How to Walk Your Dog in the Mountains with No Magic Leash*

Sariel Har-Peled†
Dept. Computer Science
University of Illinois, Urbana-Champaign
201 N. Goodwin Avenue
Urbana, IL , 61801, USA
sariel@uiuc.edu

Amir NayyeriDept. Computer Science
University of Illinois, Urbana-Champaign
201 N. Goodwin Avenue
Urbana, IL, 61801, USA
nayyeri2@uiuc.edu

Mohammad Salavatipour‡
Dept. of Computing Science
University of Alberta
mreza@cs.ualberta.ca

Anastasios Sidiropoulos
Toyota Technological Institute at Chicago
tasos@ttic.edu

ABSTRACT

We describe a $O(\log n)$-approximation algorithm for computing the homotopic Frechét distance between two polygonal curves that lie on the boundary of a triangulated topological disk. Prior to this work, algorithms where known only for curves on the Euclidean plane with polygonal obstacles.

A key technical ingredient in our analysis is a $O(\log n)$-approximation algorithm for computing the minimum height of a homotopy between two curves. No algorithms were previously known for approximating this parameter. Surprisingly, it is not even known if computing either the homotopic Frechét distance, or the minimum height of a homotopy, is in NP.

Categories and Subject Descriptors

F.2.2 [**ANALYSIS OF ALGORITHMS AND PROBLEM COMPLEXITY**]: Nonnumerical Algorithms and Problems – *Geometrical problems and computations*

General Terms

Algorithms, Theory

Keywords

Frechét distance, Approximation algorithms

*The current full version of this paper can be found online [19].

†Work on this paper was partially supported by a NSF AF award CCF-0915984.

‡Part of this work was done while visiting Toyota Technological Institute at Chicago.

SCGâĂŹ12, June 17–20, 2012, Chapel Hill, North Carolina, USA.

1. INTRODUCTION

Comparing the shapes of curves – or sequenced data in general – is a challenging task that arises in many different contexts. The ***Frechét distance*** and its variants (e.g. dynamic time-warping [22]) have been used as a similarity measure in various applications such as matching of time series in databases [23], comparing melodies in music information retrieval [26], matching coastlines over time [24], as well as in map-matching of vehicle tracking data [4, 28], and moving objects analysis [5, 6]. See [16, 1, 2] for algorithms for computing the Frechét distance.

Informally, for a pair of such curves $f, g : [0,1] \to \mathcal{D}$, for some ambient space $(\mathcal{D}, \mathsf{d})$, their Frechét distance is the minimum length leash needed to traverse both curves in sync. To this end, imagine a person traversing f starting from $f(0)$, and a dog traversing g starting from $g(0)$, both traveling along these curves without ever moving backwards. Then, the Frechét distance is the infimum over all possible traversals, of the maximum distance between the person and the dog. Specifically, given a bijective continuous parameterization $\phi : [0,1] \to [0,1]$. the ***width*** of this reparameterization, i.e., the longest leash needed by this reparameterization, is $\mathrm{width}(\phi) = \sup_{x \in [0,1]} \mathsf{d}\Big(f(x), g(\phi(x))\Big)$, As such, the Frechét distance between f and g is defined to be

$$\mathsf{d}_{\mathcal{F}}(f, g) = \inf_{\phi:[0,1]\to[0,1]} \mathrm{width}(\phi),$$

where ϕ ranges over all orientation-preserving homeomorphisms.

While this distance makes complete sense when the underlying distance is the Euclidean metric, it becomes less useful if the distance function is more interesting. For example, imagine walking a dog in the woods. The leash might get tangled as the dog and the person walk on two different sides of a tree. Since the Frechét distance cares only about the distance between the two moving points, the leash would "magically" jump over the tree.

Homotopic Frechét distance.

To address this shortcoming, a natural extension of the above notion called ***homotopic Frechét distance*** was in-

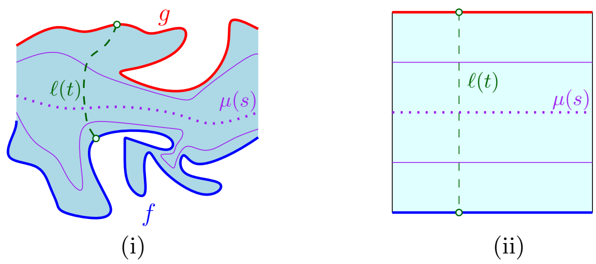


Figure 1: (i) Two curves f and g, and (ii) the parametrization of their homotopic Frechét distance.

troduced by Chambers *et al.* [7]. Informally, revisiting the above person-dog analogy, we consider the infimum over all possible traversals of the curves, but this time, we require that the person is connected to the dog via a leash, i.e. a curve that moves continuously over time. Furthermore, one keeps track of the leash during the motion, where the purpose is to minimize the maximum leash length needed.

To this end, consider a continuous mapping $\psi : [0,1]^2 \to \mathcal{D}$. For parameters $s, t \in [0,1]$ consider the one dimensional functions $\ell_t(y) = \psi(t, y)$ and $\mu_s(x) = \psi(x, s)$. The functions $\ell(y) \equiv \ell_t(y)$ and $\mu(x) \equiv \mu_s(x)$ are parametrized curves that are the natural restrictions of ψ to one dimension, by the x and y coordinates, respectively. We require that $\mu(0) = f$ and $\mu(1) = g$. The ***homotopic width*** of ψ is $\mathrm{width}(\psi) = \max_{t \in [0,1]} \|\ell(t)\|$, and the ***homotopic Frechét distance*** between f and g is

$$\mathsf{d}_{\mathcal{H}}(f, g) = \inf_{\psi:[0,1]^2 \to \mathcal{D}} \mathrm{width}(\psi)\,,$$

where the infimum is over all such mappings, and $\|\cdot\|$ denotes the length of a curve.

Clearly, $\mathsf{d}_{\mathcal{H}}(f, g) \geq \mathsf{d}_{\mathcal{F}}(f, g)$; in fact, $\mathsf{d}_{\mathcal{H}}(f, g)$ can be arbitrary larger than $\mathsf{d}_{\mathcal{F}}(f, g)$. We remark that $\mathsf{d}_{\mathcal{H}}(f, g) = \mathsf{d}_{\mathcal{F}}(f, g)$ for any pair of curves in the Euclidean plane, as we can always pick the leash to be a straight line segment between the person and the dog. In other words, the map ψ in the definition of $\mathsf{d}_{\mathcal{H}}$ can be obtained from the map ψ in the definition of $\mathsf{d}_{\mathcal{F}}$ via an appropriate affine extension. However, this is not true for general ambient spaces, where the leash might have to pass over obstacles, hills, etc.

The homotopic Frechét distance is refereed to as the ***morphing width*** of f and g, and it bounds how far a point on f has to travel to its corresponding point in g under the morphing of ψ [13]. The length of $\mu(s)$ is the ***height of the morph at time s***, and the ***height*** of such a morphing is $\mathsf{height}(\mu) = \max_{s \in [0,1]} |\mu(s)|$. The ***homotopy height*** between f and g bounded by $\ell(0)$ and $\ell(1)$ is

$$\mathsf{h}(f, g, \ell(0), \ell(1)) = \inf_{\mu} \mathsf{height}(\mu)\,,$$

where μ varies over all possible morphings between f and g, such that each curve in μ has one end on $\ell(0)$ and one end in $\ell(1)$. See Figure 1 for an example. Note that if we do not constraint the endpoints of the curves during the morphing to stay on $\ell(0)$ and $\ell(1)$, the problem of computing the minimum height homotopy is trivial. One can contract f to a point, send it to g from the shortest (f, g)-path, and then expand it to g. To keep the notation simple, we use $\mathsf{h}(f, g)$ when f and g have common endpoints.

Intuitively, the homotopy height measures how long the curve has to become as it deforms from f to g, and it was introduced by Chambers and Letscher [8, 9]. Observe that if we are given the starting and ending leashes $\ell(0)$ and $\ell(1)$ then the homotopy height of f and g, is the homotopic Frechét distance between $\ell(0)$ and $\ell(1)$.

Here, we are interested in the problems of computing the homotopic Frechét distance and the homotopic height between two simple polygonal curves that lie on the boundary of an arbitrary triangulated topological disk.

Why are these measures interesting?.

For the sake of the discussion here, assume that we know the starting and ending leash of the homotopy between f and g. The region bounded by the two curves and these leashes, form a topological disk, and the mapping realizing the homotopic Frechét distance is a mapping of the unit square to this disk $\mathcal{D}$. This mapping specifies how to sweep over $\mathcal{D}$ in a geometrically "efficient" way (especially if the leash does not sweep over the same point more than once), so that the leash (i.e., the sweeping curve) is never too long [13]. As a concrete example, consider the two curves as enclosing several mountains between them on the surface – computing the homotopic Frechét distance corresponds to deciding which mountains to sweep first and in which order.

Furthermore, this mapping can be interpreted as surface parameterization [15, 27] and can thus be used in applications such as texture mapping [3, 25]. In the texture mapping problem, we wish to find a continuous and invertible mapping from the texture, usually a two-dimensional rectangular image, to the surface.

Another interesting interpretation is when f is a closed curve, and g is a point. Interpreting f as a rubber band in a 3d model, the homotopy height between f and g here is the minimum length the rubber band has to be so that it can be collapsed to a point (here, the rubber band stays on the surface as this is happening). In particular, a short closed curve with large homotopic height to any point in the surface is a "neck" in the 3d model.

To summarize, these measures seem to provide us with a fundamental understanding of the structure of the given surface/model.

