Exploring curved schematization of territorial outlines

Citation for published version (APA):

DOI:
10.1109/TVCG.2015.2401025

Document status and date:
Published: 01/01/2015

Document Version:
Accepted manuscript including changes made at the peer-review stage

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.

Download date: 03. Aug. 2019
Exploring Curved Schematization

Arthur van Goethem, Wouter Meulemans, Bettina Speckmann, and Jo Wood

Abstract—Hand-drawn schematized maps traditionally make extensive use of curves. However, there are few automated approaches for curved schematization; most previous work focusses on straight lines. We present a new algorithm for area-preserving curved schematization of geographic outlines. Our algorithm converts a simple polygon into a schematic crossing-free representation using circular arcs. We use two basic operations to iteratively replace consecutive arcs until the desired complexity is reached. Our results are not restricted to arcs ending at input vertices. The method can be steered towards different degrees of “curviness”: we can encourage or discourage the use of arcs with a large central angle via a single parameter. Our method creates visually pleasing results even for very low output complexities. To evaluate the effectiveness of our design choices we perform a geometric evaluation of the resulting schematizations. Besides the geometric qualities of our algorithm, we also investigate the potential of curved schematization as a concept. We conducted an online user study investigating the effectiveness of curved schematizations compared to straight-line schematizations. While the visual complexity of curved shapes was judged higher than that of straight-line shapes, users generally preferred curved schematizations. We observed that curves significantly improved the ability of users to match schematized shapes of moderate complexity to their unschematized equivalents.

Index Terms—Schematization, circular arcs, algorithm, user study

1 INTRODUCTION

Maps are one of the most efficient ways to communicate location-based information. They help people to make decisions in navigation, spatial planning, or risk and disaster management. Maps also communicate geopolitical information, they give a spatial dimension to rhetoric arguments, and generally aid the process of public opinion and consensus building. Effective maps immediately convey their message and hence are as simple as possible. Schematization creates a simplified representation of the original data by applying design-based rules to reduce visual complexity beyond the needs of the target scale. Linear features, such as roads and rivers, and the boundaries of regions are drawn using only a few straight line segments in a few different directions, or they are approximated by a few simple curves. Schematization attempts to direct the user’s attention away from the exact shapes of geographic entities, focussing on the relation between those entities instead. Alternatively, its purpose can be to produce a striking and memorable design for maximum impact. Furthermore, schematized maps are often used as base maps of thematic maps, to avoid the “illusion of accuracy” created by fully schematized maps geographic outlines are typically captured by a small set of simple mathematical shapes. Hand-drawn schematized maps traditionally make extensive use of curves. Curves can capture more complex shapes than straight lines, which are represented by a small set of simple mathematical shapes. We present an algorithm for schematization with circular arcs. For an input consisting of \(n\) edges, our algorithm runs in \(O(n^2)\) time. Our algorithm allows new vertices to be introduced and as such is a non-vertex-restricted method. Furthermore, it preserves the exact size of each region of the input and maintains topology. The algorithm iteratively replaces consecutive arcs until the desired complexity is reached. This replacement is based on two operations. The first is vertex-restricted and replaces two consecutive arcs by a single arc connecting the endpoints. The second operation is non-vertex-restricted and replaces three consecutive arcs at a new location. We prioritize operations according to different schemes to obtain results of varying degree of “curviness”—preferring arcs with a large or small central angle. Our method creates visually pleasing results even for very low output complexities and combines well with different rendering styles. Lastly, our algorithm can also deal with subdivisions (e.g., multiple countries). We investigate the effectiveness of our design choices by measuring the geometric accuracy of our algorithm. Introducing new vertices indeed improves the quality of the schematization.

We conducted a user study investigating the effectiveness of curved schematizations compared to straight-line
schematizations of equivalent complexity. While the visual complexity of the curved shapes was judged higher than that of shapes consisting of straight lines, users generally preferred curved schematizations. Curves also significantly improved the ability of users to match schematized shapes of moderate complexity to their unschematized equivalents.

**Organization.** We review related work and, in Section 2, introduce terminology. In Section 3 we present our algorithm and discuss various extensions. In Section 4 we evaluate its performance in terms of geometric properties more thoroughly. We discuss the user study in Section 5 and close in Section 6 with a discussion of our work.

**Related work.** Simplification of shapes has received significant attention (e.g., [2], [20], [29]). In contrast to simplification, schematization is less focused on geographic accuracy and more on visual presentation. Network schematization (e.g., metro maps) has been studied extensively (e.g., [9], [16], [19]). These algorithms are often not concerned with maintaining shape and most use only straight-line representations. An exception is the force-directed method of Fink *et al.* [9], which uses Bézier curves.

There are fewer methods for shape schematization. Buchin *et al.* [4] and Cicerone and Cermignani [6] describe algorithms for schematization where every line must adhere to one of a given set of orientations. Reimer and Meulemans [22] conjecture that parallelism drives straight-line schematization. Automated curved schematization has only recently emerged as a research topic. Van Goethem *et al.* [11] describe a framework for topologically safe curved schematization. This framework supports only vertex-restricted methods and is quite slow. For an input consisting of \( n \) edges and a desired output schematization using \( k \) curves, the algorithm runs in \( O(n^2 \cdot k) \) time. To ensure topology preservation, the edge-based Voronoi diagram of the input shape is used. The resulting Voronoi cells, however, can be overly restrictive when different parts of the shape are close together yet far apart along the polygon boundary. Consequently, a low-complexity schematization may not be possible. Mi *et al.* [17] describe an algorithm for curved abstraction. They detect basic parts and rebuild a shape up to a given detail using these parts. Their algorithm, however, does not consider topology. The use of curves is almost a given in manual cartography; examples include chromatic diagrams [21] and transit maps [23].

Smooth circular arcs have been used to represent inherently smooth polygons [13]. It is unclear whether such solutions are useful for schematization: in this context it is generally not desirable to assume that the original shape is relatively smooth. A non-smooth approach is given by Drysdale *et al.* [7]. This method imposes restrictions on “gates” which hinder a high complexity reduction. In the field of graph drawing, the use of curves has also received significant attention, e.g., [8], [10]. However, these papers typically do not preserve any measure of shape. Kämper *et al.* [14] present a method for distorting a schematization into a circular-arc cartogram, but recognizability is not a primary concern; aesthetics and legibility remain untested.

