

# Orthogonal Schematization with Minimum Homotopy Area

Bram Custers<sup>1</sup>, Jeff Erickson<sup>2</sup>, Irina Kostitsyna<sup>1</sup>,  
Wouter Meulemans<sup>1</sup>, Bettina Speckmann<sup>1</sup>, and Kevin Verbeek<sup>1</sup>

1 TU Eindhoven, the Netherlands

[b.a.custers|i.kostitsyna|w.meulemans|b.speckmann|k.a.b.verbeek]@tue.nl

2 University of Illinois at Urbana-Champaign, United States of America

jeffe@illinois.edu

## 1 Introduction

Visualizing data in its geographic context is useful for exploration, analysis and communication. Thematic maps show data as layers on top of a base cartographic map. However, in many cases high accuracy of spatial locations is restrictive and only of secondary concern for visualization: the spatial dimension may be distorted to further emphasize the data itself. This process of simplifying beyond the needs of a target scale is called cartographic *schematization* [12], and is often accompanied by stylistic restrictions on the used geometric primitives. The prototypical example is the London Underground map with its iconic octilinear style.

We typically distinguish schematization types based on the object to be schematized and the geometry restrictions. Objects are typically either graphs (e.g. transit networks) or polygons (e.g. territorial outlines such as countries). Geometry restrictions are often formulated via a set of admissible orientations [2, 11, 10], with orthogonal, hexilinear and octilinear being the most common, using angles that are a multiple of 90, 60 or 45 degrees. Alternatives include curves, popularized by Roberts [13], such as Bézier curves [6, 15] or circular arcs [7, 14, 15, 16]. Here we focus on orthogonal schematization of polygons.

When schematizing polygons, there are two main quality criteria beyond using few geometric primitives [2]: (1) the result should be a simple polygon, if the input is simple; and (2) the result should resemble the input polygon, as the schematic often functions as a base map for thematic data. Requiring simplicity of the output often causes simplification and schematization problems to be NP-hard [1, 8, 9]. Moreover, the similarity measure used to capture resemblance has a large impact on the visual quality of the schematic. This schematic may also exhibit undesirable behavior when optimizing for a measure, even in the most restricted orthogonal case [2]. Both factors have led to various heuristics being developed to schematize polygons either enforcing simplicity [2] or permitting self-intersections [4].

**Problem definition.** Given a simple orthogonal polygon  $P$  with  $n$  vertices and an integer  $k < n$ , compute an orthogonal polygon  $S^*$  with at most  $k$  vertices such that  $\sigma(P, S^*)$  is minimized. Here,  $\sigma$  denotes the minimal homotopy area [3] between the two polygons; see Section 2 for definitions. We may require  $S^*$  to be a simple polygon or allow self-intersections.

**Contributions.** In Section 3 we illustrate the differences in homotopy area between the simple and nonsimple case, and lower bound the number of intersections in a nonsimple optimal solution. In Section 4 we present a dynamic program to compute in  $O(n^5 k)$  time the optimal solution which may self-intersect. Details and full proofs will be available in the full version of this paper.

36th European Workshop on Computational Geometry, Würzburg, Germany, March 16–18, 2020.

This is an extended abstract of a presentation given at EuroCG'20. It has been made public for the benefit of the community and should be considered a preprint rather than a formally reviewed paper. Thus, this work is expected to appear eventually in more final form at a conference with formal proceedings and/or in a journal.

## 2 Preliminaries

**Ortho-polygons.** We consider only orthogonal polygons and polylines, which we abbreviate to ortho-polygons and ortho-polylines. We use  $P = \langle v_1, \dots, v_n \rangle$  to denote a simple ortho-polygon with  $n$  edges. Edge  $e_i$  connects vertices  $v_i$  and  $v_{i+1}$ , where  $v_{n+1} = v_1$ . A schematization of  $P$  is an ortho-polygon  $S$  with at most  $k$  edges. The complexity of an ortho-polygon or ortho-polyline is its number of edges.