Previous work.

Chambers *et al.* [7] gave a polynomial time algorithm to compute the homotopic Frechét distance between two polygonal curves on the Euclidean plane with polygonal obstacles. Chambers and Letscher [8, 9] introduced the notion of minimum homotopy height, and proved structural properties for the case of a pair of paths on the boundary of a topological disk. We remark that in general, it is not known whether the optimum homotopy has polynomially long description. In particular, it is not known whether the problem is in NP.

The problem of computing the (standard) Frechét distance between curves has been considered by Alt and Godau [2], who gave a polynomial time algorithm. Eiter and Mannila [14] studied the easier discrete version of this problem. Computing the Frechét distance between surfaces [17], appears to be a much more difficult task, and its complexity is poorly understood. The problem has been shown to be NP-hard by Godau [18], while the best algorithmic result is due to Alt and Buchin [1], who showed that it is upper semi-computable.

Efrat *et al.* [13] considered the Frechét distance inside a simple polygon as a way to facilitate sweeping it efficiently. They also used the Frechét distance with the underlining geodesic metric as a way to get a morphing between two curves. For recent work on the Frechét distance, see [11, 12, 20, 10] and references therein.

Our results.

In this paper, we consider the problems of computing the homotopic Frechét distance and the homotopy height between two simple polygonal curves that lie on the boundary of a triangulated topological disk $\mathcal{D}$ that is composed of n triangles.

We give a polynomial time $O(\log n)$-approximation algorithm for computing the homotopy height between f and g. Our algorithm to compute an approximate homotopy between f and g is via a simple, yet delicate divide and conquer algorithm.

We use the homotopy height algorithm as an ingredient for a $O(\log n)$-approximation algorithm for the homotopic Frechét distance problem. Intuitively, our algorithm for homotopic Frechét distance works as follows. We first approximately guess the optimum (i.e. $\mathsf{d}_{\mathcal{H}}(f, g)$). Using this guess, we classify parts of $\mathcal{D}$ as "obstacles", i.e. regions over which a short leash cannot pass. Let D' be the punctured disk obtained from $\mathcal{D}$ after removing these obstacles. The isotopy class of any leash is determined by the set of punctures that are on its left side. Observe that the leashes of the optimum solution belong to the same isotopy class. We describe a greedy algorithm to pick an isotopy class out of exponential number of choices, s.t. the homotopic Frechét distance constrained inside it is a constant factor of the homotopic Frechét distance in $\mathcal{D}$. Then, we use an extended version of our homotopy height algorithm to compute the homotopic Frechét distance.

The $O(\log n)$ factor shows up in the homotopic Frechét distance algorithm only because it uses the homotopy height as a subroutine. Thus, any constant factor approximation algorithm for the homotopy height problem implies a constant factor approximation algorithm for the homotopic Frechét distance.

As a warm-up exercise and in order to simplify the presentation we first consider the discrete version of the homotopy height problem in Section 2.1. This is how Chambers and Letscher formulated the problem. Later, in Section 2.2, we describe an algorithm to approximately find the shortest homotopy in continuous settings. In Section 3, we address the homotopic Frechét distance, both discrete and continuous. Further basic definitions are provided in the full version [19].

2. APPROXIMATING THE HEIGHT OF A HOMOTOPY

In this section we give an approximation algorithm for finding a homotopy of minimum height in a topological disc $\mathcal{D}$, whose boundary is defined by two walks L and R that share their end-points s and t. We start with the discrete case, i.e. when the disk is a triangulated edge-weighted planar graph and then generalize it to the continuous case. We then use this algorithm as a subroutine (in the next section) in our algorithm for the minimum homotopic Frechét distance problem.

2.1 The discrete case

To start, let us assume we are given an embedded planar graph G all of whose faces (except possibly the outer face) are triangles. Let $s, t \in \partial\mathsf{G}$ and L and R be the two non-crossing (s, t)-walks on $\partial\mathsf{G}$ in counter-clockwise and clockwise order, respectively. We use $\mathcal{D}$ to denote the topological disk enclosed by $\mathsf{L} \cup \mathsf{R}$. We refer to vertices of G (inside or on the boundary of $\mathcal{D}$) as vertices of $\mathcal{D}$. Our goal is to find a minimum height homotopy from L to R of non-crossing walks. Informally, the homotopy is defined by a sequence of walks, where every two consecutive walks differ by either a triangle, or an edge (being traversed twice). For a formal definition, see [19].

Lemma 2.1. *Let x and y be vertices of G that are at distance ρ. Then any homotopy between L and R has height at least ρ.*

Proof: Fix a homotopy of height δ. This homotopy contains an s-t walk ω that passes through x, and an s-t walk χ that passes through y. We have, by the triangle inequality, that $\rho \leq \|\omega[s, x]\| + \|\chi[s, y]\|$, and $\rho \leq \|\omega[x, t]\| + \|\chi[y, t]\|$. Therefore, $\rho \leq (\|\omega\| + \|\chi\|)/2 \leq \delta$, as required. ∎

Lemma 2.2. *Suppose d_1 is the maximum distance of a vertex of G from either of L or R, d_2 is the largest edge weight, and let $d = \max\{d_1, d_2\}$. Furthermore, let $\mathcal{D}$, L, R, and d be defined as above. Then any homotopy between L and R has height at least d.*

Proof: For every homotopy between L and R, and for every edge e, there exists a walk in the homotopy that passes through e. Therefore, the height of the homotopy is at least d_2. Moreover, the height is at least d_1 by Lemma 2.1. ∎

Here we present an algorithm which finds an (L, R)-homotopy of height at most $\|\mathsf{L}\| + \|\mathsf{R}\| + O(d \log n)$.

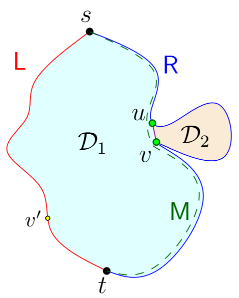
Lemma 2.3. *Let $\mathcal{D}$ be an edge-weighted triangulated topological disk with n faces such that its boundary is formed by two walks L and R that share endpoints s and t. Then, one can compute, in $O(n \log n)$ time, a homotopy from L to R of height at most $\|\mathsf{L}\| + \|\mathsf{R}\| + O(d \log n)$, where d is the largest among (a) the maximum distance of a vertex of $\mathcal{D}$ from either of L and R, and (b) the maximum edge weight.*

Proof: Let $f(\|\mathsf{L}\| + \|\mathsf{R}\|, d, n)$ denote the maximum height of such a homotopy. We will show that $f(u, d, n) = u + O(d \log n)$.

The base case $n = 0$ is easy. Indeed, if we have two edges (u, v) and (v, u) consecutive in R (or in L) we can retract these two edges. By repeating this we arrive at both L and R being identical, and we are done. The case $n = 1$ is handled in a similar fashion. After one face flip, the problem reduces to the case $n = 0$. As such, $f(\|\mathsf{L}\| + \|\mathsf{R}\|, d, 1) \leq \|\mathsf{L}\| + \|\mathsf{R}\| + d$.

For $n > 1$, compute for each vertex of G its shortest path to L, and consider the set of edges $\mathcal{E}$ used by all these shortest paths. Clearly, these shortest paths can be chosen so that $\mathsf{L} \cup \mathcal{E}$ form a tree. We consider each edge of R to be "thick" and have two sides (i.e., we think about these edges as being corridors with thickness). If $\mathcal{E}$ uses an edge of R then it uses the inner copy of this edge, while R uses the outer side. Similarly, we will consider each original vertex of R to be two vertices (one inside and other one on the boundary R). The set $\mathcal{E}$ would use only the inner vertices of R, while R would use only the outer vertices. To keep the graph triangulated we also arbitrarily triangulate inside each thick edge of R by adding corridor edges. Observe that, each corridor edge either connects two copies of a single vertex (thus has weight zero) or copies of two neighbors on R (and so has the same weight as the original edge).

Clearly, if we cut $\mathcal{D}$ along the edges of $\mathcal{E}$, what remains is a simple triangulated polygon (it might have "thin" corridors along the edges of R). One can find a diagonal uv such that each side of the diagonal contains at least $\lceil n/3 \rceil$ triangles of G (and at most $(2/3)n$). (Here, we count only the "real" triangles of G – we consider the faces of the thin corridors of the edges of R to have weight 0.) Observe that, because the faces inside corridors have weight zero, we can ensure that if uv is a corridor edge then u and v are copies of the same vertex. We use this property in the following case analysis.



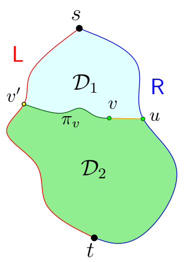
(A) Consider the case that u and v are both vertices of R. In this case, let $\mathsf{R}[u, v]$ be the portion of R in between u and v, and let $\mathcal{D}_2$ be the disk having $\mathsf{R}[u, v] \cup uv$ as its outer boundary. Let $\mathcal{D}_1$ be the disk $\mathcal{D} \setminus \mathcal{D}_2$. Let $\mathsf{M} = \mathsf{R}[s, u] \cup uv \cup \mathsf{R}[v, t]$.