For computer-aided design, Burchard *et al.* [26] discuss fitting a curve given aesthetic requirements.

A wealth of research exists on perception, rendering, and their combination. For example, Bar and Neta [1] argue that curved objects are preferred by observers since sharp bends are identified with threat. This observation would support the need for curved schematizations, though to maintain recognizability some sharp bends may still be necessary. Vessel and Rubin [28] investigate the objectiveness of taste. They conclude that for natural, real-world images, people typically agree on aesthetics; however, on abstract images, individual taste plays a large role. While schematizations are typically rather abstract depictions, they stem from a reality and should up to some degree also correspond to this reality. Hence, we think that there is some consensus possible on what qualifies as a good schematization. Vande Moere *et al.* [27] evaluate the effect of visual style in the context of information visualization.

## 2 Preliminaries

**Non-vertex-restricted.** Simplification and schematization algorithms can be classified as vertex-restricted or non-vertex-restricted. For the former, each output vertex must also be an input vertex. For the latter, output vertices may be placed freely. As illustrated in Figure 1, freely placing new vertices benefits the ability to accurately capture shapes. Our algorithm is a non-vertex-restricted method that computes the result shown in Figure 1(c).

**Algorithmic properties.** A result is topologically correct if it has no intersections and each region maintains the same adjacencies. This property is crucial to shape schematization: incorrect topology greatly interferes with legibility and recognizability of a map.

Relative sizes, and in some cases absolute sizes, may also influence the information portrayed. Hence, when schematizing shapes, we often prefer not to greatly distort the region sizes. We enforce area preservation as a strict way of maintaining relative sizes: each region in the input has the exact same area in the output. Our result is area equivalent to the input; and our algorithm is area preserving.

**Circular arcs.** A (circular) arc is a connected part of a circle. An arc is given by its center \( c \), startpoint \( s \) and endpoint \( e \), and its orientation (clockwise or counterclockwise). The central angle \( \alpha \) is the angle from segment \( cs \) to \( ce \), either the
clockwise or counterclockwise angle depending on the orientation. The circular segment is the region enclosed by the arc and line segment se (see Figure 2).

**Signed area.** We use signed area to reason about area preservation (see Figure 2). Assume we are given an open curve $C$, starting at $u$ and ending at $v$; in addition, assume there is a curve $C'$ from $v$ to $u$ such that the concatenation of $C$ and $C'$, as well as the concatenation of line segment $uv$ and $C'$ are non-self-intersecting counterclockwise closed curves. Let $A(C)$ and $A(uv)$ respectively denote the area enclosed by these closed curves (respectively, green & gray and blue & gray in the figure). The signed area of $C$ is then $A(C) - A(uv)$, that is, it is the area it adds compared to its straight-line replacement. Any valid $C'$ results in the same signed area.

For an arc, the signed area is the size of its circular segment, computed as $\frac{1}{2} \cdot r^2 \cdot (\alpha - \sin \alpha)$ for radius $r$ and central angle $\alpha$. The signed area is positive for counterclockwise arcs and negative for clockwise arcs. Two points and a signed area uniquely determine an arc (Property 1).

**Property 1.** Given two points $u$ and $v$ and a value $A$, there is a unique arc from $u$ to $v$ with signed area $A$.

For a sequence of (at most 3) arcs, we sum the signed area of each arc and the signed area of the polyline formed by the vertices of the arcs. To this end, consider a polyline $P = \langle u_1, \ldots, u_k \rangle$ with $k \leq 4$. The signed area of the polyline is computed as $\frac{1}{2} \sum_{i=1}^{k-1} (u_i \times u_{i+1})$ where $\times$ denotes the 2-dimensional cross product.

**Distance measures.** A number of distance measures exist to quantify the similarity between two polygons. In this paper we use the symmetric difference and the Fréchet distance. For two polygons $P$ and $Q$, the symmetric difference is defined as the area that is covered by their union but not by their intersection: $|P \cup Q| - |P \cap Q|$. The Fréchet distance interprets the boundary of a polygon $P$ as a continuous function $P : S^1 \rightarrow \mathbb{R}^2$, where $S^1$ denoted the unit circle. Let $\Psi$ denote the set of all orientation-preserving homeomorphisms on $S^1$. The Fréchet distance is then defined as $\inf_{\psi \in \Psi} \max_{t \in S^1} \|P(t) - Q(\psi(t))\|$. This is the pointwise Fréchet difference.

**3 SCHEMATIZATION ALGORITHM**

We describe an algorithm that schematizes a simple polygon using circular arcs. The algorithm maintains a closed curve $S = \langle a_1, \ldots, a_n \rangle$ consisting of circular arcs $a_i$. Line segments are considered to be arcs with an infinite radius. The input—a simple polygon—is therefore also a closed curve. The complexity of $S$ is its number of arcs $|S| = n$. Every consecutive pair of arcs, say $a_i-1$ and $a_i$, meet at a vertex $v_i$. Arc $a_i$ is oriented from $v_i$ to $v_{i+1}$. We treat the sequence circularly, e.g. $a_{n+1} = a_1$ and $v_{n+1} = v_1$.

Our algorithm uses two types of operations to decrease the complexity of a shape (Section 3.1). An operation replaces a sequence of arcs by a shorter sequence. As such, it makes only local changes. By iteratively applying these operations we obtain a schematization (Section 3.2). We describe two optional postprocessing steps in Section 3.3.

**3.1 Area-preserving operations**

Here we describe the operations used by our algorithm. Each operation is executed on two or three consecutive arcs. An operation replaces a chain of arcs by a new chain using one arc less. The vertices at which the sequence starts and ends must remain at their original position.

We wish to ensure that the area encompassed by curve $S$ remains the same. To this end, we enforce that each operation is area preserving: the signed area of a replaced arc sequence must be equal to the signed area of the replacement. By extension, this means that any sequence of operations preserves the area as well.

The operations change the curvature of arcs. This is desirable as it implies that they can be used to turn straight-line shapes into curved shapes.

