϕ .

3. A HOMOTOPIC APPROACH

Under a scaling function f , it is clear that the image of the regions \mathcal{R} can cover no more than the entire base region Ω . Thus, there is a natural limit to how much scaling of the regions is possible within a given base Ω ; these limiting scales are called end-states:

Definition. A scaling function $f : \mathcal{R} \rightarrow \Omega$ is an *end-state* if $f(\mathcal{R}) = \Omega$.

Remark. The purpose of introducing end-states is for a complete description of the problem. Indeed, it is not intended to represent importance in mapping in its own right.

By abuse of terminology, we sometimes refer to the partition $\{f(R_i)\}$ of Ω as the end-state. FIGURE 3(a) shows the original regions, and a possible end-state is given in (c). So far we have spoken of individual

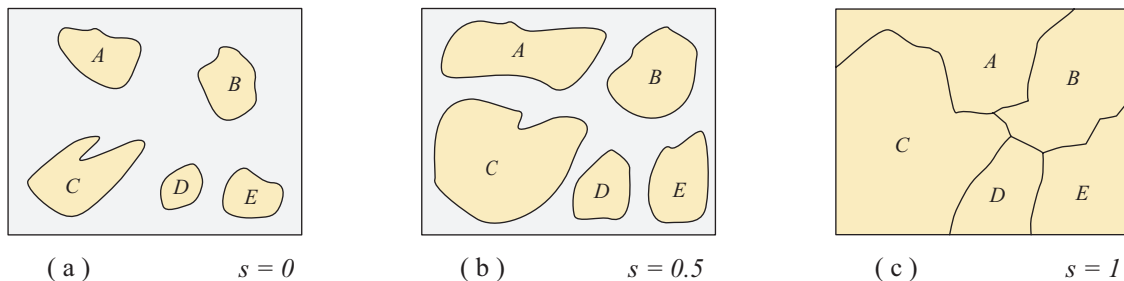


FIGURE 3. An intermediate deformation step between initial and end-states.

functions f which are meant to depict scale changes. It is desirable to be able to enlarge the regions by any amount provided the regions still fit within Ω . That is, we require a *homotopy* connecting the original state (given by the inclusion function) with an end-state (with the regions filling the base region).

Definition. Let X and Y be topological spaces. A *homotopy* between two continuous functions $g_1 : X \rightarrow Y$ and $g_2 : X \rightarrow Y$ is a continuous function $r : X \times [0, 1] \rightarrow Y$ from the product of the space X with the unit interval $[0, 1]$ to Y , where $r(x, 0) = g_1(x)$ and $r(x, 1) = g_2(x)$ for all points x in X .

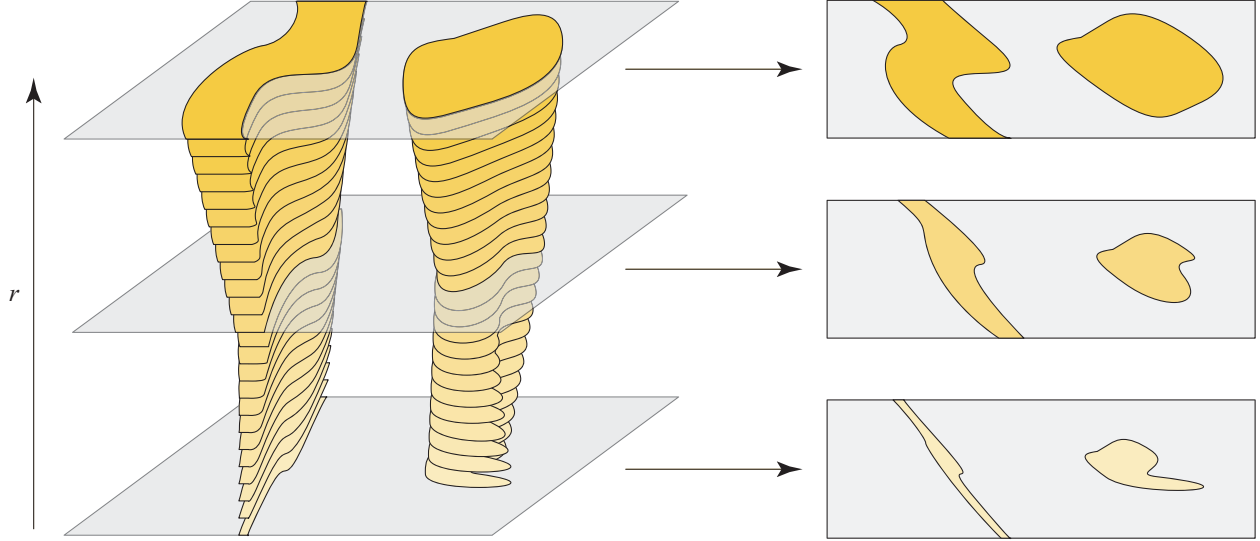


FIGURE 4. Homotopic deformation of two regions over time.