Clearly, the distance of any vertex of $\mathcal{D}_1$ from L is at most d. By induction, there is a homotopy of height $f(\|\mathsf{L}\| + \|\mathsf{M}\|, d, (2/3)n)$ from L to M. Similarly, the distance of any vertex of $\mathcal{D}_2$ from uv is smaller than its distance to L. As such, by induction, there is a homotopy between uv and $\mathsf{R}[u, v]$ of height at most $f(\|\mathsf{R}[u, v]\| + d, d, (2/3)n)$. Clearly, we can extend this to a homotopy of M to R of height

$$\|\mathsf{R}[s, u]\| + f(\|\mathsf{R}[u, v]\| + d, d, (2/3)n) + \|\mathsf{R}[v, t]\| .$$

Putting these two homotopies together results in the desired homotopy from L to R.

(B) If v is a vertex of $\mathcal{E}$ and u is a vertex of R. As such, v is an inner vertex of R (that belongs to $\mathcal{E}$) and u is an outer vertex of R. Recall that we can assume that v and u are inner and outer copies of a same vertex of R. Let π_v be the shortest path in $\mathcal{D}$ from v to L, and let v' be its endpoint on L.



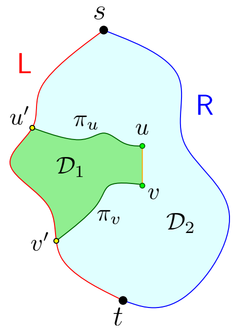
Consider the disk $\mathcal{D}_1$ having the "left" boundary $\mathsf{L}_1 = \mathsf{L}[s, v'] \cup \pi_v \cup vu$ and $\mathsf{R}_1 = \mathsf{R}[s, u]$ as its "right" boundary. This disk contains at most $(2/3)n$ triangles, and by induction, it has a homotopy of height $f(\|\mathsf{L}_1\| + \|\mathsf{R}_1\|, d, (2/3)n)$. To see why we can apply the recursion, observe that u and v are copies of the same vertex of R. That is, all shortest paths of vertices inside $\mathcal{D}_1$ to L are completely inside $\mathcal{D}_1$. Particularly, the distance of all vertices in $\mathcal{D}_1$ to L_1 are at most d.

Similarly, the topological disk $\mathcal{D}_2$ with the left boundary $\mathsf{L}_2 = uv \cup \pi_v \cup \mathsf{L}[v', t]$ and the right boundary $\mathsf{R}_2 = \mathsf{R}[u, t]$ has a homotopy of height $f(\|\mathsf{L}_2\| + \|\mathsf{R}_2\|, d, (2/3)n)$.

Starting with L, extending a tendril from v' to v, from v to u, and then applying the homotopy to first half of this walk (i.e., L_1) to move to R_1, and then the homotopy of $\mathcal{D}_2$ to the second part, results in a homotopy of L to R of height

$$\max \left(\begin{array}{c} \|\mathsf{L}\| + 2d, \\ f\Big(\|\mathsf{L}_1\| + \|\mathsf{R}_1\|, d, (2/3)n \Big) + \|\mathsf{L}_2\|, \\ \|\mathsf{R}_1\| + f\Big(\|\mathsf{L}_2\| + \|\mathsf{R}_2\|, d, (2/3)n \Big) \end{array} \right).$$

If the first number is the maximum, we are done. Otherwise, the above value is at most $f(\|\mathsf{L}\| + \|\mathsf{R}\| + 2d, d, 2/3n)$.



(C) Here we handle the case that u and v are both vertices of $\mathsf{L} \cup \mathcal{E}$. Then, as before, let u' and v' be the closest points on L to u and v, respectively. Now, let π_u (resp. π_v) be the shortest path from u (resp. v) to u' (resp. v').

Consider the disk $\mathcal{D}_1$ having $\mathsf{L}_1 = \mathsf{L}[u', v']$ as left boundary, and $\mathsf{R}_1 = \pi_u \cup uv \cup \pi_v$ as right boundary. This disk contains between $n/3$ and $2n/3$ triangles of the original surface. The distance of any vertex of $\mathcal{D}_1$ to L_1 (when restricted to $\mathcal{D}_1$) is at most d, and as such by induction, there is a homotopy from L_1 to R_1 of height $\alpha = f(\|\mathsf{L}_1\| + \|\mathsf{R}_1\|, d, (2/3)n) \leq f(\|\mathsf{L}[u', v']\| + 3d, d, (2/3)n)$. This yields a homotopy of height $\alpha_1 = \|\mathsf{L}[s, u']\| + \alpha + \|\mathsf{L}[v', t]\|$, from L to $\mathsf{L}_2 = \mathsf{L}[s, u'] \cup \pi_u \cup uv \cup \pi_v \cup \mathsf{L}[v, t]$. It is straight forward to check that $\alpha_1 \leq f(\|L\| + 3d, d, (2/3)n)$.

Next, let $\mathcal{D}_2$ be the disk with its left boundary being L_2 and its right boundary being $\mathsf{R}_2 = R$. Observe, that as before, the maximum distance of any vertex of $\mathcal{D}_2$ to L_2 is at most d. As before, by induction, there is a homotopy form L_2 to R_2 of height $\alpha_2 = f(\|\mathsf{L}_2\| + \|\mathsf{R}_2\|, d, (2/3)n)$. Since $\|\mathsf{L}_2\| \leq \|\mathsf{L}\| + 3d$, we have $\alpha_2 \leq f(\|\mathsf{L}\| + \|\mathsf{R}\| + 3d, d, (2/3)n)$.

In all cases the length of the homotopy is at most

$$f(\|\mathsf{L}\| + \|\mathsf{R}\| + 3d, d, (2/3)n)$$

Now, it is easy to verify that the solution to the recursion $f(u, d, n)$ that complies with all the above inequalities is $f(u, d, n) = u + O(d \log n)$, as desired.

We can compute the shortest path tree in linear time using the algorithm of Henzinger *et al.* [21]. The separating edge can also be found in linear time using DFS. So, the running time for a graph with n faces is $T(n) = T(n_1) + T(n_2) + O(n)$, where $n_1 + n_2 = n$ and $n_1, n_2 \leq 2/3n$. It follows that $T(n) = O(n \log n)$. ■

Remark 2.4. In the algorithm of Lemma 2.3, it is not necessary that we have the shortest paths from L to all the vertices of $\mathcal{D}$. Instead, it is sufficient if we have a tree structure that provides paths from any vertex of $\mathcal{D}$ to L of distance at most d in this tree. We will use this property in the continuous case, where recomputing the shortest path tree is relatively expensive.

2.2 The continuous case

In this section we extend the arguments to the continuous case. Here we are given a piecewise linear triangulated topological disk, $\mathcal{D}$, with n triangles. The boundary of $\mathcal{D}$ is composed of two paths L and R with shared endpoints s and t. Observe that the distance of any point x in $\mathcal{D}$ from L and R is not longer than the homotopy height as there is a (s, t)-path that contains x. Here, we build a homotopy of height $\|\mathsf{L}\| + \|\mathsf{R}\| + O(d \log n)$, where d is the maximum distance of any point in $\mathcal{D}$ from either L or R.

We use the following observations (see [19] for details):

(A) The shortest path from a set of $O(n)$ edges to the whole surface can be computed in $O(n^3 \log n)$ time.

(B) A shortest path (i.e., a geodesic) intersects a face along a segment and it locally looks like a segment if the adjacent faces are rotated to be coplanar.

2.2.1 *Homotopy height if edges are short*

Here, we assume that the longest edge in $\mathcal{D}$ has length at most $2d$, where d is the maximum distance for any point of $\mathcal{D}$ from either L or R.

As in the discrete case, let $\mathcal{E}$ be the union of all the shortest paths from the vertices of $\mathcal{D}$ to L (as before, we treat the edges and vertices of R as having infinitesimal thickness). For a vertex v of $\mathcal{D}$, its shortest path π_v is a polygonal path that crosses between faces (usually) in the middle of edges (it might also go to a vertex, merge with some other shortest paths and then follow a common shortest path back to L). In particular, each such shortest path might intersect a face of $\mathcal{D}$ along a single segment. As such, the polygon resulting from cutting $\mathcal{D}$ along $\mathcal{E}$, call it P, is a polygon that has complexity $O(n^2)$. A face of P is a hexagon, a pentagon, a quadrilateral, or a triangle. However, it has at most 3 edges that are portions of the edges of $\mathcal{D}$. We say the degree of a face is i if it has i edges that are portions of the edges of $\mathcal{D}$. Observe that, each triangle of $\mathcal{D}$ is now decomposed into a set of faces. Obviously, each triangle of $\mathcal{D}$ contains at most one face of degree 3. Overall, there are $O(n)$ faces of degree 3 in P.

Now consider C^*, the dual of the graph that is inside the polygon (ignore the edges on the boundary). More precisely, C^* has a vertex corresponding to each face inside the polygon P. Two vertices of C^* are adjacent if and only if their corresponding faces share a portion of an edge of $\mathcal{D}$ (this shared edge is a diagonal of the polygon resulting from the cutting). Since the maximum degree of C^* is 3, there is an edge that is a good separator. We use this edge in a similar fashion to the proof of Lemma 2.3, except that in the recursion we avoid recomputing the shortest paths (i.e., we use the old shortest paths and distances computed in the original disk), see Remark 2.4. So, we compute the shortest paths once in the beginning in $O(n^3 \log n)$ time. Then, in each step we can find the separator in $O(n^2)$ time. Namely, the total time spent on computing the separators is $T(n) = T(n_1) + T(n_2) + O(n^2)$, where $n_1 + n_2 = O(n^2)$ and $n_1, n_2 \le (2/3)(n_1 + n_2)$; that is, $T(n) = O(n^2 \log n)$. As such, the total running time is dominated by the computation of the shortest paths.