**Vertex-restricted.** The vertex-restricted operation replaces two consecutive arcs by a single arc. The signed area of the initial sequence gives rise to a required signed area for the result (see Figure 2). This uniquely determines the replacement arc (Property 1). How to compute this arc is detailed in the appendix.

**Non-vertex-restricted.** The non-vertex-restricted operation takes as input three consecutive arcs, say $a_{i-1}$, $a_i$, and $a_{i+1}$. It replaces these by two arcs $a'$ and $a''$ such that $a'$ starts at $v_{i-1}$ and $a''$ ends at $v_{i+2}$. Arcs $a'$ and $a''$ meet at a (possibly new) vertex $v'$. We have three degrees of freedom when inserting the new arcs. The position of $v'$ can be chosen freely (2 degrees) and the same holds for the signed area of one of the two arcs. As the entire operation must preserve area, this directly implies a signed area for the second arc. Ideally, we would find the optimal replacement for the three arcs. However, this seems infeasible and we instead apply the following heuristic to find a good replacement.

We define a line $L$ on which vertex $v'$ needs to be placed. Line $L$ runs through the center of $a_i$ and a point $p$ on $a_i$. (If $a_i$ is a line segment, $L$ is the perpendicular at $p$.) We would like the point $p$ to be closer to $a_{i+1}$ if $a_{i-1}$ and $a_i$ are alike. Similarly, it should be closer to $a_{i-1}$ if $a_i$ and $a_{i+1}$ are alike. We measure the symmetric difference $\Delta_1$ caused by replacing $a_{i-1}$ and $a_i$ by their area-equivalent arc. Similarly, we obtain the symmetric difference $\Delta_2$ caused by replacing $a_i$ and $a_{i+1}$ by their area-equivalent arc. We use for $p$ the point that is a fraction of $\frac{\Delta_1}{\Delta_1 + \Delta_2}$ along the perimeter of $a_i$ as measured from $v_i$ (see Figure 3). In particular, if $\Delta_1$ is zero, then $p$ is $v_{i+1}$; if $\Delta_2$ is zero, then $p$ is $v_i$. If both $\Delta_1$ and $\Delta_2$ are zero, we use a fraction of $\frac{1}{2}$.
Point $p$ also defines the signed area of $a'$ and $a''$ (see Figure 4). The signed area for $a'$ is the signed area between $v_{i-1}$ and $v_i$ caused by $a_{i-1}$, $a_i$, and $a_{i+1}$, respectively. Similarly, the signed area for arc $a''$ is defined by segment $pv'$, $a_i$ from $p$, and $a_{i+1}$. This uniquely defines a solution for any point $v'$ (Property 1).

For $v'$, we use the point on $L$ that minimizes the symmetric difference for the resulting arcs. To this end, we assume that the symmetric difference is unimodal on $L$ (i.e., it has a unique minimum) and perform a golden search—a “binary search” for unimodal functions. The solution obtained may still contain intersections (i.e., it is not planar). Non-planar solutions must be rejected to maintain the correct topology. When $v_{i-1}$ and $v_{i+2}$ are on the same side of solution line $L$, this is the best we can do. It is unclear how to efficiently find a legal solution and in specific cases no legal solution may be found (see Figure 5). However, when $v_{i-1}$ and $v_{i+2}$ are on opposite sides of the solution line, we can always transform the solution into a planar solution.

**Lemma 1.** Assume $v_{i-1}$ and $v_{i+2}$ lie on different sides of line $L$. Then there exists a planar solution with $v'$ on $L$.

**Proof:** As the problem is invariant under rotation and translation, we assume that both $v_{i-1}$ and $v_{i+2}$ lie on the $x$-axis (have $y$-coordinate zero), and we assume that $v_{i-1}$ is left of $v_{i+2}$. Since $v_{i-1}$ and $v_{i+2}$ lie on different sides, solution line $L$ cannot be horizontal.

Moving point $v'$ downward along line $L$ causes the signed area of both arcs to decrease. Symmetrically, moving $v'$ upward causes an increase in signed area. Hence, there is a unique point $l_1$ on $L$ where the first arc has signed area zero (and thus is a straight line). Analogously, there is a unique point $l_2$ on $L$ where the second arc has signed area zero. A solution higher than $l_1$ on $L$ causes a positive signed area for the first arc (and thus a counterclockwise arc), any solution below causes a negative signed area (and thus a clockwise arc). Without loss of generality, we assume that $l_1$ is higher than $l_2$. We consider two cases based on $l_1$ and $l_2$.

Assume that $l_1$ and $l_2$ lie on different sides of the $x$-axis (Figure 6(a)). Consider the solution with $v'$ on the $x$-axis. As $l_1$ lies above $v'$, the first arc is clockwise for this solution; similarly, the second arc is counterclockwise. Thus, they lie in different half-planes of the $x$-axis and this solution cannot intersect itself.

For the second case, we assume that $l_1$ and $l_2$ lie on the same side of the $x$-axis (see Figure 6(b)). Without loss of generality, we assume that they lie above. Consider the solution at $l_2$. Here the second arc is a straight line and the first arc is a clockwise arc. As $l_2$ lies above the $x$-axis, the first and the second arc lie in different half-planes defined by line $v_{i-1}l_2$ and cannot intersect.

By Lemma 1, for any non-planar solution with $v_{i-1}$ and $v_{i+2}$ on opposite sides of $L$, we can determine one or more positions on $L$ that give a planar solution. We select the position $s$ on line $L$ with minimum distance $|s-v'|$. We assume that there are at most two intervals on $L$ which contain planar solutions. If $s$ is in an interval adjacent to the optimal solution $v'$, we use binary search to obtain the planar solution with the least symmetric difference in the interval $[v',s]$. If this is not the case, we apply a binary search to obtain a better planar solution. However, this solution is not guaranteed to be optimal. In our experiments, this occurred only in rather contrived cases.

### 3.2 Iterative schematization

Here we describe an algorithm to perform the operations of the previous section to reduce the complexity of a shape. **Initialization.** The algorithm initializes by computing the result for each possible operation. That is, for all pairs of neighboring arcs we compute the result of the vertex-restricted operation; for each sequence of three arcs we compute the result of the non-vertex-restricted operation. For each operation we compute the region of symmetric
difference with the original arcs and store it with the result. To prevent topologically unsafe operations we count the number of arcs that overlap this region. This number is called the blocking number. An operation maintains planarity of the shape if its blocking number is zero. We call such an operation admissible.