**Homotopy area.** Homotopy area [3], measures the similarity between curves. To define this measure, interpret polygons  $P$  and  $S$  as continuous functions mapping the unit interval  $[0, 1]$  to  $\mathbb{R}^2$  for an arbitrary fixed starting point on the polygon. A *homotopy*  $H : [0, 1] \times [0, 1] \rightarrow \mathbb{R}^2$  between polygons  $P$  and  $S$  is defined as a continuous deformation from  $P$  to  $S$  over a time  $t \in [0, 1]$  such that  $H(a, t = 0) = P(a)$  and  $H(b, t = 1) = S(b)$ . Here,  $a$  and  $b$  are the parameters for the parameterization of  $P$  and  $S$ .

The homotopy area of  $H$  is defined as the total area that is swept by the deformation, with multiplicity. The minimal homotopy,  $H^*$ , between two curves is a homotopy with the smallest homotopy area; the minimal homotopy area,  $\sigma(P, S)$ , is then the area swept by  $H^*$ .

As shown by Chambers and Wang [3], the minimal homotopy between two *simple* curves can be decomposed into smaller *subhomotopies* that deform the curve in a consistent direction relative to the deforming curve. These subhomotopies are delimited by *anchorpoints*, which are a subset of the intersections between  $P$  and  $S$ . These anchorpoints are stationary in the minimal homotopy and must therefore occur in the same order along both curves (Observation 3.1 in [3]). The subhomotopies are minimal homotopies for the curves between the delimiting anchorpoints. Moreover, any minimal homotopy without (or between) anchorpoints is *sense preserving* (Lemma 3.2 in [3]): intuitively, during the morph between the two curves, a point either consistently moves locally to the left (or stands still) or consistently to the right (or stands still). A careful inspection of the proofs in [3] reveals that these results also hold for the case of self-intersecting curves, as the arguments are fully local.

The homotopy area of any subhomotopy is given by the areas of the cells of the arrangement induced by the affected subcurves, multiplied by their *winding number* (Lemma 4.3 in [3]). The winding number of a cell is intuitively defined as the (possibly negative) number of counterclockwise rotations one makes when standing at a point in the cell and following the boundary of the curve.

Formally, homotopy area requires the given curves to be at least  $C^1$  continuous, for the derivatives to be well defined. This is not the case for polygons, but we may imagine each corner to be an infinitesimally small smooth curve instead; see also [5].

## 3 Comparing simple and nonsimple schematization

**Simplicity.** Is the optimal schematization  $S$  always simple if the input polygon  $P$  is simple? Unfortunately, this is not the case. Let  $P$  be an ortho-polygon with a ratio  $\delta \ll 1$  between the longest and shortest edge length (see Figure 1). Let  $S$  and  $S'$  be the optimal simple and nonsimple schematization of  $P$ . In the worst case,  $\sigma(P, S)/\sigma(P, S') = \Omega(1/\delta)$ .

**Self-intersections.** The optimal nonsimple schematization  $S'$  can have  $\Omega(k^2)$  self-intersections, asymptotically matching the trivial upper bound (see Figure 2). As length  $d$  can be arbitrarily large with respect to area  $A$ ,  $S'$  smooths out the top lines. For  $\kappa$  repetitions of the top lines and  $\kappa$  repetitions of the comb-like structure below, we obtain  $\Theta(\kappa^2)$  intersections. With  $k = \Theta(\kappa)$  and  $n = \Theta(\kappa)$ , the lower bound follows.

