Volume-based similarity of linear features on terrains

Citation for published version (APA):

DOI:
10.1145/3274895.3274937

Document status and date:
Published: 06/11/2018

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

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

Link to publication

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

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

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

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.
Volume-based Similarity of Linear Features on Terrains

Willem Sonke
TU Eindhoven
The Netherlands
w.m.sonke@tue.nl

Marc van Kreveld
Utrecht University
The Netherlands
m.j.vankreveld@uu.nl

Tim Ophelders
TU Eindhoven
The Netherlands
t.a.e.ophelders@tue.nl

Bettina Speckmann
TU Eindhoven
The Netherlands
b.speckmann@tue.nl

Kevin Verbeek
TU Eindhoven
The Netherlands
k.a.b.verbeek@tue.nl

ABSTRACT
Linear features on terrains model the boundaries of ground cover regions, delineate glaciers, or form the boundary of rivers and lakes. When computing the similarity between such linear features, it is important to also take their context into account: the terrain. We hence explore the possibilities of volume-based distance measures for linear features on a terrain. Our measures construct suitable base surfaces between the linear features, which can slice through the input terrain and also hover above. The similarity between two linear features is then captured by the volume of “earth” above the base surface and below the terrain, and possibly also by the volume of “air” below the base surface and above the terrain. We suggest six ways of choosing a suitable base surface. These choices give rise to different measured volumes and can be useful in different application scenarios.

1 INTRODUCTION
Computing the similarity between linear features is an important part of many analysis methods for geographic data. Linear features appear, for example, as the boundary of ground cover regions, separating different types of vegetation or delineating glaciers. They also describe the course of rivers, and form the boundary of river beds and lakes. Linear features can change over time: glaciers expand or retract, certain kinds of vegetation cover larger areas, while other kinds are being decimated. rivers change their course and, correspondingly, their beds. Hence, similarity of linear features can be used in trend analysis. When computing the similarity between the boundaries of a glacier in two different years, it is important to also consider the context in which the change occurred: both the terrain itself and the geomorphological processes that caused or are caused by this change.

In mountainous terrain the region in which a certain species of moss occurs is influenced both by the terrain itself, but also by wind exposure or changing temperatures. The effort (or time) needed for a river bed to move its course to a new river bed depends on the amount of sediment that must be eroded between these river beds. The third (terrain) dimension may also represent something else than elevation, for example water temperature. When fish or whales migrate they may choose different routes through the ocean, possibly preferring colder or warmer waters. In this case the (dis)similarity between two routes is influenced both by the amount and the temperature of the water in between.

In most cases, the similarity between two linear features is determined by “whatever lies in between the two features”. If the two features are simply two paths in the plane, without any further context, then a classic geometric distance measure such as Hausdorff, Dynamic Time Warping, or Fréchet can be appropriate. The area between the two linear features can also be meaningful. However, in a 3-dimensional setting, where the context of the two features to be compared is given by a terrain, we need to be able to take this context into account. In this paper we hence propose a set of volume-based distance measures for linear features on terrains. Specifically, our measures construct suitable base surfaces between the two linear features to be compared. These base surfaces can slice through the input terrain and also hover above, to model the costs of geomorphological processes such as erosion and fill. The similarity between two linear features is then captured by the “earth” above the base surface and possibly also by the “air” trapped below (see Fig. 1). It is useful to have a collection of different similarity measures available, so that the most appropriate one can be used depending on the application. Our paper extends this collection with new volume-based measures for 3-dimensional situations.

Results and organization. We assume that our input consists of a two-dimensional surface \( T \) with a height function \( h : \mathbb{R}^2 \rightarrow \mathbb{R} \),
representing a terrain, and two paths \( \pi_1 \) and \( \pi_2 \), describing two linear features on the surface. For ease of explanation, we furthermore assume that \( \pi_1 \) and \( \pi_2 \) share their start and end points but are otherwise disjoint. Our methods can in fact easily be applied to intersecting paths as well, by considering all parts between intersections separately and summing the resulting volumes. Furthermore, if \( \pi_1 \) and \( \pi_2 \) do not share their start (end) points, we can draw a path \( \pi_{\text{start}}(\pi_{\text{end}}) \) connecting the start (end) points of \( \pi_1 \) and \( \pi_2 \), in order to delineate the part of \( T \) that is between \( \pi_1 \) and \( \pi_2 \).

We explore the possibilities of volume-based distance measures between \( \pi_1 \) and \( \pi_2 \) with respect to \( T \). Earth and air are usually not equivalent, so we aim to study both symmetric and asymmetric measures. Specifically, in Section 2.1 we first consider the most basic of surfaces, namely planes, and argue how to place them optimally. We describe three such planes: the horizontal plane that fits the paths \( \pi_1 \) and \( \pi_2 \) best, the general plane that fits \( \pi_1 \) and \( \pi_2 \) best, and the plane that minimizes the total (vertical) area between the plane and \( \pi_1 \) and \( \pi_2 \). In Section 2.2 we consider more general base surfaces which do contain \( \pi_1 \) and \( \pi_2 \) (which usually cannot be the case with planes). We consider the minimal area surface between \( \pi_1 \) and \( \pi_2 \), as well as two other types of surfaces, which are motivated by water flow on terrains; the latter surfaces by definition do not trap air below them. Note that the input surface \( T \) itself is also a possible base surface: the only surface for which \( \pi_1 \) and \( \pi_2 \) are considered identical (since there is no earth above and no air below \( T \) with respect to \( T \)). The input surface is hence a poor choice of base surface.