**Iteration.** Each operation is scored based on the symmetric difference between the resulting arcs and the section of the original shape that it represents. Hence, we maintain for each arc in the current shape, the part in the original it represents. Initially, each line segment represents itself. When performing a vertex-restricted move, the new arc represents the union of the parts represented by the replaced arcs. When performing a non-vertex-restricted move, each new arc represents a fraction of the union of the parts that the old arcs represented. We base this fraction on the perimeter length of the new arcs. Note that this does not guarantee an optimal matching, but works well in practice.

The algorithm iteratively selects the admissible operation \( o \) with the lowest score. Before the operation is executed, other operations must be updated. If an operation involves an arc that is also involved in \( o \), it is discarded. Otherwise, the blocking number of the operation is decreased by the number of arcs of \( o \) that overlap its region. Operation \( o \) is now executed, creating new arcs. We update the blocking numbers again, but now increase the values where required. Finally, we construct new operations involving at least one of the newly created arcs and initialize their blocking numbers. The algorithm proceeds until either no operation is possible or a target complexity has been reached. Pseudocode is given by Algorithm 1.

In each iteration, at most 5 arcs affect the blocking numbers of the operations; only a constant number of operations is constructed, taking linear time per operation. Hence, the algorithm runs in \( O(|S|^2) \) time.

Ideally we would prove that any planar shape has at least one admissible operation. Unfortunately, this is not the case. In theory our algorithm could terminate at an undesirably high complexity. However, we did not observe this in practice with territorial outlines: they are typically quite sparse with well-separated boundaries.

**Weighting.** Since our algorithm selects operations based only on their score, we can reweight operations to introduce a preference. We propose the weight \( \overline{c} \), where \( \overline{c} \) denotes the maximum central angle of the new arcs created by an operation and \( c \) is a parameter. We multiply the score given above by this weight. The weight gives a preference based on central angles and yields a more curvy result (large central angles) or less curvy result (small central angles). By varying parameter \( c \), we can aim for a certain style of schematization. For positive \( c \), results with small central angles are preferred over results with large central angles. This results in a flat schematization style: arcs are relatively straight and have low curvature. Negative \( c \) obtains the exact opposite, resulting in a curvy style with arcs that are far from straight. For \( c = 0 \), the weight is always one and no change occurs, no preference is given and the result could be either style or even a mix. We refer to this as regular schematization. This weighting scheme only steers the selection of operations and does not give a guarantee on the resulting shape. In particular, it does not behave monotonously in \( c \); for example, \( c = 2 \) does not necessarily yield a result that is less curvy than \( c = 1 \). However, in our experiments, we found that the preferred style clearly emerges when using a \( c \) of either -1, 0, or 1 (see Figure 7); we use only these values.

**Alternative termination.** The algorithm stops when the shape has complexity \( k \) or less (or no admissible operation exists). However, it may also be desirable to stop based on some measure of shape similarity instead. In this scenario, we compute for each operation the Fréchet distance of its result to the original part. If this is larger than a given threshold, we disallow the operation. This ensures that the Fréchet distance between the schematization and the input shape does not exceed the given threshold for the maintained map between arc and original. The Fréchet distance between one or two circular arcs and a polygonal line with \( n \) edges can be computed in \( O(n \log n) \) [25]. Therefore, using a bound on the Fréchet distance increases the execution time to \( O(|S|^2 \log |S|) \) for a given input polygon \( S \).

**Subdivisions.** We described our algorithm for simple polygons. However, it can also be applied to planar subdivisions (e.g., multiple countries); we allow only operations that (re)move vertices of degree two. Other vertices are fixed in their given position, but may still be used as endpoints for an operation. An example is given Figure 11.

### 3.3 Postprocessing

**Smoothing.** In our method we fit an arc without considering its neighbors. This may cause very small or large angles between neighboring arcs. Small angles create shallow dents in the outline, causing an increase in complexity without adding to the shape. Very sharp angles can negatively affect

![Figure 7. Three different styles for a 10-arc schematization of Spain.](image-url)
perception of the schematization [1]. To avoid small angles, we use the following postprocessing.

We inspect each vertex \( v \) of the result and check whether the angle between tangents is less than a threshold (we use 20 degrees). If this is the case, we move the vertex to a new position, \( v' \), such that the bend becomes smooth. It is desired that this does not change the tangents at the other endpoints of the involved arcs as it could cause a different bend to lose its smoothness. The solution space for \( v' \) is in fact a circle \( C \) [8]. We choose for \( v' \) the nearest point on \( C \); this fully determines the incident arcs.

This smoothing method is not area-preserving; such a solution need not exist given that tangents are maintained. Our proposed method does not heavily distort the area; this distortion may be acceptable depending on the application. Figure 8 illustrates the benefit of smoothing; the area distortion is less than 0.4%.

**Rendering.** Obtaining a schematization is only a first step in developing a map or visualization. The presentation of the resulting map also plays a crucial role. We present some results obtained by combining our schematizations with different rendering styles.

In Figure 9 we combine the sketchy rendering by Wood et al. [30] with a flat 12-arc schematization of France. We believe the flat schematization style matches well with the hand-drawn style provided by the renderer. A manual sketch usually does not use arcs with a large central angle as these are generally harder to draw. The implied imprecision of the sketch reinforces the geographic inaccuracy of the schematization. In fact, the degree of “sketchiness” may be used to convey the severity of the geographic inaccuracy [3], [30]. In Figure 10 we rendered a regular 12-arc schematization of Vietnam with a variable stroke thickness. This also yields a hand-drawn appearance. However, in contrast to the sketchy style, this style feels more controlled and thus more curviness can be allowed. Inspired by Christophe et al. [5], we apply a pop-art style of rendering to a curvy schematization of Italy (see Figure 11). The high curviness of the drawing is unusual, but this matches the unusual coloring typically seen in pop art.

Much of the pop-art movement (e.g., Lichtenstein) focused on mimicking the automated, often cheap and crude, printing process with carefully crafted hand-drawn paintings. The “unusual” colors used can be seen as a reflection of the simple, bold, three color printing processes. One could argue that the highly curved edges of the Italy example are reflecting and exaggerating that bold simplicity. In almost complete contrast, the sketchy work (and all non-photorealistic rendering) is mimicking hand-crafted drawing with an automated process. Here we deliberately avoid the bold exaggerated curves to emphasize that process.