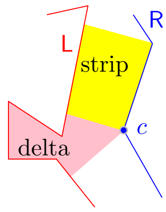
The proof of Lemma 2.3 then goes through literally in this case. Since all the edges have length at most $2d$, by assumption, we get the following.

Lemma 2.5. *Let $\mathcal{D}$ be a topological disk with n faces where every face is a triangle (here, the distance between any two points on the triangle is their Euclidean distance). Furthermore, the boundary of $\mathcal{D}$ is formed by two walks L and R (that share two endpoints). Let d be the maximum distance of any point of $\mathcal{D}$ from either L or R. Furthermore, assume that all edges of $\mathcal{D}$ have length at most $2d$. Then, one can compute a continuous homotopy from L to R of height $\le \|\mathsf{L}\| + \|\mathsf{R}\| + O(d \log n)$ in $O(n^3 \log n)$ time.*

2.2.2 *Homotopy height if there are long edges*

Algorithm.

For any two points in $\mathcal{D}$ consider a shortest path π connecting them. The signature of π is the ordered sequence of edges (crossed or used) and vertices used by π, see [19]. For a point $\mathsf{p} \in \mathsf{R}$, let $\mathsf{s}_\mathsf{L}(\mathsf{p})$ denote the signature of the shortest path from p to L. The signature $\mathsf{s}_\mathsf{L}(\mathsf{p})$ is well defined in R except for a finite set of *medial* points, where there are two (or more) distinct shortest paths from L to p. In particular, let Π_R be the set of all shortest paths from any medial point on R to L. Observe that, the medial points are the only points that the signature of the shortest path from R to L changes in any non-degenerate triangulation.



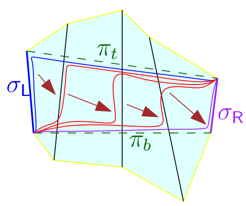
Cutting $\mathcal{D}$ along the paths of Π_R breaks $\mathcal{D}$ into ***corridors***. If the intersection of a corridor with R is a point (resp. segment) then it is a ***delta*** (resp. ***strip***). In a strip C, all the shortest paths to L from the points in the interior of the segment $C \cap \mathsf{R}$ have the same signature. Intuitively, strips have a natural way to morph from one side to the other. We further break each delta into chunks and pockets.

So, consider a delta C with an apex c (i.e., the point of R on the boundary of C). For a point $x \in \mathsf{L} \cap C$, we define its signature (in relation to C), to be the signature of the shortest path from x to c (restricted to lie inside C). Again, we partition $\mathsf{L} \cap C$ into maximum intervals that have the same signature. If a newly created region has a single intersection point with both L and R, then it is a ***pocket***, otherwise, it is a ***chunk***.

Consider such a chunk C. Its intersection with L is a segment, and its intersection with R is a point (i.e., the apex c of the delta). Observe that the distance of any point of $\mathsf{L} \cap C$ to c is at most $2d$, and this provides a natural way to morph $\mathsf{L} \cap C$ to c. A pocket, on the other hand, is a topological disk that its intersections with L and R are both single points, and the two boundary paths between these intersections are of length $\le 2d$. Pockets are handled by using the recursive scheme developed for the discrete case.

Applying the above partition scheme to all the deltas results in breaking $\mathcal{D}$ into strips, chunks and pockets. Next, order the resulting regions according to their order along L, and transform each one of them at time, such that starting with L we end up with R.

(A) **Morphing a chunk/strip** S: Let $\sigma_\mathsf{L} = \mathsf{L} \cap S$ and $\sigma_\mathsf{R} = \mathsf{R} \cap S$. There is a natural homotopy from $\pi_t \cup \sigma_\mathsf{L}$ to $\sigma_\mathsf{R} \cup \pi_b$.



The strip/chunk S has no vertex of $\mathcal{D}$ in its interior, and as such it is formed by taking planar quadrilaterals and gluing them together along common edges. Observe that by the triangle inequality, all such edges of any of these quadrilaterals are of length $\le \max(\|\sigma_\mathsf{L}\|, \|\sigma_\mathsf{R}\|) + 4d$. It is now easy to check that we can collapse each such quadrilateral in turn to get the required homotopy. Since each of π_t and π_b is composed of two shortest paths, there is a linear number of such quadrilaterals, and each collapse can be done in constant time. See figure for an example.

(B) **Morphing a pocket**: A pocket has perimeter $4d$, and there is a point on its boundary, such that the distance of any point in it to this base point is at most $2d$. By the the triangle inequality, we have that for a topological disk $\mathcal{D}$, such that all the points of $\mathcal{D}$ are in distance at most $2d$ from some point c. Then the longest edge in $\mathcal{D}$ has length at most $4d$. As such, all the edges inside a pocket can not be longer than $4d$. We can now apply Lemma 2.5 to such a pocket. This results in the desired homotopy.

Analysis. The shortest paths from R to L can be computed in $O(n^3 \log n)$ time. The shortest paths inside a delta to its apex can be computed in $O(n^2 \log n)$. Since there is a linear number of deltas, the total running time for building the strips is $O(n^3 \log n)$.

Lemma 2.6. *The number of paths in* Π_{R} *is* $O\Big(|\mathsf{V}(\mathcal{D})|\Big)$, *where* $\mathsf{V}(\mathcal{D})$ *is the set of vertices of* $\mathcal{D}$.

Proof: Let $\{\zeta_1, \zeta_2, \ldots, \zeta_k\}$ be the paths in Π_{R} sorted by the order of their endpoints along R. Observe that these paths are geodesics and such one can assume that they are interior disjoint. Now, if $l_i \in L$ and $r_i \in R$ are the endpoints of ζ_i, for $i = 1, \ldots, k$, then these endpoints are sorted along their respective curves. In particular, let $\mathcal{D}_i$ be the disk having $\mathsf{L}[s, l_i] \cup \zeta_{i+1} \cup \mathsf{R}[s, r_i]$ for boundary. We have that $\mathcal{D}_1 \subseteq \mathcal{D}_2 \subseteq \cdots \subseteq \mathcal{D}_k$. The signatures of ζ_i and ζ_{i+2} must be different as otherwise they would be consecutive. Furthermore, because of the inclusion property, if an edge or a vertex of $\mathcal{D}$ intersects ζ_i but does not intersect ζ_{i+1} then, it can not intersect any later path. As such, every other path in Π_{R} can be charged to vertices or edges that are added or removed from the signature of the respective path. Since an edge or a vertex can be added at most once, and deleted at most once, this implies the desired bound on the number of paths. ■

Arguing as in Lemma 2.6, we have that the total number of parts (i.e., chunks, corridors and pockets) generated by the above decomposition is $O(|\mathsf{V}(\mathcal{D})|)$.

Lemma 2.7. *Consider a strip or a chunk S generated by the above partition of $\mathcal{D}$. Let $\sigma_{\mathsf{L}} = \mathsf{L} \cap S$ and $\sigma_{\mathsf{R}} = \mathsf{R} \cap S$. Let π_t and π_b be the top and bottom paths forming the two sides of S that do not lie on* R *or* L.

(A) We have $\|\pi_b\| \leq 2d$ and $\|\pi_t\| \leq 2d$.

(B) If $\|\sigma_{\mathsf{L}}\| > 0$ or $\|\sigma_{\mathsf{R}}\| > 0$ then there is no vertex of $\mathcal{D}$ in the interior of S.

(C) If $\|\sigma_{\mathsf{L}}\| > 0$ or $\|\sigma_{\mathsf{R}}\| > 0$ then there is a homotopy from $\pi_t \cup \sigma_{\mathsf{L}}$ to $\sigma_{\mathsf{R}} \cup \pi_b$ of height $\max(\|\sigma_{\mathsf{L}}\|, \|\sigma_{\mathsf{R}}\|) + 4d$. *This homotopy can be computed in linear time.*

Proof: (A) If the strip was generated by the first stage of partitioning then the claim is immediate.

Otherwise, consider a delta C with an apex c. For any point $x \in \mathsf{L} \cap C$ we claim that there is a path of length $\leq 2d$ to c. Indeed, consider the shortest path π_x from x to R in $\mathcal{D}$. If this path goes to c the claim holds immediately. Otherwise, the shortest path (that has length at most d) must cross either the top or bottom shortest path forming the boundary of C that are emanating from c. We can now modify π_x, so that after its intersection point with this shortest path, it follows it back to c. Clearly, the resulting shortest path has length at most $2d$ and lies inside the resulting chunk.

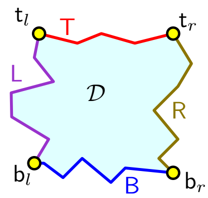
(B) Indeed, the boundary paths π_t and π_b have the same signature (formally, they are the limit paths of same signature). Since $\mathcal{D}$ is non-degenerate, if there was any vertex in the middle, then the path on one side of the vertex, and the path on the other side of the vertex can not possibly have the same signature.

(C) Immediate from the algorithm description. ■

We thus get the following result.

Theorem 2.8. *Suppose that we are given a piecewise linear triangulated surface with the topology of a disk, such that its boundary is formed by two walks* L *and* R. *Then, there is a continuous homotopy from* L *to* R *of height* $\leq \|\mathsf{L}\| + \|\mathsf{R}\| + O(d \log n)$. *This homotopy can be computed in* $O\big(n^3 \log n\big)$ *time.*

3. COMPUTING THE HOMOTOPIC FRECHÉT DISTANCE



In this section, fix $\mathcal{D}$ to be a triangulated topological disk with n faces. Let the boundary of $\mathcal{D}$ be composed of T, R, B and L, four internally disjoint walks appearing in clockwise order along the boundary. Also, let $\mathsf{t}_l = \mathsf{L} \cap \mathsf{T}$, $\mathsf{b}_l = \mathsf{L} \cap \mathsf{B}$, $\mathsf{t}_r = \mathsf{R} \cap \mathsf{T}$ and $\mathsf{b}_r = \mathsf{R} \cap \mathsf{B}$.[①] See figure on the right.