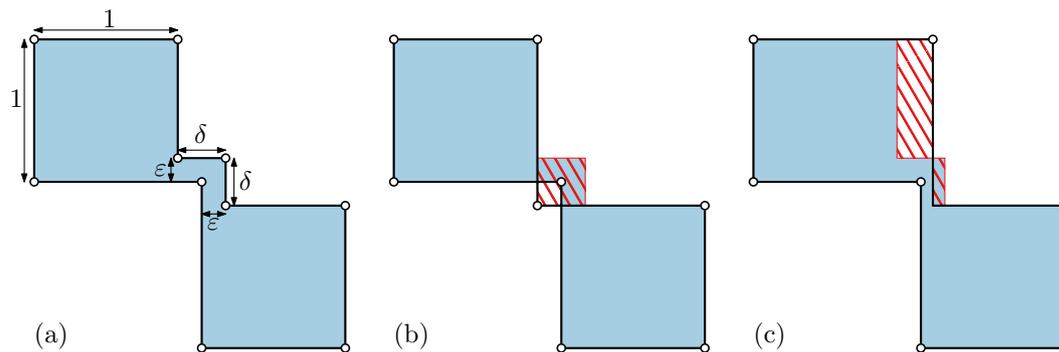


Figure 1 (a) Ortho-polygon with 10 edges. (b) Optimal nonsimple schematization  $S'$  with 8 edges self-intersects and has homotopy area  $\delta^2$ . (c) Optimal simple schematization  $S$  has homotopy area greater than  $(\delta - \varepsilon) \cdot (1 - \varepsilon) = \Omega(\delta)$  for small  $\varepsilon$ .

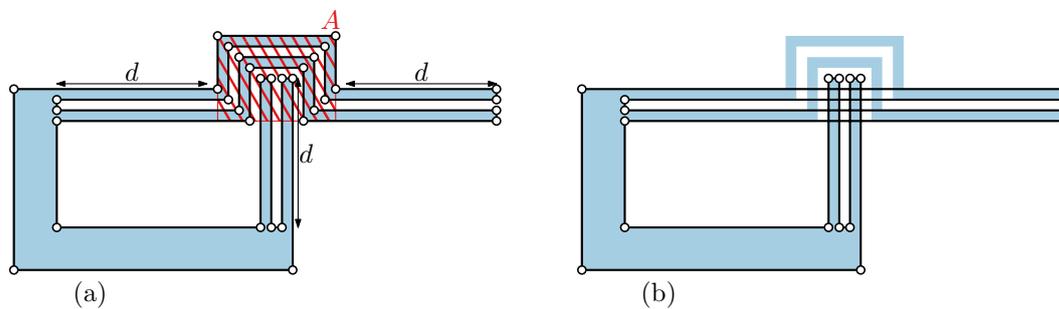


Figure 2 (a) Ortho-polygon with  $n = 34$  and  $\kappa = 4$ . Length  $d$  is larger than area  $A$ ; drawing shows  $d$  to be smaller for clarity of illustration. (b) Optimal nonsimple schematization  $S'$  for  $k = 18$ , with  $\kappa^2 = \Omega(k^2)$  intersections.

## 4 Dynamic program for nonsimple $S^*$

We now compute an optimal schematization  $S^*$  for a given value of  $k$  that is allowed to self-intersect. Leveraging Chambers and Wang [3], we first prove that optimal solutions have a canonical form. Then, we discuss a dynamic program that establishes the following result.

► **Theorem 4.1.** *Given a simple ortho-polygon  $P$  with  $n$  vertices and an integer  $k < n$ , we can compute the schematization  $S^*$  with minimal homotopy area in  $O(n^5k)$  time, if  $S^*$  is allowed to self-intersect.*

### 4.1 Canonical form

► **Lemma 4.2.** *Each edge of  $S^*$  has at least one anchorpoint.*

**Proof sketch.** If an edge  $e$  has no anchorpoints, then it must be part of a single subhomotopy of the minimal homotopy. By sense preservation (Lemma 3.1 of [3]), moving  $e$  in the direction of sense preservation decreases the subhomotopy area. ◀

Hence, in an optimal solution, subhomotopies span at most two edges in  $S^*$  and the matching subcurve is simple. Thus, self-intersections of  $S^*$  can occur only between edges separated by anchorpoints. An *anchorsegment* is a nonempty subsegment of an edge  $e$  of  $S^*$  that coincides with an edge  $e'$  of  $P$ , such that all points on this anchorsegment are anchorpoints; the direction of  $e$  and  $e'$  must match, if this segment is more than a single point.