These preliminary results look promising, but a more thorough study of the relation between rendering and schematization styles is necessary. This need not be limited to only curved schematization, but could also include straight-line schematization.

4 **GEOMETRIC EVALUATION**

In the previous section we introduced a new schematization algorithm. Figure 12 illustrates some results.\footnote{For the interested reader, more results have been made available online at www.win.tue.nl/~Ewemeulema/results.html} To assess the...
quality of our algorithm, we evaluate our design choices by measuring geometric similarity. As we assume a purely geometric input —without semantic annotation— this is a good description of the technical qualities of our algorithm.

We consider four aspects. First, we briefly assess two geometric similarity measures, the Fréchet distance and the symmetric difference. Second, we evaluate how effective our non-vertex-restricted algorithm is at placing new vertices. Third, we determine the effect of our style parameter. Finally, we measure the impact of smoothing on the geometric similarity and the area distortion.

To allow comparison between different results in this section, the Fréchet distance and symmetric difference are respectively stated as a percentage of the diagonal of the minimal bounding box, and of the input area. 

**Measuring accuracy.** To measure the quality of an operation, we use a geometric scoring function. We use the symmetric difference as a similarity measure and the Fréchet distance as an alternative stopping criteria. In Figure 13 the progression of both measures during the schematization process is shown for a regular schematization of Vietnam. The graph shows the difference in nature between both distance measures. The Fréchet distance is a measure of the position of worst fit along the polygon. Operations may only alter the Fréchet distance if they alter or exceed this worst fit. In contrast, the symmetric difference measures the total sum of errors along the polygon. This makes a better differentiation possible between operations, which is reflected in the more consistently rising distance function.

Both the Fréchet distance and the symmetric difference increase nearly monotonously during the execution: both our operations are relatively effective and do not introduce large artifacts that are “hidden” by later iterations.

It is important to note that both measures only capture geometric similarity. Geometric similarity, however, does not always fully capture recognizability. Key recognition points may be critical for a high-quality schematization, even though their impact on the geometric shape is comparatively small. For example, the inclusion of the Thames estuary (due to its correlation with London) may be important for a schematization of Great Britain.

**Introducing new vertices.** The feasible solutions for a non-vertex-restricted algorithm are a superset of those for vertex-restricted algorithms. Consequently, non-vertex-restricted algorithms should be able to generate better, or at least equal, quality schematizations (e.g., Figure 1). As the solution space is infinite, however, any non-vertex-restricted algorithm can only test a subset of the possible vertex placements. Non-vertex-restricted operations should be judged on their ability to find high-quality new vertex positions. Here we evaluate the effectiveness of the non-vertex-restricted operation within our algorithm. The results of our algorithm are compared with results obtained using only vertex-restricted operations. We assume that, by keeping all other variables equal, the resulting difference in quality is due to the non-vertex-restricted operation.

We compare the Fréchet distance and the symmetric difference of solutions with different complexities (see Table 1). For most solutions the Fréchet distance is better for non-vertex-restricted solutions of the same complexity. The same holds for the symmetric difference though the results are less decisive. On first glance, the addition of the non-vertex-restricted operation increases the quality of schematization. Results vary slightly for different complexi-
Table 1. The Fréchet distance and symmetric difference of vertex-restricted (VR) and non-vertex-restricted (NVR) solutions. Results highlighted in bold are shown in Figure 14. For each outline-complexity and distance measure pair the best result is underlined.

<table>
<thead>
<tr>
<th></th>
<th>Fréchet distance</th>
<th>Symmetric difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>VR</td>
<td>NVR</td>
</tr>
<tr>
<td></td>
<td>VR</td>
<td>NVR</td>
</tr>
<tr>
<td>Antarctica (Regular)</td>
<td>6 arcs</td>
<td>7.3</td>
</tr>
<tr>
<td></td>
<td>8 arcs</td>
<td>3.6</td>
</tr>
<tr>
<td></td>
<td>10 arcs</td>
<td>3.2</td>
</tr>
<tr>
<td>Australia (Flat)</td>
<td>6 arcs</td>
<td>6.1</td>
</tr>
<tr>
<td></td>
<td>8 arcs</td>
<td>5.9</td>
</tr>
<tr>
<td></td>
<td>10 arcs</td>
<td>3.5</td>
</tr>
<tr>
<td>Britain (Regular)</td>
<td>6 arcs</td>
<td>11.2</td>
</tr>
<tr>
<td></td>
<td>8 arcs</td>
<td>11.2</td>
</tr>
<tr>
<td></td>
<td>10 arcs</td>
<td>9.5</td>
</tr>
<tr>
<td>China (Curvy)</td>
<td>6 arcs</td>
<td>6.0</td>
</tr>
<tr>
<td></td>
<td>8 arcs</td>
<td>4.9</td>
</tr>
<tr>
<td></td>
<td>10 arcs</td>
<td>4.5</td>
</tr>
<tr>
<td>Florida (Curvy)</td>
<td>6 arcs</td>
<td>13.8</td>
</tr>
<tr>
<td></td>
<td>8 arcs</td>
<td>13.8</td>
</tr>
<tr>
<td></td>
<td>10 arcs</td>
<td>10.1</td>
</tr>
</tbody>
</table>

Ities due to (un)fortunate replacements by either method.

To get a better understanding of these results, further visual inspection of the data is required. We highlight some of the most interesting results (see Figure 14). The schematization of Antarctica with 10 arcs has contradicting quality measurements for the Fréchet distance and the symmetric difference. The difference in Fréchet distance is likely caused by the better fit in the south of Antarctica. Both schematizations have placed emphasis on different parts of the boundary, but the non-vertex-restricted schematization has an overall better representation of the shape. The schematization of China using 6 arcs catches interest as both measures favor the non-vertex-restricted solution. Though the vertex-restricted result more severely distorts the northeastern part, it provides a better fit for the inlets. Finally, we point the attention to the large difference in scoring measures for the Florida schematizations. A visual inspection shows that both scoring measures clearly line up with the visual quality of the schematizations. The introduction of new vertices allows for a better fit of the input, while still maintaining the strongly curved look. The other inputs appear to be inline with the scoring functions and generally favor the non-vertex-restricted schematizations.