3.1 Approximating the Regular Frechét Distance

Let $\mathsf{d}_{\mathcal{F}}(\mathsf{T}, \mathsf{B})$ (resp. $\mathsf{d}_{\mathcal{H}}(\mathsf{T}, \mathsf{B})$) be the regular (resp. homotopic) Frechét distance between T and B (when restricted to $\mathcal{D}$). Clearly, $\mathsf{d}_{\mathcal{F}}(\mathsf{T}, \mathsf{B}) \leq \mathsf{d}_{\mathcal{H}}(\mathsf{T}, \mathsf{B})$. The following lemma implies that the Frechét distance can be approximated within a constant factor.

Lemma 3.1. *Let $\mathcal{D}$, n,* T *and* B *be as above. Then, for the continuous case, one can compute, in $O\big(n^3 \log n\big)$ time, reparametrizations of* T *and* B *of width $\leq 2\delta$, where $\delta =$* $\mathsf{d}_{\mathcal{F}}(\mathsf{T}, \mathsf{B})$.

Proof: Let Π be the set of shortest paths from all points of T to the curve B. As in the proof of Lemma 2.6, let Π_{T} be the set of all shortest paths from medial points on T to B. Arguing as in Lemma 2.6, we have that the set Π_{T} is composed of a linear number of paths. The paths in Π_{T} do not cross and so partition $\mathcal{D}$ into a set of regions. Each region is bounded by a portion of T, a portion of B and two paths in Π_{T}. A region is a *delta* if the two paths of Π_{T} in its boundary share a single endpoint (on T), it is a ***pocket*** if they share two endpoints (one on T and one on B), and it is *strip* if they share no endpoints.

Obviously, the (endpoints of the) paths in Π covers all of T. The paths in Π also cover all of B except for the bases of deltas. Now, for each delta we compute the set of all shortest paths from the vertices of its base to its apex inside the delta. Let Π_{B} be the set of all such paths in all deltas.

[①]We use the same notation to argue about both the discrete problem and the continuous problem.

Clearly, the union of Π_B and Π_T is a set of non-crossing paths whose endpoints cover all the vertices of T and B.

The shortest path from any point of T to B is at most δ. So, all paths in Π have length at most δ. Similarly, the shortest path from a point of B to T is at most δ. Now, consider a delta C with apex c. Let b be a point on the base of C (and so on B). The shortest path π_b from b to T has length at most δ. Let x be the first point that π_b intersects a boundary path of C, π_C. Now, $\pi_b[b, x] \cdot \pi_C[x, c]$ has length at most 2δ and it is inside C. We conclude that all paths in Π_B have length at most 2δ.

The paths in $\Pi_B \cup \Pi_T$ decompose $\mathcal{D}$ into strips and corridors. The left and right portions of a strip is of length at most 2δ, and its top and bottom sides have as such Frechét distance at most 2δ from each other. Similarly, the leash can jump over a pocket from the left leash to the right leash. Doing this to all corridors and pockets, results in reparametrizations of L and R such that their maximum length of a leash for these reparametrizations are at most 2δ. This implies that the Frechét distance is at most 2δ, and we have an explicit reparametrization that realizes this distance.

As for the running time, in $O(n^3 \log n)$ time, one can compute all shortest paths from T to the whole surface. Then, one can, in $O(n^2 \log n)$ time, compute the shortest paths inside each of the linear number of deltas. It follows that the total running time is $O(n^3 \log n)$. ■

Lemma 3.2. *Let $\mathcal{D}$ be a triangulated topological disk with n faces, and T and B be two internally disjoint walks on the boundary of $\mathcal{D}$. Then, one can compute, in $O(n)$ time, reparametrizations of T and B that approximate the discrete Frechét distance between T and B. The computed reparametrizations have width $\leq 3\delta$, where δ is the Frechét distance between T and B.*

Proof: First, compute the set of shortest paths, $\Pi_T = \{\pi_1, \pi_2, \cdots, \pi_k\}$, from vertices of T to the path B. Now, let π_i and π_{i+1} be two consecutive paths, that is the endpoints of π_i and π_{i+1}, a_i and a_{i+1}, are adjacent vertices on T. For all $1 \leq i < k$, we add the paths $\pi_i^+ = (a_i, a_{i+1}) \cdot \pi_{i+1}$ to the set Π_T to obtain Π_T^+. Observe that each path in Π_T^+ has length at most 2δ; it is composed of zero or one edge of T and a shortest path from a vertex of T to B. Further, Π_T^+ partitions the graph into regions, similar to the continuous case. Now for each vertex of B that is not an endpoint of a path in Π_T^+, we compute the shortest path inside its region to T. Because the region is bounded by paths of length at most 2δ, the length of such a shortest path is at most 3δ. If Π_B is the set of all such shortest paths, then $\Pi_T^+ \cup \Pi_B$ is a leash sequence of height at most 3δ.

We use the algorithm of Henzinger *et al.* [21] to compute the shortest paths from T in linear time. Since all regions are disjoint, we can compute all the shortest paths inside different regions in $O(n)$ time, as well. ■

3.2 Homotopic Frechét distance if there are no mountains

The following lemma implies that if all the vertices in $\mathcal{D}$ are not too far from the two curves, then one can transform the Frechét distance into the continuous variant (i.e., without jumps in the leash).

Lemma 3.3. *Let $\mathcal{D}$ be a triangulated topological disk with n faces, and T and B be two internally disjoint walks on the boundary of $\mathcal{D}$. Further, assume for all $\mathsf{p} \in \mathcal{D}$, p's distance to both T and B is $\leq x$. Then, one can compute reparametrizations of T and B of width $O(x \log n)$. The running time is $O(n^4 \log n)$ for the continuous case. For the discrete case the running time is $O(n \log n)$.*

In particular, if $x = O(\mathsf{d}_{\mathcal{H}}(\mathsf{T}, \mathsf{B}))$ then this is an $O(\log n)$-approximation to the optimal homotopic Frechét distance.

Proof: Using the algorithm of Lemma 3.1 (or Lemma 3.2 for the discrete case) we compute (not necessarily continuous) reparametrization of T and B of (regular) Frechét distance $\leq \delta$, where $\delta = O(x)$.

Note that the leash movement $s(\cdot)$ associated with this reparametrizations, is not required to deform continuously in s. For a given time $t \in [0, 1]$, let $s^-(t) = \lim_{t' \to t^-} s(t')$ and $s^+(t) = \lim_{t' \to t^+} s(t')$. By definition, s is discontinuous at t if and only if $s^-(t) \neq s^+(t)$. In the discrete case, let $s^-(t) = s(t)$ and $s^+(t) = s(t + 1)$. In this case, we say s is discontinuous at t if $s^-(t)$ and $s^+(t)$ are more than one flip operation apart. In both cases, we also say that s has a *gap* at t.

Observe that a gap at time t can be filled by attaching an $(s^-(t), s^+(t))$-homotopy to s at time t. Observe that all the vertices inside the disc with boundary $s^-(t) \cup s^+(t)$ have distance $O(x)$ to T and B and so to $s^-(t)$ and $s^+(t)$, so Lemma 2.8 and Lemma 2.3 imply that an $(s^-(t), s^+(t))$-homotopy with height $O(x \log n)$ can be computed.

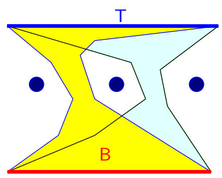
Suppose that s has a gap at t. In the discrete case, the disc with boundary $s^-(t) \cup s^+(t)$ contains at least one face. In the continuous case, that disc contains at least one vertex. Thus, in both cases, there are $O(n)$ gaps.

We start with s and fill in all linear number of gaps to obtain a continuous leash sequence of height $O(x \log n)$.

In the discrete case, the initial leash sequence can be computed in $O(n)$ time. Since the gaps are disjoint, they can be filled in overall $O(n \log n)$.

In the continuous case, the initial leash sequence can be computed in $O(n^3 \log n)$ time. Each of the $O(n)$ number of gaps can be filled in $O(n^3 \log n)$ time. ■

3.3 A Decision Procedure for the Homotopic Frechét distance in the presence of mountains



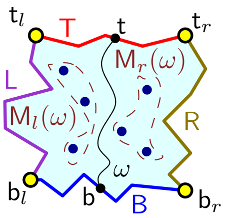
For a parameter $\tau \geq 0$, a vertex $v \in \mathsf{V}(\mathcal{D})$ is ***τ-tall*** if and only if its distance to T or B is larger than τ (intuitively τ is a guess for the value of $\mathsf{d}_{\mathcal{H}}(\mathsf{T}, \mathsf{B})$). Here, we consider the case where there are τ-tall vertices. Intuitively, one can think about tall vertices as insurmountable mountains. Thus, to find a good homotopy between T and B, we have to choose which "valleys" to use (i.e., what homotopy class the solution we compute belongs to if we think about tall vertices as punctures in the disk). As a concrete example, consider the figure on the right, where there are three tall vertices, and two possible solutions are being shown. As suggested by the figure, we have to make a combinatorial decision: Which tall vertices are going to be to the "left" of computed homotopy, and which tall vertices are going to be on the other side.