We can view r as a *continuous* deformation of the function g_1 to the function g_2 . Alternatively, one may regard the image $g_1(X)$ as being continuously transformed into the image $g_2(X)$ over some parameter varying in $[0, 1]$. FIGURE 4 shows such a deformation, where three particular generalizations of the map are shown in detail to the right. In our setting, we set $X = \mathcal{R}$ and $Y = \Omega$, where g_1 is the inclusion function ϕ and g_2 is an end-state. By viewing the second parameter of r as a scaling value s ranging within the interval $[0, 1]$, r describes a continuous deformation of ϕ into an end-state. Consequently, as the scaling s increases from 0 to 1, the images of \mathcal{R} grow within Ω , using up the surrounding map space, until all of Ω is filled.

Remark. It is vital to note that the scale values of the *actual map* do not vary from 0 to 1; the value s is but a parameter which keeps track of the actual scales. So if one desires to calculate the evolution of maps from some scale value a to scale value b , we reparametrize these values to range from 0 to 1. Thus, the unit interval $[0, 1]$ is simply used for the sake of consistency.

We are now in position to pose the fundamental problem from a mathematical context.

Problem. Let \mathcal{R} be a collection of pairwise disjoint regions inside a map Ω given by an inclusion function $\phi : \mathcal{R} \rightarrow \Omega$. Construct a homotopy $r : \mathcal{R} \times [0, 1] \rightarrow \Omega$ having the following properties:

- (1) $r(\cdot, 0)$ is the inclusion function ϕ .
- (2) $r(\cdot, s)$ is a scaling function for all values s in the interval $(0, 1)$.
- (3) $r(\cdot, 1)$ is an end-state.

FIGURE 3 shows a possible sequence of homotopic deformation.

4. AN EQUATION OF MOTION

4.1. Our main tool in understanding the adjacency of regions so far has been the Voronoi diagram. Now the *medial axis* is introduced, a finer tool which is used to extract more topological information; see [4] for more details.

Definition. The *medial axis* \mathcal{M} of the regions \mathcal{R} in Ω is the closure of the set of points in Ω that have more than one closest point in \mathcal{S}_∂ . The *outer medial axis* \mathcal{M}_o is $\mathcal{M} \cap \mathcal{R}^c$. The *inner medial axis* \mathcal{M}_i of the region R_i is $\mathcal{M} \cap R_i$.

FIGURE 5(a) shows two regions in Ω with part (b) showing the Voronoi diagram of the regions in dashed lines. FIGURE 5(c) displays the outer medial axis \mathcal{M}_o and (d) shows the inner medial axes \mathcal{M}_i in the interior of each region. Notice that the Voronoi diagram is a subset of the outer medial axis. Indeed, whereas the Voronoi diagram only measures the shape between two *distinct* regions, the medial axis also measures shape deformations between a region and itself.

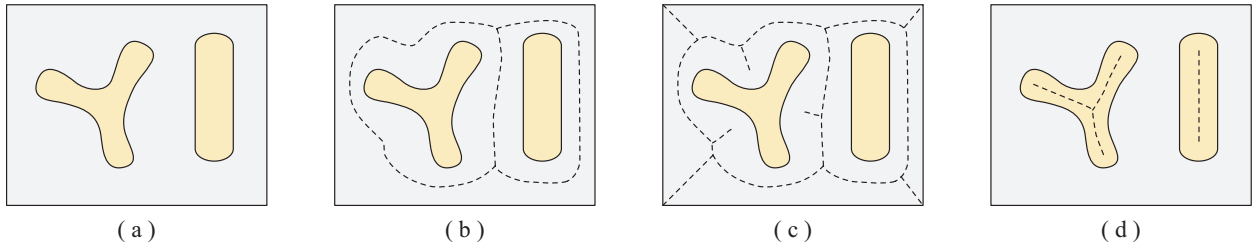


FIGURE 5. (a) Regions along with the (b) Voronoi diagram, (c) outer medial axis, and (d) inner medial axes.

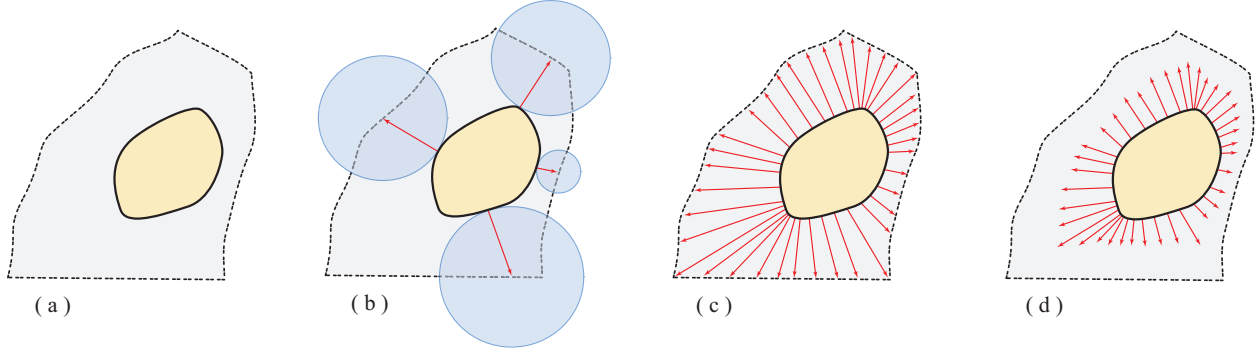
A geometric approach towards a solution of the central problem is proposed below. In using a geometric rather than combinatorial approach, we require that all the curves \mathcal{S}_∂ be differentiable. Roughly, this procedure assigns a velocity vector to the points on the boundary of each region at every instant. The velocity is the sum of three terms: one which points outward from S_i (contributing a net increase in the area of R_i), one which points inward (contributing a net decrease), and a compensation factor to maintain relative position with adjacent regions. We now show how to obtain these velocity vectors at each scale level.