We conclude from these results that our non-vertex-restricted operation often benefits the similarity of the result with respect to the input. The results also indicate that geometric similarity is a decent measure of quality, though it is not precise enough to function as a ground truth. We opted for the symmetric difference to score the operations, but any efficiently computable measure could be chosen instead. Improved distance measures, possibly including additional information, may also improve our results.

The effect of style. The presented algorithm allows different styles of curved schematization to be promoted. The chosen style is strongly correlated with the use of the non-vertex-restricted operation (see Figure 15). The percentage of operations that introduce new vertices is largely independent of the polygon shape and remains near constant during the schematization process. This is likely due to the imbalance in our scoring function for the non-vertex-restricted operation. By taking the maximum central angle of the two resulting arcs, we enforce a more...
strictly conforming style for the flat schematization. Flat schematization requires that no strongly curved arcs are present in the solution. Curvy schematization, in contrast, does not require only strongly curved arcs. This imbalance is reflected in the use of the non-vertex-restricted operation.

While the curvy style introduces more vertices, this is not to say these are all essential. While some of the operations capture more complex shapes in a single iteration (see Figure 16(a)), some operations are solely a result of promoting the more strongly curved arcs (see Figure 16(b)). Slightly shifting vertices, however, is not deemed an artifact as it strictly matches the goals we set for the curved style without strongly affecting the resulting schematization.

The effect of smoothing. As a postprocessing step, we smooth the result by displacing vertices where the angle between the tangents of the incident arcs is smaller than a given parameter (the smoothing angle). As we increase the smoothing angle from 0° to 180°, more vertices are displaced to create a smooth outline. Figure 17 shows the results using different smoothing angles for a regular schematization of Australia. At 20 degrees three vertices are smoothed (two along the southern and one along the eastern shore) but the overall shape is affected only slightly. When we increase the smoothing angle to 30 degrees a fourth vertex is smoothed, but the geometric similarity is negatively affected for the west coast. If all vertices are smoothed, the shape loses most of its characteristic features and bears little resemblance to the geographic outline.

To investigate the effect of smoothing, we considered four outline-complexity pairs: Australia with 10 and 20 arcs, France with 12 arcs and Vietnam with 15 arcs. For each of these pairs we computed a schematization of the given complexity using the three curved styles. For the resulting 12 schematizations, we increased the smoothing angle from 0 to 180 degrees in increments of one degree. At each step, we measured area distortion (increase or decrease in area), the symmetric difference and the Fréchet distance with respect to the geographic outline (see Figure 18). Note that the results for Vietnam in the curvy style could not be smoothed beyond 124 degrees as smoothing the vertices would introduce intersections.

Ideally, the area distortion should be 0%, implying area-equivalent results. Our smoothing method, however, does not preserve area. We observe that smoothing shallow bends does not cause a large area distortion on the evaluated results. As the smoothing angle increases, the area distortion becomes more severe, in some cases even rising over 20%. In many cases, fully smooth schematic outlines may not be desirable as some characteristic features are defined by sharp angles. For example, the northern tip of Northern Territory in Australia is best represented using a sharp bend. A very small deviation (below 0.5%) may be observed for the unsmoothed results. This is caused by the imprecision in computer arithmetic and by the need to use numerical methods to compute replacement arcs (see Appendix). We believe that this is within a tolerable range for "area-
equivalent” results. If desired, the distortion can be reduced by using more precise arithmetic and numerical methods.

Area distortion does not reflect geometric similarity to the geographic outline. Therefore, we also measured the symmetric difference and the Fréchet distance. Both measures indicate a decrease in similarity for smoothing angles over 30 degrees. We opt to use a smoothing angle of 20 degrees, to err on the side of caution and prevent unwarranted distortion. The area distortion provides a lower bound on the symmetric difference. However, even cases that have a low area distortion may exhibit a large symmetric difference (e.g., Vietnam in the curvy style).

5 USER STUDY

In the previous section we evaluated various geometric aspects of our algorithm. Here, we wish to investigate the potential of curved schematization as a concept, rather than as a product of our algorithm. To assess both the appeal and the utility of curved schematization, we conducted a user study to evaluate three hypotheses.

1) People prefer the visual appeal of curved schematizations of shapes over their straight-line equivalents.

2) Curved schematization improves the recognizability of shapes for any given degree of simplification.

3) Curved schematization creates shapes that are judged to be simpler than their straight-line equivalents.

We explore these hypotheses to support or refute our initial assumption that creating curved rather than straight-line schematizations, increases user engagement (visual appeal) and reduces visual clutter (simplicity), while maintaining the information-carrying capacity (recognizability).

5.1 Generating schematizations

In our study we compared four styles of schematization.

Curved images. We generated curved schematizations with our algorithm in the three styles: curvy \((c = -1)\), flat \((c = 1)\), and regular \((c = 0)\). To all results, we applied the smoothing step. Due to an implementation error, the figures used in this study were computed using the sum of central angles in the weighting scheme, instead of the maximal central angle. This does not affect the figures for the regular (unweighted) style. For the flat and curvy style, the use of non-vertex-restricted operations is more strongly discouraged or encouraged respectively. For a flat style mostly vertex-restricted operations are used (see Section 4), thus the effect on these shapes is relatively minor. For the curvy style, the effect is more pronounced as a large number of non-vertex-restricted operations is used. Maximum weighting produces better results (see e.g. Figure 19). Therefore, the potential of the curvy style is somewhat “underestimated” in this study.

Straight-line images. To minimize influences due to the type of algorithm used, we chose to use the same algorithm for straight-line schematizations as we used for the curved schematizations. With straight lines, it is not possible to construct an area-preserving operation that replaces two straight lines by one. That is, there is no direct equivalent with straight lines to our vertex-restricted operation with arcs. The non-vertex-restricted operation, however, does admit a straight-line variant with a uniquely defined optimal solution. The solution space for the connecting vertex \(v'\) is a line \(L\) parallel to the line through \(v_{i-1}\) and \(v_{i+2}\). We determine the optimal position on \(L\) using a golden search. The initial search interval is determined by projecting \(v_{i-1}\) and \(v_{i+2}\) onto \(L\). While the center of this interval has a score higher than one of the endpoints, this endpoint is shifted along \(L\) to increase the size of the search interval by \(\|s - t\|\). This ensures that the search interval contains a (locally) minimal value. We did not apply the smoothing step for straight-line results as “smoothing” a vertex would be equivalent to removing it.