In the discrete case, we subdivide each edge in the beginning so that if an edge has length $> 2\tau$, then the vertex inserted in the middle of it is τ-tall. Observe that, if $\tau \geq \mathsf{d}_{\mathcal{H}}(\mathsf{T}, \mathsf{B})$ then no leash of the optimum homotopic motion can afford to contain a τ-tall vertex. We use M^τ to denote the set of all τ-tall vertices in $\mathsf{V}(\mathcal{D})$.

Now, let ω and ω' be two walks connecting points on T and B. We say that ω and ω' are ***isotopic*** in $\mathcal{D} \setminus \mathsf{M}^\tau$ if and only if they are homotopic in $\mathcal{D} \setminus \mathsf{M}^\tau$ after contracting T and B. Consequently, the disc with boundary $\mathsf{T} \cup \mathsf{B} \cup \omega \cup \omega'$ contains no tall vertices if ω and ω' are isotopic.

Definition 3.4. *Given a subset $X \subseteq \mathsf{M}^\tau$, consider a path ζ from T to B, such that X is contained in one side of $\mathcal{D} \setminus \zeta$ (i.e., cutting $\mathcal{D}$ along ζ breaks it into two connected components), and $\mathsf{M}^\tau \setminus X$ is contained in the other side. The set of all such paths is the* **isotopy class** *of X. For each such X, the isotopy class of X is a τ-isotopy class.*

Let $\pi_{\mathrm{L},\mathsf{h}}$ (resp. $\pi_{\mathrm{R},\mathsf{h}}$) be the **left geodesic** *(resp.* **right geodesic***) of an isotopy class h; that is, $\pi_{\mathrm{L},\mathsf{h}}$ denotes the shortest path in h from t_l to b_l (resp. from t_r to b_r).*



Let ω be any walk in h from $\mathsf{b} \in \mathsf{B}$ to $\mathsf{t} \in \mathsf{T}$. We define the ***left tall set*** of h, denote $\mathsf{M}_l(\mathsf{h}) = \mathsf{M}_l(\omega)$ to be the set of all τ-tall vertices to the left of ω; inside the disc with boundary $\mathsf{L} \cup \mathsf{T}[\mathsf{t}_l, \mathsf{t}] \cup \omega \cup \mathsf{B}[\mathsf{b}_l, \mathsf{b}]$, where L is the "left" portion of the boundary of $\mathcal{D}$, having endpoints t_l and b_l. We similarly define the ***right tall set*** of h, $\mathsf{M}_r(\mathsf{h}) = \mathsf{M}_r(\omega)$, to be the set of all τ-tall vertices to the right of ω. See figure on the right. Note that the sets $\mathsf{M}_l(\mathsf{h})$ and $\mathsf{M}_r(\mathsf{h})$ do not depend on the particular choice of ω, since all paths in h are in the same isotopy class.

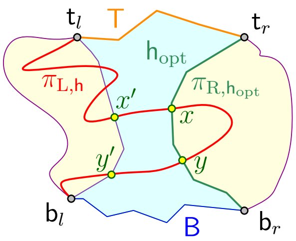
We say that h is ***τ-extendable*** from the left if and only if $\|\pi_{\mathrm{L},\mathsf{h}}\| \leq \tau$ and there is an isotopy class h', such that $\|\pi_{\mathrm{L},\mathsf{h}'}\| \leq \tau$ and $\mathsf{M}_l(\mathsf{h}) \subset \mathsf{M}_l(\mathsf{h}')$. In particular, h is ***τ-saturated*** if it is not τ-extendable and $\|\pi_{\mathrm{L},\mathsf{h}}\| \leq \tau$.

3.3.1 On the left and right geodesics

Lemma 3.5. *Let h be a τ-saturated isotopy class, where $\tau \geq \mathsf{d}_{\mathcal{H}}(\mathsf{T}, \mathsf{B})$. Then, $\|\pi_{\mathrm{R},\mathsf{h}}\| \leq 4\tau$.*

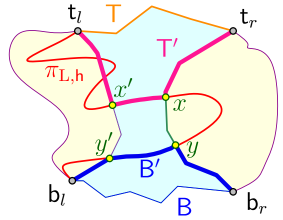
Proof: Let $\mathsf{h}_{\mathrm{opt}}$ be the isotopy class of the leashes in the optimum solution. Of course, no leash in the optimum solution contains a τ-tall vertex, and therefore, all leashes in the optimal solution are isotopic.

Since h is saturated, the set $\mathsf{M}_l(\mathsf{h})$ is not a proper subset of $\mathsf{M}_l(\mathsf{h}_{\mathrm{opt}})$. If $\mathsf{M}_l(\mathsf{h}) = \mathsf{M}_l(\mathsf{h}_{\mathrm{opt}})$ then $\mathsf{h} = \mathsf{h}_{\mathrm{opt}}$, and in particular $\|\pi_{\mathrm{R},\mathsf{h}}\| = \|\pi_{\mathrm{R},\mathsf{h}_{\mathrm{opt}}}\| \leq \tau$.



Otherwise, the set $\mathsf{M}_l(\mathsf{h}) \cap \mathsf{M}_r(\mathsf{h}_{\mathrm{opt}})$ is not empty. It follows that $\pi_{\mathrm{L},\mathsf{h}}$ crosses $\pi_{\mathrm{R},\mathsf{h}_{\mathrm{opt}}}$.

Let x be the first intersection point between $\pi_{\mathrm{L},\mathsf{h}}$ and $\pi_{\mathrm{R},\mathsf{h}_{\mathrm{opt}}}$, as one traverses $\pi_{\mathrm{L},\mathsf{h}}$ from t_l to b_l. Let x' be the last intersection point of $\pi_{\mathrm{L},\mathsf{h}}[\mathsf{t}_l, x]$ with $\pi_{\mathrm{L},\mathsf{h}_{\mathrm{opt}}}$. Similarly, y is the last intersection point between $\pi_{\mathrm{L},\mathsf{h}}$ and $\pi_{\mathrm{R},\mathsf{h}_{\mathrm{opt}}}$, and y' is the first intersection of $\pi_{\mathrm{L},\mathsf{h}}[y, \mathsf{b}_l]$ and $\pi_{\mathrm{L},\mathsf{h}_{\mathrm{opt}}}$. Observe that the interiors of $\pi_{\mathrm{L},\mathsf{h}}[x', x]$ and $\pi_{\mathrm{L},\mathsf{h}}[y, y']$ does not intersect the curves $\pi_{\mathrm{L},\mathsf{h}_{\mathrm{opt}}}$ and $\pi_{\mathrm{R},\mathsf{h}_{\mathrm{opt}}}$.



As the curves $\pi_{\mathrm{L},\mathsf{h}}$ and $\pi_{\mathrm{R},\mathsf{h}}$ are isotopic (by definition), the disk with the boundary $\mathsf{T} \cdot \pi_{\mathrm{L},\mathsf{h}} \cdot \mathsf{B} \cdot \pi_{\mathrm{R},\mathsf{h}}$ does not contain any tall vertex, and $\mathsf{T} \cdot \pi_{\mathrm{L},\mathsf{h}} \cdot \mathsf{B}$ is homotopic to $\pi_{\mathrm{R},\mathsf{h}}$.

Consider the following walk $\mathsf{T}' = \pi_{\mathrm{R},\mathsf{h}_{\mathrm{opt}}}[\mathsf{t}_r, x] \cdot \pi_{\mathrm{L},\mathsf{h}}[x, x'] \cdot \pi_{\mathrm{L},\mathsf{h}_{\mathrm{opt}}}[x', \mathsf{t}_l]$. The walk T' is homotopic to T. Similarly, $\mathsf{B}' = \pi_{\mathrm{L},\mathsf{h}_{\mathrm{opt}}}[\mathsf{b}_l, y'] \cdot \pi_{\mathrm{L},\mathsf{h}}[y', y] \cdot \pi_{\mathrm{R},\mathsf{h}_{\mathrm{opt}}}[y, \mathsf{b}_r]$ is homotopic to B. It follows that $\pi_{\mathrm{R},\mathsf{h}}$ is homotopic to $\mathsf{T}' \cdot \pi_{\mathrm{L},\mathsf{h}} \cdot \mathsf{B}'$. As $\pi_{\mathrm{R},\mathsf{h}}$ is the shortest path in its homotopy class with these endpoints, it follows that

$$\|\pi_{\mathrm{R},\mathsf{h}}\| \leq \|\mathsf{T}' \cdot \pi_{\mathrm{L},\mathsf{h}} \cdot \mathsf{B}'\| \leq \|\pi_{\mathrm{L},\mathsf{h}}\| + \left(\|\pi_{\mathrm{L},\mathsf{h}_{\mathrm{opt}}}\| + \|\pi_{\mathrm{L},\mathsf{h}}\| + \|\pi_{\mathrm{R},\mathsf{h}_{\mathrm{opt}}}\|\right) \leq 4\tau,$$

as T' and B' are disjoint, and $\mathsf{T}' \cup \mathsf{B}' \subseteq \pi_{\mathrm{R},\mathsf{h}_{\mathrm{opt}}} \cup \pi_{\mathrm{L},\mathsf{h}_{\mathrm{opt}}} \cup \pi_{\mathrm{L},\mathsf{h}}$. ■

Lemma 3.6. *Let h be a τ-isotopy class, such that we have $\max(\|\pi_{\mathrm{L},\mathsf{h}}\|, \|\pi_{\mathrm{R},\mathsf{h}}\|) \leq x$, where $x \geq \tau \geq \mathsf{d}_{\mathcal{H}}(\mathsf{T}, \mathsf{B})$. Let $\mathcal{D}'$ be the disk with boundary $\mathsf{T} \cdot \pi_{\mathrm{R},\mathsf{h}} \cdot \mathsf{B} \cdot \pi_{\mathrm{L},\mathsf{h}}$. Then, all the points inside $\mathcal{D}'$ are closer than $O(x)$ to both T and B in $\mathcal{D}'$.*

Proof: We first consider the continuous case.