Growth Outward: FIGURE 6(a) displays a region R_i and part of the outer medial axis \mathcal{M}_o (the dashed line). For each point x in S_i , let $\hat{\mathbf{n}}$ be its outward directed unit normal. Let $u(x)$ be the first point of intersection between the ray emanating from x in the $\hat{\mathbf{n}}$ direction and \mathcal{M}_o . Note that $u(x)$ is the center of the unique medial ball whose boundary contains x , as depicted in FIGURE 6(b). Assign a vector $\mathbf{v}_{\text{grow}}(x)$ of length $\|u(x) - x\|$ to each point x on S_i , as shown in part (c).¹ Finally, uniformly rescale each of these vectors over the entire curve S_i so that the total growth value adds to 1; that is,

$$\int_{x \in S_i} \mathbf{v}_{\text{grow}}(x) \cdot \hat{\mathbf{n}} \, dc = 1.$$

Part (d) shows the vectors after the rescaling.

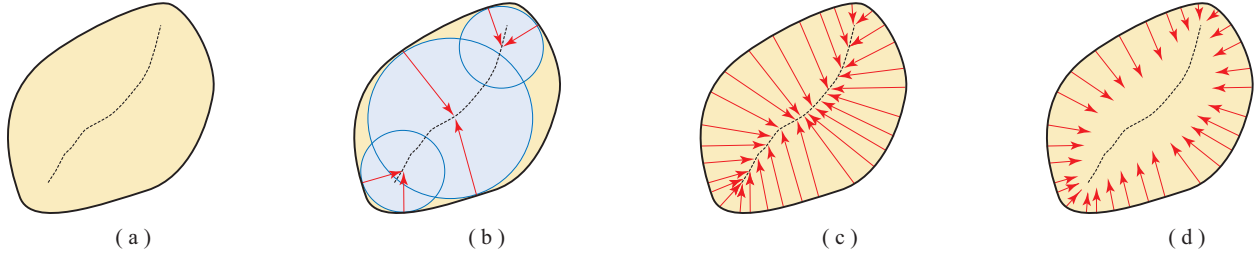
¹This function might be non-differentiable at a few places, and thus needs to be replaced by an arbitrarily close differentiable one. Geometrically, this can be viewed as smoothing of the corners of the medial axis.

FIGURE 6. Assignment of the $\mathbf{v}_{\text{grow}}(x)$ velocity vectors.

Shrink Inward: Consider FIGURE 7(a) showing the region R_i along with its inner medial axis \mathcal{M}_i . For each point x in S_i , let $u(x)$ be the unique point on inner medial axis \mathcal{M}_i such that a medial ball centered at $u(x)$ touches x ; see FIGURE 7(b). Since S_i is differentiable, $u(x)$ is unique. Assign a vector $\mathbf{v}_{\text{shrink}}(x)$ to each x in S_i of length $\|u(x) - x\|$, as exhibited in part (c), and renormalize these values over the curve S_i so that

$$\int_{x \in S_i} \mathbf{v}_{\text{shrink}}(x) \cdot \hat{\mathbf{n}} \, dc = -1.$$

Part(d) shows the vectors after rescaling.

FIGURE 7. Assignment of the $\mathbf{v}_{\text{shrink}}(x)$ velocity vectors.

4.2. We desire the adjacency graph to remain unchanged throughout the homotopy deformation. A natural way to preserve adjacency is to ensure that certain edges of the outer medial axis never contract to a point. The Voronoi edges correspond to the set of points that have more than one closest point to *two distinct* regions. It is clear that these edges must be preserved, but they alone are not enough.

Definition. The *homology graph* \mathcal{M}_H is obtained by iteratively deleting vertices of degree one of the outer medial axis. The edges of \mathcal{M}_H are called *homology edges*.