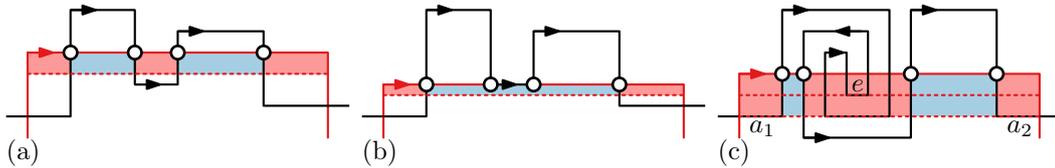
► **Lemma 4.3.** *Each edge of  $S^*$  has an anchorsegment of non-zero length.*

**Proof.** For a contradiction, assume we have an edge  $e$  in  $S^*$  that does not overlap an edge of  $P$  with the same direction. Let  $H^*$  denote the minimal homotopy between  $S^*$  and  $P$ . By Lemma 4.2, we know that  $e$  has one or more anchorpoints and thus we may consider the two or more subhomotopies  $H_1, \dots, H_m$  that involve part of  $e$ . Each  $H_i$  is between two simple curves and thus sense preserving (Lemma 3.2 of [3]).

Without loss of generality, assume  $e$  is horizontal. As the subcurves are simple, each subhomotopy area is the multiplication of area and winding number, summed over all faces in the arrangement (Lemma 4.1 and Lemma 4.3 of [3]). Consider moving  $e$  up or down: this move may change the homotopy area for each  $H_i$ , but cannot cause local intersections in the subcurves. Hence, the total change  $\Delta h$  in the area of all minimal subhomotopies is the change in face area with winding-number multiplicity. Either  $\Delta h$  is zero in both directions, or  $\Delta h$  is positive in one direction and negative in the other.

In the latter case, we find a contradiction with the optimality of  $S^*$ , so assume  $\Delta h = 0$ . We may freely move the edge up or down, until one of the considered arrangements changes; see Figure 3(a,b). At this moment  $e$  must overlap some edge  $e'$  of  $P$ , also considered in one of the original subhomotopies. Let  $S'$  denote the new schematization, with minimal homotopy  $H'$ . Now, either (a part of) this overlap is stationary in  $H'$  and thus an anchorsegment, or the entire edge still moves in  $H'$ . The former case implies an overlap in more than a single point: otherwise,  $\Delta h$  does not change. That is, the arrangement may have a face split, but these have the same winding numbers. In the latter case, we can continue shifting our edge  $e$  as  $\Delta h$  is zero (or positive in the same direction, contradicting optimality). See Figure 3(c) for an example.

Note that we cannot make an edge of  $S^*$  disappear during this motion. ◀



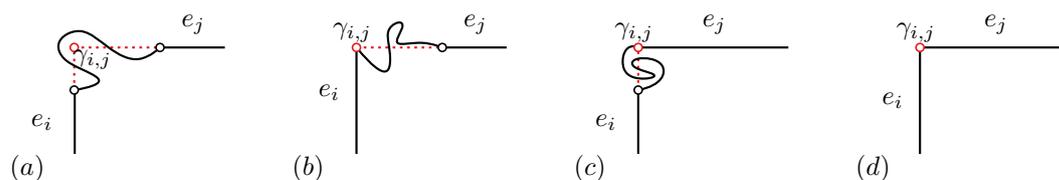
■ **Figure 3** Moving an edge in the schematization. The black line shows part of the input polygon; the red line is part of the schematization. Dots indicate anchorpoints. (a) Moving the horizontal edge down to the dashed line decreases (red) and increases (blue) homotopy area for subhomotopies. The areas are exactly balanced:  $\Delta h = 0$ . (b) The edge coincides with an edge of the polygon, causing the arrangement of the middle subhomotopy to change: its cell collapsed. Moving the edge further down increases homotopy. (c) Moving the edge down causes the edge to overlap with input edge  $e$ . This edge moves in the underlying homotopy and thus can be ignored. Moving the edge further down results in anchor segments  $a_1$  and  $a_2$ .

Lemma 4.3 implies a canonical form for  $S^*$ , in which each edge is anchored to an edge of  $P$  through its anchorsegment. Consequently, every vertex of  $S^*$  lies on the grid  $G$  induced by the edges of  $P$ . The *forward halfline* of an edge  $e_i$  is the halfline originating from  $v_i$ , overlapping  $e_i$ . Similarly, the *backward halfline* of an edge  $e_i$  originates from  $v_{i+1}$ , overlapping  $e_i$ . An edge  $e$  of  $S^*$  must start on the backward halfline of its anchored edge in  $P$  and end on the forward halfline, and the direction of  $e$  matches the direction of its anchored edge.