By the definition of τ-isotopy, the disk $\mathcal{D}'$ has no τ-tall vertices. Furthermore, by the definition of x, we have that the distance of any point on T to B, restricted to paths in $\mathcal{D}'$ is at most δ_1, where $\delta_1 = x + \mathsf{d}_{\mathcal{F}}(\mathsf{T}, \mathsf{B}) \leq 2x$. Indeed, the shortest path from any point on T to B in $\mathcal{D}$, either stays inside $\mathcal{D}'$, or alternatively intersects either $\pi_{\mathrm{L},\mathsf{h}}$ or $\pi_{\mathrm{R},\mathsf{h}}$.

We can now deploy the decomposition of $\mathcal{D}'$ into strips and pockets, as done in Section 2.2.2. Every strip is being swept by a leash of length $\leq \delta_2 = 2\delta_1 \leq 4x$ (the factor two is because a strip might rise out of a delta), and as such the claim trivially holds for points inside a strip.

Every pocket P has perimeter of length at most $\|\partial\mathsf{P}\| \leq \delta_3 = 2\delta_2 = 8x$ (the perimeter also contains two points of T and B and they are in distance at most δ_2 from each other in either direction along the perimeter).

So consider such a pocket P. Since $\mathcal{D}'$ contains no τ-tall vertices, P does not contain any tall vertex. Let e be an edge in P (or a subedge if it intersects the boundary of P). The two endpoints of e are in P, and such an endpoint is either a (not tall) vertex or it is contained in $\partial\mathsf{P}$. In either case, these endpoints are in distance at most x from $\partial\mathsf{P}$, and as such they are in distance at most $\delta_4 = 2x + \|\partial\mathsf{P}\|/2 = 2x + \delta_2 \leq 6x$ from each other (inside P). We conclude that $\|\mathsf{e}\| \leq \delta_4$, and as such, any point in e is in distance at most $\delta_6 = \|\mathsf{e}\|/2 + x + \delta_2 \leq 3x + x + 8x \leq 12x$ from T and B.

Now, consider any point p in P, and consider the face F that contains it. Since the surface is triangulated, F is a triangle. Clipping $\triangle$ to P results in a planar region F' that has perimeter $\leq \delta_5 = 3\delta_4 + \|\partial\mathsf{P}\| \leq 3 \cdot 6x + \delta_3 \leq (18+8)x \leq 26x$ (note, that an edge might be fragmented into several subedges, but the furthest two points along a single edge is at most δ_4 using the same argumentation as above). As such, the furthest a point of P can be from an edge of P is at most $\delta_7 = \delta_5/2\pi \leq 5x$. As such, the maximum distance of a point of P from either T or B (inside $\mathcal{D}'$) is at most $\delta_6 + \delta_7 \leq 12x + 5x = 17x$.

The discrete case is easier. Any edge of length $\geq 2\tau$ was split, by introducing a middle vertex, which must be τ-tall.

As such, there are no long edges using by any curves in the same isotopic class. A simplified version of the above proof now implies the claim. ■

3.3.2 The decision algorithm

The proof of the following three lemmas, can be found in the full version of the paper [19].

Lemma 3.7. *Let $\mathcal{D}, n, \mathsf{T}, \mathsf{L}, \mathsf{B}, \mathsf{R}, \mathsf{t}_l, \mathsf{b}_l, \tau$ be as above, and let $X \subseteq \mathsf{V}(\mathcal{D})$ be the set of τ-tall vertices. Consider the shortest path ζ (between t_l and b_l) that belongs to any homotopy class h such that $X \subseteq \mathsf{M}_l(\mathsf{h})$. Then, path ζ can be computed in $O(n^4 \log n)$ (resp. $O(n^2 \log n)$) time in the continuous (resp. discrete) case.*

Lemma 3.8. *Let $\mathcal{D}$ be a triangulated topological disk with n faces, and T and B be two internally disjoint walks on $\mathcal{D}$'s boundary. Given $\tau > 0$, one can compute a τ-saturated isotopy class in $O(n^5 \log n)$ (resp. $O(n^3 \log n)$) time the continuous (resp. discrete) case.*

Lemma 3.9. *Let $\mathcal{D}$ be a triangulated topological disk with n faces, and T and B be two internally disjoint walks on the boundary of $\mathcal{D}$. Given a real number $x > 0$, one can either:*
(A) Compute a homotopy from T to B of width $O(x \log n)$.
(B) Return that $x < \mathsf{d}_{\mathcal{H}}(\mathsf{T}, \mathsf{B})$.
The running time of this procedure is $O(n^5 \log n)$ (resp. $O(n^3 \log n)$) in the continuous (resp. discrete) case.

Proof: Assume $x \geq \delta_H = \mathsf{d}_{\mathcal{H}}(\mathsf{T}, \mathsf{B})$, and we use x as a guess for this value δ_H. Using Lemma 3.8, one can compute a x-saturated isotopy class, h. Lemma 3.5 implies that both $\pi_{\mathrm{L},\mathsf{h}}$ and $\pi_{\mathrm{R},\mathsf{h}}$ are at most $4x$. Let $\mathcal{D}' \subseteq \mathcal{D}$ be the disc with boundary $\mathsf{T} \cup \cup \pi_{\mathrm{L},\mathsf{h}} \cup \mathsf{B} \cup \pi_{\mathrm{R},\mathsf{h}}$. By Lemma 3.6, all vertices in $\mathcal{D}'$ are in distance $O(x)$ from T and B (i.e., there are no $O(x)$-tall vertices in $\mathcal{D}'$). Finally, Lemma 3.3 implies that a continuous leash sequence of height $O(x \log n)$ between T and B, inside $\mathcal{D}'$, can be computed.

Thus, if x is larger than $\mathsf{d}_{\mathcal{H}}(\mathsf{T}, \mathsf{B})$ then this algorithm returns the desired approximation. Otherwise, it fails only if $x < \mathsf{d}_{\mathcal{H}}(\mathsf{T}, \mathsf{B})$. In the case of such failure, return that x is too small. ■

3.4 A strongly polynomial algorithm for the Frechét distance

Identifying the tall vertices.

Observe that using the algorithm of Lemma 3.9, we can decide given a candidate value δ_H for $\mathsf{d}_{\mathcal{H}}(\mathsf{T}, \mathsf{B})$ if it is too large, too small, or leads to the desired approximation. Indeed, if the algorithm returns an approximation of values $O(\delta_H \log n)$ but fails for $\delta_H/2$, we know it is the desired approximation.

So, compute for each vertex $v \in \mathsf{V}(\mathcal{D})$ its tallness; that is α_v would be the maximum distance of v to either T or B. Sort these values, and using binary search, compute the vertex w, with the minimum value α_w, such that Lemma 3.9 returns a parametrization with homotopic Frechét distance $O(\alpha_w \log n)$. If the algorithm of Lemma 3.9 returns that α_w/n is too small of a guess, then $[\alpha_w/n, \alpha_w \log n]$ contains δ_H. In this case, we can use binary search to find an interval $[\gamma/2, \gamma]$ that contains δ_H and use Lemma 3.9 to obtain the desired approximation. Similarly, if v is the tallest vertex shorter than w, then we can assume that $\alpha_v n$ is too small of a guess, otherwise we are again done as $[\alpha_v, \alpha_v n]$ contains δ_H.

As such, in the following, we know that the desired distance δ_H lies in interval $[x, y]$ where $x = \alpha_v n$ and $y = \alpha_w/n$, and for every vertex u of $\mathcal{D}$ it holds that (i) $\alpha_u < x/n$, or (ii) $\alpha_v > yn$. Naturally, we consider all the vertices that satisfy (ii) as tall vertices, by setting $\tau = 2x/n$. In the following, let M denote the set of these τ-tall vertices.

Rough approximation via greedy addition of vertices.

For a vertex $v \in \mathsf{V}(\mathcal{D})$, define $\mathrm{cost}(v)$ to be the length of the shortest path between t_l and b_l that has v on its left side. Similarly, for a set of vertices $X \subseteq \mathsf{V}(\mathcal{D})$, let $\mathrm{Cost}(X)$ be the length of the shortest path between t_l and b_l that has X on its left side. For a specific v or X, one can compute $\mathrm{cost}(v)$ and $\mathrm{cost}(X)$ by invoking the algorithm of Lemma 3.7 once.

The following is easy to verify.

Lemma 3.10. *For $X, Y \subseteq \mathsf{V}(\mathcal{D})$, we have $\mathrm{Cost}(X \cup Y) \leq \mathrm{Cost}(X) + \mathrm{Cost}(Y)$.*

Consider the algorithm that starts with $X_0 = \emptyset$. In the ith iteration, the algorithm computes the vertex $v_i \in \mathsf{M} \setminus X_{i-1}$, such that $\mathrm{Cost}(X_{i-1} \cup \{v_i\})$ is minimized, and set $X_i = X_{i-1} \cup \{v_i\}$. Let h_i be the isotopy class having X_i on its left side, and $\mathsf{M} \setminus X_i$ on its right side.