The top diagrams of FIGURE 8 show \mathcal{M}_o for different regions A , B , and C , whereas diagrams on the bottom display their associated adjacency graph. The Voronoi edges are marked as single dashed lines and the remaining edges of \mathcal{M}_o are marked as thickened dashed lines. These thickened edges correspond to a set of points that have more than one closest point to the *same* region, and thus do not contribute to the adjacency graph. The thickened edges in FIGURE 8(a) and (b) are homology edges, and are distinct from the non-homology edges marked in (c). Although the homology edges do not directly contribute to the

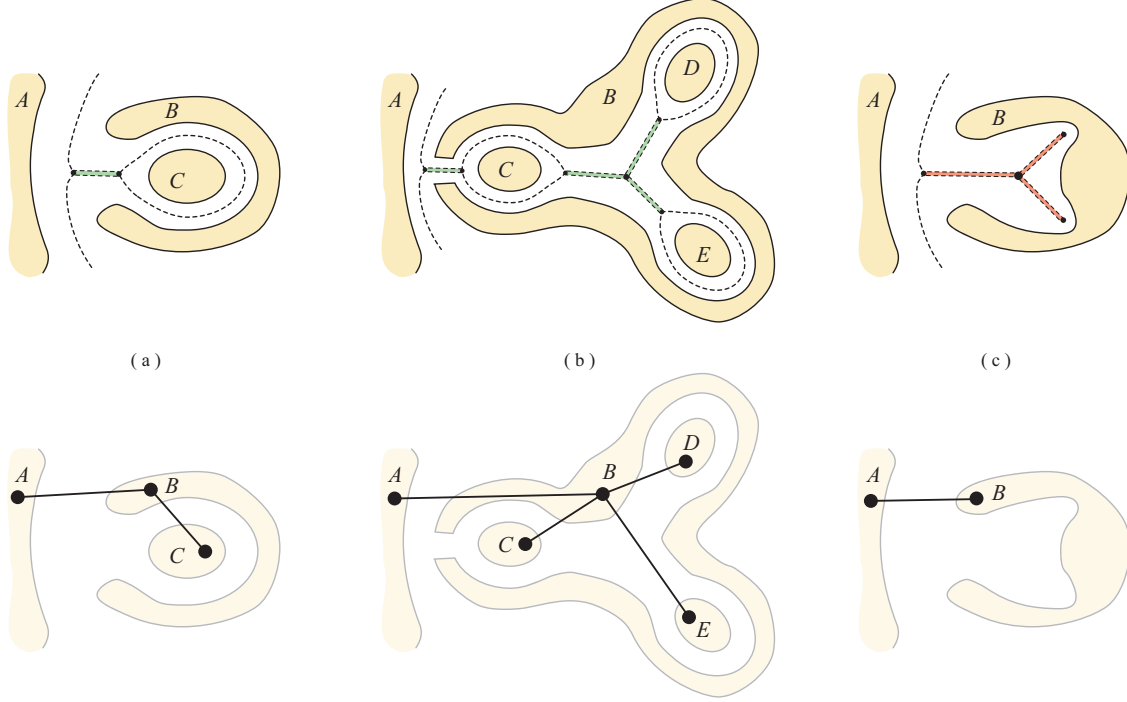


FIGURE 8. The marked edges of \mathcal{M}_o in (a) and (b) are homology edges, but not Voronoi edges. The marked edges in (c) are not homology edges.

adjacency graph, they help preserve it. For example, the marked homology edge in FIGURE 8(a) is keeping regions A and C from becoming adjacent. Similarly, marked homology edges of FIGURE 8(b) are separating regions A and C , along with separating regions C , D and E from each other. A non-homology edge, on the other hand, like those marked in FIGURE 8(c), are not needed for preservation of relative position; indeed, we obtain the following:

Theorem. *If the homology graph is preserved then so is the adjacency graph.*

Proof. Each leaf of the graph \mathcal{M}_o corresponds to a set of points that have more than one closest point to the *same* region. Since a vertex of a leaf has degree one and is free, contracting the edge along the free vertex will not alter the adjacency of regions. Doing this iteratively for all leaves results in the \mathcal{M}_H . \square

Adjacency Preservation: We assign an additional velocity term, based on the length of a homology edge and its distance to the nearest point on \mathcal{S} . Choose a decreasing continuous function $h : \mathbb{R}^+ \rightarrow \mathbb{R}$ where $h(0+) = +\infty$. For a point y on an edge e of \mathcal{M}_o , let ℓ_e be the length of e . Define $H : \mathcal{M}_o \rightarrow \mathbb{R}$, where

$$(4.1) \quad H(y) = \begin{cases} h(\ell_e) & e \text{ is a homology edge} \\ 0 & \text{otherwise.} \end{cases}$$

For a point x in S_i , let $u(x)$ be the first intersection of the outer medial axis \mathcal{M}_o with the ray pointing normally outward from x . Now, define the constant Δ_i to be

$$\Delta_i = \int_{x \in S_i} H(u(x)) \mathbf{v}_{\text{grow}}(x) \cdot \hat{\mathbf{n}} \, dc$$

and define

$$\mathbf{v}_{\text{protect}}(x) = H(u(x)) \mathbf{v}_{\text{grow}} + \Delta_i \mathbf{v}_{\text{shrink}}.$$

This term, as it is meant to preserve adjacency, is constructed in order to not affect the rate of change of area. In particular, note that

$$\int_{x \in S_i} \mathbf{v}_{\text{protect}}(x) \cdot \hat{\mathbf{n}} \, dc = 0.$$

Remark. So far we have defined velocity vectors at a given scale level. It is clear that these vectors depend implicitly on scale level s since the geometry of the regions depend on it. Thus, although we have denoted velocity vectors as $\mathbf{v}_*(x)$, a more accurate notation would be $\mathbf{v}_*(x, s)$. However, we will continue to suppress the dependence on s for clarity of notation.