### 4.2 Dynamic Programming

First, we pick a midpoint of an edge of  $G$  that is on an edge of  $P$  to cut  $P$  into an ortho-polyline  $P'$ . In the DP, we select only vertices of  $G$  as intermediate vertices. Thus, an optimal schematization of  $P'$  with  $k + 1$  edges is a schematization of  $P$  with  $k$  edges. Testing all  $O(n^2)$  midpoints yields the optimal schematization  $S^*$ , as any anchorsegment must contain such a midpoint. In the remainder  $P$  is an ortho-polyline, to be schematized with  $k$  edges.

**Partial solutions.** A *partial solution* is an ortho-polyline  $S$  where each edge has an anchorsegment. The first anchorsegment starts at  $v_1$  on  $e_1$ . The last anchorsegment anchors to some edge  $e_i$ , but is not yet complete. Instead, the last vertex of  $S$  is a gridpoint  $\gamma_{i,j}$  of  $G$  on the backward halfline of  $e_j$  and forward halfline of an  $e_i$ ,  $i < j$ . If  $\gamma_{i,j}$  lies on  $e_j$ , then the anchorsegment must contain  $\gamma_{i,j}$ ; otherwise, the start of  $e_j$  must eventually be part of the anchorsegment. Between two anchorsegments of  $S$  we can compute the minimal subhomotopy area, even if the last is not yet complete. That is, consider two subsequent anchorsegments of  $S$ , anchored to edges  $e_i$  and  $e_j$  of  $P$  with  $i < j$ . The corresponding subhomotopy  $\sigma_{i,j}$  can be computed purely from this information. This follows from one of four cases (see Figure 4) depending on the intersection  $\gamma_{i,j}$ . If the halfines do not intersect, we call such a pair incompatible and use  $\sigma_{i,j} = \infty$ . Subhomotopy areas  $\sigma_{1,i}$  and  $\sigma_{j,n+1}$  are independent of the choice of starting point  $v_1$ , though the corresponding  $\gamma$ -values do. As anchorsegments are directed and ordered, we need to consider only pairs of compatible edges ( $\gamma_{i,j}$  exists).



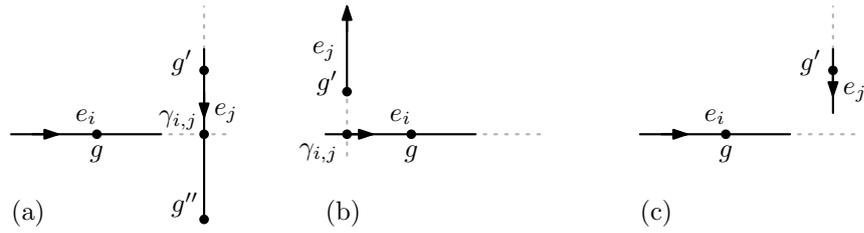
■ **Figure 4** Four cases for compatible edges  $e_{i'}$  and  $e_i$  to compute  $\sigma_{i',i}$ . The gridpoint  $\gamma_{i',i}$  (red marker) can be outside both edges (a), on one of both edges (b,c), or on both edges (d).

However, we must be careful not to reverse an edge of  $S$ . Suppose  $S$  ends with edges anchored at  $e_{i''}$ ,  $e_i$  and  $e_j$ , such that the first two edges are compatible, as well as the last two. If  $\gamma_{i,j}$  is after  $\gamma_{i'',i}$  in the direction of  $e_i$ , then the edge of  $S$  anchored on  $e_i$  is directed along  $e_i$ . However, if this is not the case, this edge of  $S$  is reversed and does not follow the canonical form. Hence, we do not need to consider these cases.