5.2 Experimental protocol

We constructed trials using a range of shapes based on country outlines with varying degrees of schematization and curviness parameterizations. We recruited unpaid volunteers to take part in an unsupervised online survey, largely through social networks and email lists. The estimated completion time was about 15 minutes per person. We used a mixed-design varying shape and schematization style in three tasks. Before the main evaluation tasks, we asked for basic demographic and background information: age, gender, background ("visual art/design", “geography/cartography”, “computer science/IT” and “other”) and experience with geographic maps, schematized maps, and schematized representations. This allowed us to define the term “schematized” before it was used in subsequent tasks. While asking these questions first may risk introducing some bias, we considered the benefit in clarifying the terms used in the tasks as outweighing this potential cost.

Task 1 involved selecting one of two presented shapes as the one they found more aesthetically appealing. Specifically, users were asked to “Click on the image that you consider aesthetically more appealing. Do not overthink your answers and try to decide within a few seconds.” Each user was given 20 comparisons. The images were generated from 6 different country outlines in four different styles and two complexity levels (6- or 7-arc schematization and 11- or 12-arc schematization). This yielded 48 images in total. The pair of images being compared were of the same country outline, but with varying combinations of schematization styles. Thus, there were 288 different pairwise comparisons from which the sample of 20 was drawn without replacement (144 distinct pairs with randomized position).
Task 2 involved matching a schematized shape with one of four unschematized possible originals. Specifically, users were told “In this section, we show you a single schematic image. Below this image are four shapes. The schematic image represents one of these four shapes. Click on the shape that is represented by the schematic.” The four alternatives were displayed in a random order and represented different, but equivalently complex originals. They were constructed to resemble typical but fictitious country outlines (see Figure 20). For these 24 outlines, we generated 5-arc, 10-arc, and 15-arc schematizations in each of the four styles. The schematized stimulus outline was drawn from a random selection without replacement of 288 possibilities (6 “countries”, 4 variants per “country”, 4 styles and 3 complexity levels) but were selected to ensure even distribution over the complexity levels. Each user was given 18 questions for Task 2. This task counterbalances for oversimplification (e.g. representing everything with a perfect circle). We chose a high number of outlines to avoid biasing the outlines to a particular style.

Task 3 involved identifying which of two possible alternative shapes they regarded as the “simpler”. Specifically, users were asked to “Click on the image that you consider “simpler”. Do not overthink your answers or count elements. Give your intuitive answer and try to decide within a few seconds.” Random shapes were drawn from a pool of 72 (6 countries, 4 schematization styles, 3 levels of Fréchet distance). The thresholds on the Fréchet distance were 0.015, 0.03 and 0.06 times the diameter of the input, yielding “high”, “medium” and “low” levels of similarity (to the input) respectively. The same level of similarity and same basic country outline were used in any given pairwise comparison leading to a possible pool of 216 question combinations (108 distinct pairs with randomized position). Each user was given 20 comparisons that were selected without replacement. Figure 21 illustrates the various styles and levels of similarity. We chose to use a consistent level of similarity instead of complexity. A single arc may be considered more complex than a single line segment; however, arcs have a higher expressive power. Using a consistent complexity risks biasing judgements towards straight-line schematizations, though these sacrifice similarity (recognizability) to achieve this simplicity.

5.3 Results
Of the 322 people who attempted the questionnaire, we processed the results from the 303 who completed all questions. This provided 6060 responses for Task 1 (visual appeal), 5454 responses for Task 2 (recognizability) and 6060 for Task 3 (simplicity). The breakdown of respondents by age, gender and background is shown in Figure 22. Caution has to be exercised in relating responses to background because there is some dependency between gender and background as well as age and background (e.g. the majority of male participants are computer scientist).
group of 68 years and older had only 3 participants and was excluded from analysis.

To investigate Hypothesis 1, we used loglinear Bradley-Terry (LLBT) modeling [12] of the 6060 pairwise aesthetic preference comparisons to produce ranked “worth” scores for each of the four schematization styles. The worth score allows the consistency of preference to be assessed in forming an overall ranking of the four classes. Figure 23 shows the ranking of the four styles in terms of aesthetic preference, broken down by both gender and age. Consistently, the straight schematization was regarded as the least aesthetically attractive. There was a significant influence of gender on the most preferred style with women preferring the curvy style over other curved styles and men preferring the slightly less exaggerated regular and flat styles. This difference may in part also reflect the differing backgrounds of men and women in the study, but we conclude from this work that while there is strong evidence for a preference of curves over straight lines, there appears to be no universal agreement over which style of curviness is preferred.

In contrast, there was much stronger agreement among respondents as to what style of schematization appears to be the “simplest”. All age groups regarded the (strongly curved) curvy schematization as being the least simple and the straight-line representation to be among the simplest (see Figure 24). This suggests we can reject Hypothesis 3 that curved schematization would be judged as simpler than its straight-line equivalents. Future work might explore whether self-reported judgements of simplicity match the cognitive load required in interpreting such shapes.

We analyzed the success of respondents in matching schematized to non-schematized shapes (Task 2) to investigate the Hypothesis 2 on recognizability. Answers with response time over 4 minutes were excluded. Figure 25 shows accuracy and response time as the complexity and schematization style was varied. Chi-squared tests were applied to the schematization type, complexity, as well as both in combination. As expected, as the complexity increases, the ability to correctly match shapes decreases and the time taken to make a matching judgement rises (Chi-squared $p < 0.01$). For very highly schematized shapes (5 arcs) as well as those with the least schematization (15 arcs), the ability to match successfully is somewhat independent of style of schematization. This is largely expected since the matching task is either too challenging or too easy under these conditions. For 10-arc schematizations though, there was a significant difference in matching ability between schematization types (Chi-squared $p < 0.01$) with all curved styles improving the ability to match shapes over their straight-line equivalent. This suggests that at the right complexity, curved schematization may significantly improve the recognizability compared to straight-line simplifications. However, this must be balanced against the perception that straight-line shapes appear simpler than curved shapes (as demonstrated via Task 3).