4.3. We are now in place to state the entire deformation method. Given parameters α_i, β_i and γ_i , the equation of motion for points x in S_i is given by the partial differential equation

$$(4.2) \quad \frac{\partial}{\partial s} r(x, s) = \alpha_i \mathbf{v}_{\text{grow}}(x) + \beta_i \mathbf{v}_{\text{shrink}}(x) + \gamma_i \mathbf{v}_{\text{protect}}(x)$$

satisfying the given initial conditions $r(x, 0) = \phi(x)$. FIGURE 9 shows the assignment of velocity vectors to the boundary of a region, along with its deformation over scale change. Although this equation of motion

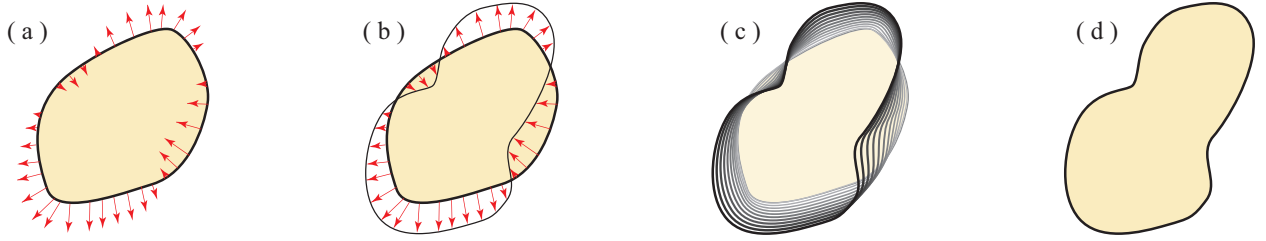


FIGURE 9. Assignment of velocity vectors and its subsequent deformation.

is defined only on the boundaries \mathcal{S} of each region, one can extend it to interiors of the regions, as follows: Each point x in R_i lies on a unique segment which connects a point $u(x)$ on \mathcal{M}_i to a point $p(x)$ on S_i . This segment is a radius of the medial ball centered at $u(x)$, going through x , touching $p(x)$; see the line segments in FIGURE 7(c). We apply a straight-line deformation to this line segment. Thus, for each x in R_i , let

$$(4.3) \quad \frac{\partial}{\partial s} r(x, s) = \frac{\text{dist}(x, u(x))}{\text{dist}(p(x), u(x))} \frac{\partial}{\partial s} r(p(x), s),$$

where $\frac{\partial}{\partial s} r(p(x), s)$ is defined in Eq. (4.2). We note some features of this method:

- (1) The parameter α_i is called the *growth-rate* of region R_i as the rate of growth due to its corresponding term in the equation is exactly α_i . The desired growth-rate will be specified by the user, yielding an extra degree of freedom.
- (2) The parameter β_i is called the *shrink-rate* as the rate of growth is exactly $-\beta_i$. By adjusting β_i the user may change the quality of the deformation. A high value for β_i will allow a region to be very flexible and to quickly recede from other regions. On the other hand, if $\beta_i = 0$, the boundaries \mathcal{S} may never recede, forcing regions to deform around the original images $\phi(\mathcal{R})$; each region will recede only due to $\mathbf{v}_{\text{protect}}$.

- (3) The parameter γ_i is also defined by the user and can be adjusted to control the quality of deformations. If γ_i is large, deformations may appear unnatural as part of the region may expand very quickly to protect an edge. On the other hand, if γ_i is small, edge protection will not happen until an edge has nearly contracted. Indeed, such sudden changes will make numerical integration of these equations very difficult. The function h , also defined by the user, appearing in $\mathbf{v}_{\text{protect}}$, further determines the quality of deformation due to the edge protection term.

The following result is an immediate consequence of our construction:

Theorem. *Let k be any constant. For each region R_i , choose α_i and β_i such that $\alpha_i - \beta_i = k \mu(R_i)$. Then the solution to Eq. (4.2) will preserve the area-ratios of the regions throughout the deformation.*

Proof. Let α_i, β_i and γ_i be parameters for region R_i , and let $r(x, s)$ be a solution to Eq. (4.2). Then the rate of change of the area of R_i is given by

$$\begin{aligned} \frac{d}{ds} \mu(r(R_i, s)) &= \int_{x \in S_i} \frac{\partial}{\partial s} r(x, s) \cdot \hat{\mathbf{n}} \, dc \\ &= \alpha_i \int_{x \in S_i} \mathbf{v}_{\text{grow}}(x) \cdot \hat{\mathbf{n}} \, dc + \beta_i \int_{x \in S_i} \mathbf{v}_{\text{shrink}}(x) \cdot \hat{\mathbf{n}} \, dc + \gamma_i \int_{x \in S_i} \mathbf{v}_{\text{protect}}(x) \cdot \hat{\mathbf{n}} \, dc \\ &= \alpha_i - \beta_i \\ &= k \mu(R_i). \end{aligned}$$