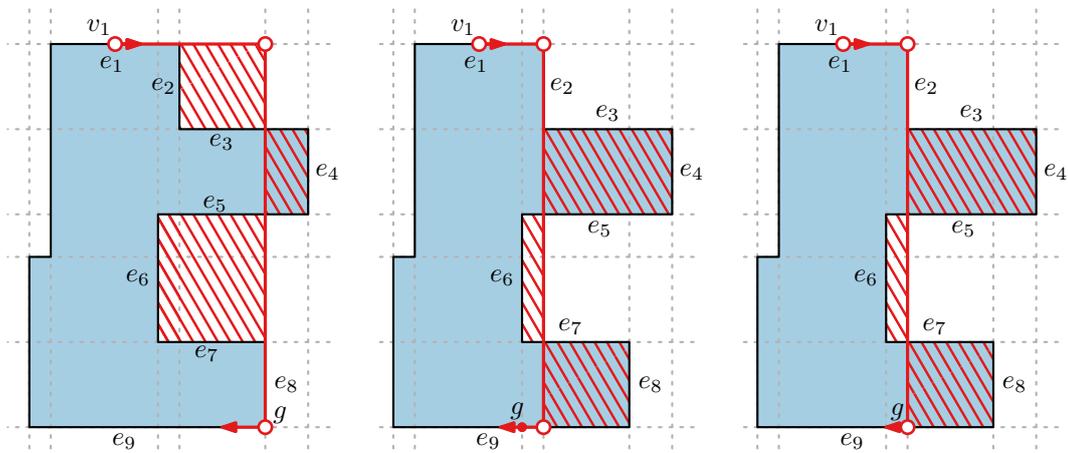
To decide whether we can extend partial solution  $S$  ending at  $e_i$  thus depends on where the last vertex of  $S$  is located with respect to  $e_i$ . We thus pair indices  $i$  with gridpoints  $g'$ . We call a pair  $(i, g')$  a *compatible predecessor* for  $(j, g)$  if  $i < j$ ,  $\gamma_{i,j}$  occurs after  $g'$  on the forward halfline of  $e_i$ , and  $\gamma_{i,j}$  does not occur before  $g$  on the backward halfline of  $e_j$ ; see Figure 5 for examples. Observe that the latter two conditions are independent of each other.

**Dynamic program.** We characterize a subproblem of our DP as  $D[j, g, l]$ : the minimal homotopy area for the optimal partial solution  $S$  with at most  $l$  edges, such that  $g$  lies on the backward halfline of  $e_j$  and the last vertex of  $S$  is not before  $g$  on this backward halfline. Figure 6 shows an example for such a partial solution. The main question is how to compute  $D[j, g, l]$  based on “smaller” instances  $D[i, g', l']$ . We are effectively choosing anchorsegments one at a time. As these occur in order, smaller instances have  $1 \leq i < j$  and  $l' = l - 1$ .

We should consider compatible predecessors for  $(j, g)$  described by  $D[i, g', l - 1]$ . If  $g''$  is closer to  $\gamma_{i,j}$  than  $g'$ ,  $g''$  is less restrictive for the partial solution and hence  $D[i, g'', l - 1] \leq$



**Figure 5** Examples of compatibility of predecessors. Black dots denote gridpoints, arrows give the directions of the edges. (a)  $(i, g)$  is a compatible predecessor of  $(j, g'')$ , but not of  $(j, g')$ , since  $\gamma_{i,j}$  occurs before  $g'$  in the backward direction of  $e_j$ . (b)  $(i, g)$  is not a compatible predecessor of  $(j, g')$  since  $\gamma_{i,j}$  occurs after  $g$  on  $e_i$  in the backward direction. (c)  $(i, g)$  is not a compatible predecessor of any gridpoint on  $e_j$ , since  $\gamma_{i,j}$  does not exist.



**Figure 6** Partial solutions for input  $P$  (blue) with  $v_1$  at the top. Shown in red are  $D[9, g, 3]$  for various gridpoints  $g$ . Homotopy areas are given by the red, hatched area.

$D[i, g', l - 1]$ . Thus, the appropriate neighbor of  $\gamma_{i,j}$  in  $G$  suffices if  $\gamma_{i,j}$  lies on  $e_i$ , or the endpoint of  $e_i$  otherwise; we denote this least restrictive neighbor by  $\lambda_{i,j}$ .

We obtain the following dynamic program. If  $j < l$ , the schematization should have more edges than the input so far, thus we set  $D[j, g, l]$  to  $\infty$ . If  $l = j = 1$ , we have not created a second anchorsegment and thus  $D[j, g, l] = 0$ . Otherwise, we set  $D[j, g, l]$  to  $\min_{1 \leq i < j} \sigma_{i,j} + D[i, \lambda_{i,j}, l - 1]$ , testing all least restrictive compatible predecessors.