This study may slightly underestimate the performance of the curvy style, due to the weighting scheme used (see Section 5.1). Although we expect that this does not overly affect the results, further studies on curved schematization are desirable. Curved schematization is still not fully understood and a more extensive study would be very interesting.

6 Conclusions

We presented an algorithm for automated curved schematization using circular arcs. Our algorithm preserves area
and topology and is able to introduce new vertices. Using a single parameter we can steer the algorithm to obtain “flat”, “regular” or “curvy” results. The schematization style can be reinforced by choosing an appropriate rendering style. We illustrated this by combining three rendering styles with our schematizations. The results are visually pleasing and seem to imply that indeed an interaction exists between rendering and schematization style.

We evaluated the geometric accuracy of our algorithm. The results indicate that the introduction of new vertices improves the geometric accuracy of the resulting schematizations. A threshold of 20 degrees is suitable to ensure that polygon area is not heavily distorted during smoothing.

We conducted a user study to investigate what level of curviness is preferred in a schematization. Users were asked to determine preference and visual complexity of schematizations of different territorial outlines and curviness. We also tested the recognizability depending on the curviness and the degree of schematization.

The results seem to indicate that the use of curves is preferred in schematizations. Straight-line schematizations were consistently deemed the least aesthetically pleasing. (However, no stylization such as parallelism [22] or octilinearity [4] was used.) There was no concluding proof that a specific type of curviness was deemed more pleasing. Although curved schematizations were preferred aesthetically, straight-line schematizations were deemed visually less complex. This suggests a trade-off between visual complexity and aesthetically pleasing results.

The use of curves has a significant effect on the recognizability of schematized shapes. While recognizability was neither helped nor hindered by the use of curves at low and high complexities, for the mid-level schematizations the use of curves increased recognizability. We conjecture that within a certain range of schematization, curves are better at characterizing shape. This would imply that the use of curves is not only aesthetically pleasing, but increases the information carrying capacity of a schematization.

**Future work.** In this paper we have explored the concept of curved schematization design and compared it to straight-line schematizations. It would also be interesting to investigate how our algorithm compares to other schematization algorithms, e.g. [4], [6], [11].

The quality of schematization may depend on the chosen style. Flat curved schematization works for most shapes, while the curvy style seems unsuitable slightly more often. For example, a curvy schematization of France (Figure 26) does not capture its shape well. It is unclear how to determine beforehand which type of schematization is suitable for a given shape; we leave this to future work.

Further research into the use of curves to represent shape is necessary, since so far most efforts have focused on straight-line representations. Our user study seems to imply that the use of curves might allow for a higher quality representation and as such promotes the use of curves in future research. However, not just circular arcs, but also other types of curves (e.g., elliptical arcs, Bézier curves) could prove interesting.

**APPENDIX**

We study the following problem: given two points \( u \) and \( v \) and a value \( A \), find the arc from \( u \) to \( v \) with signed area \( A \).

If \( A \) is zero, the solution is trivial (and degenerate): it is the line segment connecting \( u \) and \( v \). If \( A \) is positive, the arc must be counterclockwise; if \( A \) is negative, the arc must be clockwise. These cases are symmetric and thus we assume without loss of generality that \( A \) is positive.

Let \( s = \frac{\|u - v\|}{2} \) be the point halfway between \( u \) and \( v \) and let \( t = (u_y - v_y, v_x - u_x) \) be the direction from \( u \) to \( v \) rotated counterclockwise by 90 degrees. The bisector of \( u \) and \( v \) is given by \( s + \tau \cdot t \) and defines all possible center points for arcs from \( u \) to \( v \). The central angle \( \alpha \) varies monotonously increasing from 0 at \( \tau = -\infty \) to \( 2 \cdot \pi \) at \( \tau = \infty \). If we know \( \alpha \), then computing the desired arc is straightforward. However, it is not straightforward to compute \( \alpha \) for given \( u \), \( v \), and \( A \). We wish to solve the following equation for \( \alpha \):

\[
\frac{8 \cdot A}{d^2} = \frac{\alpha - \sin \alpha}{\sin^2 \frac{\alpha}{2}},
\]

where \( d = \|v - u\| \) is the distance between \( u \) and \( v \). Unfortunately, this seems algebraically impossible. However, numerical methods can be applied to solve the equation since the lefthand side is constant and the righthand side is monotonously increasing in \( \alpha \). This monotonicity also implies that the solution is unique. What remains is to argue that this claim of monotonicity is indeed true.

Intuitively, increasing the central angle of an arc without changing its endpoints increases its signed area. This is because the circular segment of a higher sagitta is a strict superset of the circular segment of a lower sagitta.

More formally, we argue as follows. Consider two central angles \( \alpha_1 \) and \( \alpha_2 \), assuming \( \alpha_1 < \alpha_2 \). Let \( a_1 \) and \( a_2 \) denote the corresponding arcs. Since circles intersect at most twice, these arcs also intersect at most twice. As they share endpoints \( u \) and \( v \), the arcs cannot intersect anywhere else. What remains to be argued is that some point on \( a_1 \) lies in the circular segment of \( a_2 \). Basic geometry shows that the tangent of \( a_1 \) (resp. \( a_2 \)) at \( u \) makes an angle of \( \frac{1}{2} \alpha_1 \) (resp. \( \frac{1}{2} \alpha_2 \)) to the line segment \( uv \). Since \( \alpha_1 < \alpha_2 \), these tangents cause \( a_1 \) to be inside the circular segment of \( a_2 \) near \( u \). Since no more intersections can occur (and no arc can cross the line between \( u \) and \( v \)), we know that \( a_1 \) lies fully in the circular segment of \( a_2 \). Hence, the circular segment of \( a_1 \) is contained in the circular segment of \( a_1 \) and therefore, the signed area of \( a_2 \) is larger than the signed area of \( a_1 \). This proves monotonicity.

---

![Figure 26](image_url)

6-arc France: regular (middle) and curvy style (right). France is unsuitable for the curvy style produced by our algorithm.