Solving this first-order equation, we get

$$\mu(r(R_i, s)) = ks \mu(R_i) + \mu(R_i) = (ks + 1) \mu(R_i).$$

It follows that area-ratios will be preserved. □

5. END-STATES AND CARTOGRAMS

5.1. Although our goal is to find a homotopy which encapsulates continuous change of scales, the question of finding an end-state is itself interesting. It is easy to see that the end-state is not necessarily unique. FIGURE 10(a) shows an example of an initial configuration of regions, along with three (b – d) possible end-states, all of which capture different information. This illuminates the fundamental challenge to calculating

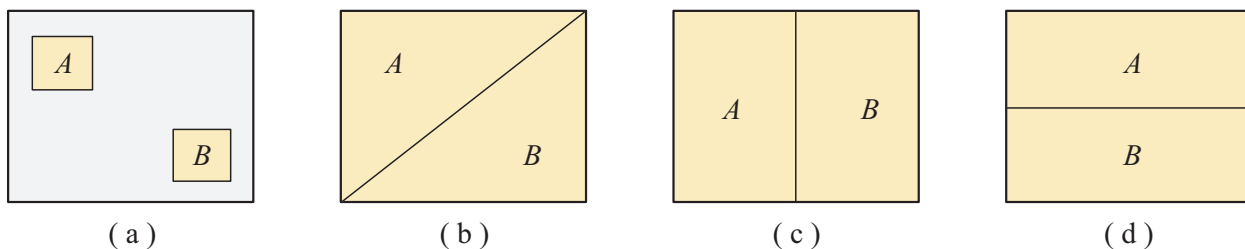


FIGURE 10. Given original regions (a), along with three possible end-states.

a final state for the regions, namely, that the different pieces of information we want to preserve on the map can often work against each other. Consequently, the notion of a *canonical* solution to this problem must necessarily include a prioritization of the desired map information to be preserved. Although an end-state solution is not unique, it always exists.

Theorem. *Given a collection of pairwise disjoint regions in Ω , there exists an end-state.*

Proof. The collection of Voronoi regions²

$$V_i = \{x \in \Omega \mid \text{dist}(x, S_i) \leq \text{dist}(x, \mathcal{S} \setminus S_i)\}$$

for a set of initial regions $\{R_i\}$ forms a partition of Ω . The Voronoi regions are not a viable end-state since they do not necessarily have the desired area-ratios. However, if one deforms the Voronoi regions in such a way that the adjacency graph is maintained and each resulting region has the desired area, then an end-state can be obtained.

Let a_i be the area of V_i and let b_i be the area for R_i at its end-state. If $a_i = b_i$ for all i , we are done. Else, at least one V_p has too much area and one V_q has too little. Let $\kappa = \min\{a_p - b_p, b_q - a_q\}$. Since Ω is connected, the adjacency graph of the regions is connected; thus, a path γ in the adjacency graph exists connecting V_p to V_q . We claim that κ -area can be transferred along the chain from V_p to V_q .

Let V_{p+1} be the region adjacent to V_p along γ ; thus, the union of V_p and V_{p+1} can be considered as a new region. It is easy to see that area can be transferred between V_p and V_{p+1} by taking the edge shared by the two regions and deforming it into a new edge so that V_p loses κ -area and V_{p+1} gains κ -area. This process is continued throughout the path γ . Thus, after each step, at least one more region (either V_p or V_q) has the correct area. This process terminates in at most $n - 1$ steps, where n is the number of regions. By design, the final partition has exactly the same adjacency information and is thus an end-state. \square

5.2. A cartogram is a map in which the sizes of regions appear proportional to some other parameter different from the areas of the regions. Notions of maintaining shape, topology, and adjacency of regions are clearly valued in such maps; see Tobler [12] for a nice overview. A combinatorial approach preserving topology is given in Edelsbrunner and Waupotitsch [2], whereas Gastner and Newman [3] use a method based on diffusion equations. Notice that this end-state problem is closely related to cartograms. The end-state $f : \mathcal{R} \rightarrow \Omega$ forces area-ratios of the regions $\{f(R_i)\}$ to be identical to the original regions $\{R_i\}$. For cartograms, however, the areas of $\{f(R_i)\}$ are prescribed by other data or parameters (such as population). Based on our equation of motion, it is possible to construct cartograms in which the regions displayed in the cartogram have exactly the desired areas.