Lemma 3.11. *The cheapest homotopic Frechét parametrization among $\mathsf{h}_1, \ldots, \mathsf{h}_n$ has width $O(\mathsf{d}_{\mathcal{H}}(\mathsf{T}, \mathsf{B})\, n \log n)$.*

Proof: Consider the set Y that is the subset of tall vertices on the left side of the optimal solution. Let i be the first index such that $Y \subseteq X_i$ and $Y \not\subseteq X_{i-1}$. Let v be any vertex in $Y \setminus X_{i-1}$. We have that

$$\begin{aligned}\mathrm{Cost}(X_i) &\leq \mathrm{Cost}(X_{i-1} \cup \{v\}) \leq \mathrm{Cost}(X_{i-1}) + \mathrm{cost}(v)\\ &\leq \mathrm{Cost}(X_{i-1}) + \mathrm{Cost}(Y)\\ &\leq \mathrm{Cost}(X_{i-2}) + 2\mathrm{Cost}(Y) \leq \cdots \leq i\mathrm{Cost}(Y)\\ &\leq n\mathrm{Cost}(Y)\,.\end{aligned}$$

Now, setting $\tau = \mathrm{Cost}(X_i)$, it follows that the set X_i is τ-saturated. Applying Lemma 3.5, implies that $\|\pi_{\mathrm{R},\mathsf{h}_i}\| \leq 4\tau$. Observe, that the disk defined by T, $\pi_{\mathrm{L},\mathsf{h}_i}$, B, $\pi_{\mathrm{R},\mathsf{h}_i}$ can not contain any tall vertex (by construction).

Now, plugging this into Lemma 3.3 implies the homotopic Frechét width of h_i (starting with $\pi_{\mathrm{L},\mathsf{h}_i}$ and ending up with $\pi_{\mathrm{R},\mathsf{h}_i}$) is $O(\tau \log n)$, which implies the claim since $\mathrm{Cost}(X_i) \leq n\mathrm{Cost}(Y) \leq n\mathsf{d}_{\mathcal{H}}(\mathsf{T}, \mathsf{B})$. ■

The algorithm.

We approximate the homotopic Frechét width of each one of the classes $\mathsf{h}_1, \ldots, \mathsf{h}_n$. We identify the value x with the smallest width, and we do a binary search in the interval $[x/n^2, x]$ for the homotopic Frechét distance.

Theorem 3.12. *Let $\mathcal{D}$ be a topological disk that is triangulated and has n faces, and T and B be two internally disjoint walks on the boundary of $\mathcal{D}$. One can compute a homotopic Frechét parametrization of width $O(\mathsf{d}_{\mathcal{H}}(\mathsf{T}, \mathsf{B}) \log n)$, where $\mathsf{d}_{\mathcal{H}}(\mathsf{T}, \mathsf{B})$ is the homotopic Frechét distance between T and B in $\mathcal{D}$.*

The running time of this procedure is $O(n^6 \log n)$ (resp. $O(n^4 \log n)$) in the continuous (resp. discrete) case.

Proof: The algorithm requires $O(n^2)$ calls to Lemma 3.7, which takes $O\big(n^6 \log n\big)$ (resp. $O\big(n^4 \log n\big)$) time in the continuous (resp. discrete) case. Then, the algorithm requires Lemma 3.3 to compute the homotopic Frechét distance of the classes $\mathsf{h}_1, \ldots, \mathsf{h}_n$. The algorithm also performs $O(\log n)$ calls to the algorithm of Lemma 3.9. ■

4. CONCLUSIONS

We presented a $O(\log n)$ approximation algorithm for approximating the homotopy height and the homotopic Frechét distance between curves on piecewise linear surfaces. It seems quite believable that the approximation quality can be further improved, and we leave this as the main open problem of our work. Since our algorithm works both for the continuous and discrete cases, it seems natural to conjecture that this algorithm should work more general surfaces and metrics.

5. REFERENCES

[1] H. Alt and M. Buchin. Semi-computability of the Frechét distance between surfaces. In *Proc. 21st Euro. Workshop on Comput. Geom.*, 2005.

[2] H. Alt and M. Godau. Computing the Fréchet distance between two polygonal curves. *Internat. J. Comput. Geom. Appl.*, 5:75–91, 1995.

[3] Chakib Bennis, Jean-Marc Vézien, Gérard Iglésias, and André Gagalowicz. Piecewise surface flattening for non-distorted texture mapping. In Thomas W. Sederberg, editor, *Proc. SIGGRAPH '91*, volume 25, pages 237–246, 1991.

[4] S. Brakatsoulas, D. Pfoser, R. Salas, and C. Wenk. On map-matching vehicle tracking data. In *Proc. 31st VLDB Conference*, pages 853–864, 2005.

[5] K. Buchin, M. Buchin, and J. Gudmundsson. Detecting single file movement. In *Proc. 16th ACM SIGSPATIAL Int. Conf. Adv. GIS*, pages 288–297, 2008.

[6] K. Buchin, M. Buchin, J. Gudmundsson, Maarten L., and J. Luo. Detecting commuting patterns by clustering subtrajectories. In *Proc. 19th Annu. Internat. Sympos. Algorithms Comput.*, pages 644–655, 2008.

[7] E. W. Chambers, E. C. de Verdière, J. Erickson, S. Lazard, F. Lazarus, and S. Thite. Walking your dog in the woods in polynomial time. In *Proc. 24th Annu. ACM Sympos. Comput. Geom.*, pages 101–109. ACM, 2008.

[8] E. W. Chambers and D. Letscher. On the height of a homotopy. In *Proc. 21st Canad. Conf. Comput. Geom.*, 2009.

[9] E. W. Chambers and D. Letscher. Erratum for on the height of a homotopy. `http://mathcs.slu.edu/~chambers/papers/hherratum.pdf`, 2010.

[10] A. F. Cook, A. Driemel, S. Har-Peled, J. Sherette, and C. Wenk. Computing the Fréchet distance between folded polygons. In *Proc. 12th Workshop Algorithms Data Struct.*, pages 267–278, 2011.

[11] A. F. Cook and C. Wenk. Geodesic Fréchet distance inside a simple polygon. *ACM Trans. Algo.*, 7:9:1–9:19, 2010.

[12] A. Driemel, S. Har-Peled, and C. Wenk. Approximating the Fréchet distance for realistic curves in near linear time. In *Proc. 26th Annu. ACM Sympos. Comput. Geom.*, pages 365–374, 2010.

[13] A. Efrat, L. J. Guibas, S. Har-Peled, J. S.B. Mitchell, and T.M. Murali. New similarity measures between polylines with applications to morphing and polygon sweeping. *Discrete Comput. Geom.*, 28:535–569, 2002.

[14] T. Eiter and H. Mannila. Computing discrete Frechét distance. Tech. Report CD-TR 94/64, Christian Doppler Lab. Expert Sys., TU Vienna, Austria, 1994.

[15] M. S. Floater. Parameterization and smooth approximation of surface triangulations. *Comput. Aided Geom. Design*, 14(4):231–250, 1997.

[16] M. Frechét. Sur quelques points du calcul fonctionnel. *Rendic. Circ. Mat. Palermo*, 22:1–74, 1906.

[17] M. Frechét. Sur la distance de deux surfaces. *Ann. Soc. Polonaise Math.*, 3:4–19, 1924.

[18] M. Godau. *On the complexity of measuring the similarity between geometric objects in higher dimensions*. PhD thesis, Free University of Berlin, 1999.

[19] S. Har-Peled, A. Nayyeri, M. Salavatipour, and A. Sidiropoulos. How to walk your dog in the mountains with no magic leash. `http://valis.cs.uiuc.edu/~sariel/papers/11/homotopy/`, 2012.

[20] S. Har-Peled and B. Raichel. The Fréchet distance revisited and extended. In *Proc. 27th Annu. ACM Sympos. Comput. Geom.*, pages 448–457, 2011. `http://www.cs.uiuc.edu/~sariel/papers/10/frechet3d/`.

[21] M. R. Henzinger, P. Klein, S. Rao, and S. Subramanian. Faster shortest-path algorithms for planar graphs. *J. Comput. Sys. Sci.*, 55:3–23, August 1997.

[22] E. J. Keogh and M. J. Pazzani. Scaling up dynamic time warping to massive dataset. In *Proc. of the Third Euro. Conf. Princip. Data Mining and Know. Disc.*, pages 1–11, 1999.

[23] M.S. Kim, S.W. Kim, and M. Shin. Optimization of subsequence matching under time warping in time-series databases. In *Proc. ACM symp. Applied comput.*, pages 581–586, 2005.

[24] A. Mascret, T. Devogele, I. Le Berre, and A. Hénaff. Coastline matching process based on the discrete Fréchet distance. In *Proc. 12th Int. Sym. Spatial Data Handling*, pages 383–400, 2006.

[25] D. Piponi and G. Borshukov. Seamless texture mapping of subdivision surfaces by model pelting and texture blending. In *Proc. SIGGRAPH 2000*, pages 471–478, August 2000.

[26] J. Serrà, E. Gómez, P. Herrera, and X. Serra. Chroma binary similarity and local alignment applied to cover song identification. *IEEE Transactions on Audio, Speech & Language Processing*, 16(6):1138–1151, 2008.

[27] Alla Sheffer and Eric de Sturler. Surface parameterization for meshing by triangulation flattening. In *Proc. 9th International Meshing Roundtable*, pages 161–172, 2000.

[28] C. Wenk, R. Salas, and D. Pfoser. Addressing the need for map-matching speed: Localizing global curve-matching algorithms. In *Proc. 18th Int. Conf. Sci. Statis. Database Manag.*, pages 879–888, 2006.