**Running time.** We first compute all  $\sigma_{i,j}$  values in  $O(n^4 \log n)$  total time using [3]: each pair of subcurves has  $O(n)$  intersections. For the DP, computing  $\lambda_{i,j}$  takes  $O(1)$  time using  $G$  and we can lookup  $\sigma_{i,j}$ : computing all  $O(n^2 k)$  cells takes  $O(n^3 k)$  time. Running the DP for all  $O(n^2)$  starting points thus takes  $O(n^5 k)$  time.

---

**References**

---

- 1 Quirijn Bouts, Irina Kostitsyna, Marc van Kreveld, Wouter Meulemans, Willem Sonke, and Kevin Verbeek. Mapping Polygons to the Grid with Small Hausdorff and Fréchet Distance. In *Proceedings of the 24th Annual European Symposium on Algorithms*, volume 57 of *LIPIcs*, pages 22:1–22:16, 2016.
- 2 Kevin Buchin, Wouter Meulemans, André Van Renssen, and Bettina Speckmann. Area-preserving simplification and schematization of polygonal subdivisions. *ACM Transactions on Spatial Algorithms and Systems*, 2(1), April 2016.
- 3 Erin Wolf Chambers and Yusu Wang. Measuring similarity between curves on 2-manifolds via homotopy area. In *Proceedings of the 29th Annual Symposium on Computational Geometry*, pages 425–434. ACM, 2013.
- 4 Serafino Cicerone and Matteo Cermignani. Fast and simple approach for polygon schematization. In *Proceedings of the International Conference on Computational Science and Its Applications*, LNCS 7333, pages 267–279. Springer, 2012.
- 5 Brittany Fasy, Selcuk Karakoc, and Carola Wenk. On minimum area homotopies of normal curves in the plane. *arXiv preprint arXiv:1707.02251*, 2017.
- 6 Martin Fink, Herman Haverkort, Martin Nöllenburg, Maxwell Roberts, Julian Schuhmann, and Alexander Wolff. Drawing metro maps using Bézier curves. In *Proceedings of the International Symposium on Graph Drawing*, LNCS 7704, pages 463–474. Springer, 2012.
- 7 Martin Fink, Magnus Lechner, and Alexander Wolff. Concentric metro maps. In *Schematic Mapping Workshop*, 2014.
- 8 Leonidas Guibas, John Hershberger, Joseph Mitchell, and Jack Scott Snoeyink. Approximating polygons and subdivisions with minimum-link paths. *International Journal of Computational Geometry & Applications*, 3(04):383–415, 1993.
- 9 Maarten Löffler and Wouter Meulemans. Discretized approaches to schematization. In *Proceedings of the 29th Canadian Conference on Computational Geometry*, 2017.
- 10 Gabriele Neyer. Line simplification with restricted orientations. In *Proceedings of the 6th International Workshop on Algorithms and Data Structures*, page 13–24. Springer-Verlag, 1999.
- 11 Martin Nöllenburg and Alexander Wolff. Drawing and labeling high-quality metro maps by mixed-integer programming. *IEEE Transactions on Visualization and Computer Graphics*, 17(5):626–641, 2011.
- 12 Andreas Reimer. *Cartographic modelling for automated map generation*. PhD thesis, Technische Universiteit Eindhoven, 2015.
- 13 Maxwell Roberts. *Underground maps unravelled: Explorations in information design*. Self published, 2012.
- 14 Thomas van Dijk, Arthur van Goethem, Jan-Henrik Haunert, Wouter Meulemans, and Bettina Speckmann. Map schematization with circular arcs. In *Proceedings of the International Conference on Geographic Information Science*, LNCS 8728, pages 1–17. Springer, 2014.
- 15 Arthur van Goethem, Wouter Meulemans, Andreas Reimer, Herman Haverkort, and Bettina Speckmann. Topologically safe curved schematization. *The Cartographic Journal*, 50(3):276–285, 2013.
- 16 Arthur van Goethem, Wouter Meulemans, Bettina Speckmann, and Jo Wood. Exploring curved schematization of territorial outlines. *IEEE Transactions on Visualization and Computer Graphics*, 21(8):889–902, 2015.