Theorem. *Given a collection of regions $\{A_1, \dots, A_n\}$ which partition Ω , an appropriate cartogram of the regions can be constructed using Eq. (4.2).*

Proof. Suppose $\mu(A_i) = a_i$ and a desired area b_i for each A_i is given, where $\sum b_i = \sum a_i$. Clearly some regions need to increase in area and others need to decrease in order to obtain the required cartogram. Without loss of generality, assume region A_1 needs to decrease the most in area out of all the regions. Thus,

$$\frac{b_1}{a_1} = \min \left\{ \frac{b_1}{a_1}, \frac{b_2}{a_2}, \dots, \frac{b_n}{a_n} \right\}.$$

We use the homotopic approach twice — once to shrink the regions in order to recede away from each other until region A_1 has reached its final area — and once to enlarge the remaining regions filling all of Ω as required. During this process, the growth rates are altered in order to end with the preferred areas.³

²Unlike the Voronoi regions defined in Eq. (2.1), we are not interested in the boundary $\partial\Omega$ now.

³It is necessary to separate the regions with some space. To do this we remove from each set A_i a small δ -neighborhood of ∂A_i . The resulting regions (still denoted as A_i) leaves every pair of regions separated by at least 2δ . As always, we smooth out corners of the regions so that their boundaries are differentiable.

RECEDE: For each region A_i , set the growth-rate to be $\alpha_i = 0$ and the shrink-rate to be $\beta_i = a_i$. Find a solution $r_1(x, s)$ to Eq. (4.2) using the regions $\{A_i\}$ on the scale interval $[0, \lambda]$, where $\lambda = 1 - (b_1/a_1)$. Since $\alpha_i - \beta_i$ is the rate of area change of region A_i (as shown by the last Theorem of Chapter 4), deforming the regions for λ unit of time changes the area of region A_i from $\mu(A_i)$ to

$$\mu(A_i) + \lambda \cdot (\alpha_i - \beta_i) = a_i + \lambda \cdot (0 - a_i) = a_i(1 - \lambda).$$

Denoting the resulting deformed regions $r_1(A_i, \lambda)$ as R_i , we note that $b_i - \mu(R_i) \geq 0$ by our choice of λ . In particular, λ is chosen such that $\mu(R_1) = b_1$.

ENLARGE: Since R_1 has attained its final area, all other regions can now be allowed to grow and fill in the cartogram to obtain the desired result. Choose α_i and β_i such that $\alpha_i - \beta_i = b_i - \mu(R_i)$. Once again, find a solution $r_2(x, s)$ to Eq. (4.2), now using the regions $\{R_i\}$ on the scale interval $[\lambda, 1 + \lambda]$. Since $\alpha_i - \beta_i$ is the rate of area change of region R_i , deforming the regions for one unit of time (from λ to $\lambda + 1$) changes the area of region R_i to

$$\mu(R_i) + 1 \cdot (\alpha_i - \beta_i) = \mu(R_i) + (b_i - \mu(R_i)) = b_i.$$

The solution will have the property that $\sum \mu(r_2(R_i, s)) < \mu(\Omega)$ for s in the interval $[\lambda, 1 + \lambda]$. The resulting cartogram is now given by the regions $r_2(R_i, \lambda + 1)$, preserving adjacencies and resulting in the desired areas. \square

6. CLOSING REMARKS

Although the focus of this paper is on theoretical approaches to continuous generalization, notably from a mathematical viewpoint, we conclude by addressing some practical matters which might arise in implementing the differential method defined above. We begin with the algorithmic issues.

- (1) APPROXIMATION: Since it is not possible to represent an arbitrary continuous function in computer memory, any algorithm for this problem must approximate the curves by some means. One standard method which is feasible in this situation is the use of Bezier splines; a classical reference is de Boor [1]. These can not only approximate the boundary curves of the regions arbitrarily close but preserve the differentiability that is needed. With regards to Voronoi diagrams and medial axes, spline (or even polygonal) approximations are also sufficient. The simulation of the deformation process proceeds by taking small time (scale) slices, moving the control points of the splines and then updating the medial axes and Voronoi diagrams. Computing the motion of each control point requires computing several integrals along the curves. The spline representation makes this fairly straight-forward as we compute the value of the functions at the control points and linearly interpolate the values between.
- (2) COMPLEXITY: In all cases, the complexity of the structures is determined by how finely the curves are sampled with spline control points. Voronoi diagrams and medial axes are both computable in polynomial time; see [6] and [15] for concrete implementations. The total running time of the algorithm is polynomial in the number of control points used to represent the shapes and in the number of time (scale) slices used in the simulation. Thus, even though the complexity of writing and implementing the given algorithm may be prohibitive for many uses, we believe it is not computationally intractable to do so.

Although this system provides an understanding of shape deformation during generalization, there are two strong restrictions that are placed upon the boundary of the regions in Ω .

- (1) DIFFERENTIABLE: The most important condition for our regions is that their boundary curve be differentiable. It is this feature that allows the use of calculating velocities in a continuous manner. On the other hand, most spatial data is stored as polygons having piecewise-linear boundary. However, it is a classical result of analysis that any piecewise-linear curve can be approximated arbitrarily close⁴ to a differentiable curve. Thus, any polygonal data can easily be transformed into a differentiable one for the use of this method.
- (2) ONE COMPONENT: The second restriction is for each region to have only one boundary component. At first, this restriction seems to sound quite severe, where usual features such as lakes and islands found in maps might not be allowed. However, upon close examination, this is not necessarily true. Recall that Eq. (4.3) extends the deformation of the boundary to the *interior* of each region. In addition, Eq. (4.3) places no restriction on the features *within* the regions; indeed, objects can be placed in each region which are highly singular and non-differentiable. The downside of this, however, is that the deformation of each object within a region is *not* controlled by its own features, but by the boundary of the region it resides in.

Overall, this paper has focused on the deformation of the shapes of regions in a map during the process of continuous scale change. The scale change transformations are given in the language of continuous mathematics, notably that of a homotopy. This homotopic solution, given by Eq. (4.2), is also equipped with three parameters, α_i, β_i and γ_i , which can be altered to satisfy the user's needs. Although only a small part of the issues surrounding continuous generalization has been studied, we hope this will lead to motivating further work in incorporating other mathematical ideas. These tools, along with others, could show an alternate approach and framework in which generalization can be cast and attacked. The interplay between the needs that should be met by cartographic scaling along with results which can be formulated and proven is a vital one. Our intentions are not to provide a complete framework for this, but to lay part of the foundation.

Acknowledgments. We are grateful to the NSF for partially supporting this project with grants DMS-0353634 and CARGO DMS-0310354. We thank Tamal Dey, Chris Jones, Marc van Kreveld, Alan Saalfeld, Jim Stasheff and Robert Weibel for helpful conversations. Satyan Devadoss also wishes to thank Jörg-Rüdiger Sack, Monika Sester, Peter van Oosterom, and Michael Worboys for organizing the Dagstuhl workshop on Spatial Data in 2006, which motivated and solidified several concepts in this work. Finally, the authors thank the anonymous reviewers for their insightful comments and corrections.

REFERENCES

1. C. de Boor. *A practical guide to splines*, revised ed., Springer, New York, 2001.
2. H. Edelsbrunner and R. Waupotitsch, A combinatorial approach to cartograms, *Computational Geometry* **7** (1997), 343-360.
3. M. Gastner and M. Newman. Diffusion-based method for producing density-equalizing maps, *Proc. Natl. Acad. Sci.* **101** (2004), 7499-7504.

⁴Two regions are arbitrarily close if the distance between the corresponding points of the regions are within some small $\varepsilon > 0$ of each other.

4. J. Goodman and J. O'Rourke (eds). *Handbook of Discrete and Computational Geometry*, 2nd ed., CRC Press, New York, 2004.
5. A. Hatcher. *Algebraic Topology*, Cambridge University Press, New York, 2002.
6. K. Hoff, T. Culver, J. Keyser, M. Lin, D. Manocha. Fast Computation of Generalized Voronoi Diagrams Using Graphics Hardware, in *International Conference on Computer Graphics and Interactive Techniques* (1999), 277 - 286.
7. C. Jones and J. Ware. Map generalization in the web age, *International Journal of GIS* **19** (2005), 859-870.
8. N. Lam, D. Catts, D. Quattrochi, D. Brown, R. McMaster. Chapter 4: Scale, in *A Research Agenda for Geographic Information Science*, R. McMaster and E. Lynn Usery (eds), CRC Press (2004), 93-128.
9. A. Okabe, B. Boots, K. Sugihara. *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*, John Wiley and Sons, New York, 1992.
10. A. Saalfeld. Consistent map generalization via white space management, preprint 1996.
11. M. Sester and C. Brenner. Continuous generalization for small mobile displays, in *Next Generation Geospatial Information*, Taylor and Francis Group, London (2006), 33-41.
12. W. Tobler. Thirty five years of computer cartograms, *Annals, Assoc. Am. Geographers* **94** (2004), 58-73.
13. M. van Kreveld. Smooth generalization for continuous zooming, *Proc. 20th Intl. Geographic Conference* (2001), 2180-2185.
14. P. van Oosterom. Variable-scale topological data structures suitable for progressive data transfer: the GAP-face tree and GAP-edge forest, *Cartography and Geographic Information Science* **32** (2005), 331-346.
15. M. Ramanathan and B. Gurumoorthy. A tracing algorithm for constructing medial axis transform of 3D objects bound by free-form surfaces, *Computer-Aided Design* **35** (2003) 619-632.
16. W. Yang and C. Gold. A system approach to automated map generalization, *Proceedings: International Workshop on Dynamic and Multi-Dimensional GIS* (1997), 229-235.

DANCIGER: STANFORD UNIV., STANFORD, CA 94305

E-mail address: danciger@math.stanford.edu

DEVADOSS: WILLIAMS COLLEGE, WILLIAMSTOWN, MA 01267 (CORRESPONDING AUTHOR)

E-mail address: satyan.devadoss@williams.edu

MUGNO: UNIV. OF MARYLAND, COLLEGE PARK, MD 20742

E-mail address: jmugno@math.umd.edu

SHEEHY: CARNEGIE MELLON UNIV., PITTSBURGH, PA 15213

E-mail address: dsheehy+@cs.cmu.edu

WARD: PRINCETON UNIV., PRINCETON, NJ 08544

E-mail address: rward@math.princeton.edu