Related work. Similarity measures for shapes have been considered in a variety of contexts. Popular geometric measures include the Hausdorff distance, the Fréchet distance, the area-of-symmetric difference, the Wasserstein distance (Earth Mover’s Distance), and the turn function distance [1–4, 20, 23]. Shape similarity measures have also been well-studied in the context of multimedia search and retrieval, see for example, the overview given by Veltkamp [25].

Also in GIS shape similarity measures have been used, for example, in cartographic generalization. Here a city outline or river shape will be displayed with less detail on a smaller-scale map, while still capturing the overall shape well. One needs to measure the similarity between the original shape and the generalized shape to determine how well the generalization still resembles the original [18, 22]. Furthermore, similarity measures are used for trajectory similarity [24], landscape ecology [12], (urban) property analysis [13], spatio-temporal processes [17], and retrieval in spatial databases [19]. There has also been some recent interest in semantic similarity [14, 21], focusing more on cognitive than on geometric aspects of similarity.

As described above, our volume-based measures are modeled to take the 3-dimensional context of the input paths into account by measuring volume in between the two paths. The use of context in geographic similarity measures was studied for trajectories in [6]. Several other computational methods that include context are based on the Fréchet distance [7, 8, 11].

2 VOLUME-BASED MEASURES

In this section we provide definitions for the six base surfaces. Like the input terrain, we can represent a base surface as a function \( \mathcal{B} : \mathbb{R}^2 \rightarrow \mathbb{R} \). Given such a function, the volume of air above the base surface is:

\[
d(\pi_1, \pi_2) = \int_D \max \left(0, h(x, y) - \mathcal{B}(x, y)\right) \, dx \, dy,
\]

where \( D \) is the part of the terrain enclosed by \( \pi_1 \) and \( \pi_2 \). Alternatively, we can measure the volume of air below the base surface by inverting the terrain and the base surface. We study two types of base surfaces: planes (Section 2.1) and general surfaces (Section 2.2).

We take planes that are close to the paths, as we want to measure volume with respect to the paths. If two paths surround a hill, we wish to measure the volume of the hill above the height of the paths; the height or volume of the hill should not influence the choice of the plane, just the amount we measure. Planes are the simplest type of base surface, but as planes cannot always capture the 3D shape of the paths well, we also consider general (non-planar) surfaces that always contain the paths.

2.1 Planes

A natural way of defining a base surface is to take a plane that is as close as possible to the paths \( \pi_1 \) and \( \pi_2 \). The simplest way to define such a plane is to use a horizontal plane \( z = h_{\text{avg}} \) at the average height \( h_{\text{avg}} \) of \( \pi_1 \cup \pi_2 \). We call the resulting plane the horizontal average plane (HAP).

The HAP is horizontal by definition, while \( \pi_1 \) and \( \pi_2 \) may be sloped, so the HAP may not stay close to the paths. Therefore we also consider arbitrary planes \( z = ax + by + c \). A suitable plane that is close to the paths is a regression plane using least-squares regression. That is, we choose \( a \), \( b \) and \( c \) such that the sum of squared errors:

\[
\int_{p \in \pi_1 \cup \pi_2} (zp - ax_p - by_p - c)^2 \, dp
\]

is minimized. We call the resulting plane a regression plane (RP). The parameters \( a \), \( b \) and \( c \) defining the plane can be computed using standard methods from statistics. We note that the HAP, being a horizontal plane at an average height, is the horizontal plane given by \( z = c \) that minimizes the sum of squared errors:

\[
\int_{p \in \pi_1 \cup \pi_2} (zp - c)^2 \, dp.
\]

In other words, the HAP is a version of the regression plane that is restricted to be horizontal.

Both the HAP and the RP are (generally) partly above and partly below the terrain. Therefore we can choose to measure only earth or include air as well. If we measure only earth, it is clear that a
higher horizontal plane will measure a smaller volume of earth. Similarly, when measuring only air, the air volume increases. In some applications one might want to treat earth and air as equals and use a choice for a plane that is symmetric. Since we choose the plane based on the paths, we can use a base plane that has as much air above it as earth below it at the paths. This means that we are choosing a horizontal plane based on the (vertical) area of earth and air at the vertical cross-section induced by \( \pi_1 \) and \( \pi_2 \). This setting is especially natural if the volume we will eventually measure is the sum of earth and air volumes with respect to the base surface. Therefore, as a third plane choice, we pick the horizontal plane given by \( z = c \) that minimizes the sum of absolute differences:

\[
\int_{p \in \pi_1 \cup \pi_2} |zp - c| \, dp.
\]

We call this plane the minimizing horizontal plane (MHP). Assuming a TIN as the terrain model, we can find the cut-off height \( c \) of a MHP by sweeping a plane from top to bottom while maintaining area of earth above (initially zero) and air below \( \pi_1 \cup \pi_2 \). These areas change with a quadratic function in \( z \), which needs to be updated at every vertex. We search for the plane where the increase of earth is the same as the decrease of air, because here the minimality is realized. The approach is the same as classification on a TIN in [16].

### 2.2 General surfaces

When \( \pi_1 \) and \( \pi_2 \) lie roughly in a plane, some plane base surfaces appear to have the desired behavior on simple examples, see Figure 2(a)–(c). However, when the paths clearly deviate from any plane, then the planes do not perform well. See Figure 3, where the paths \( \pi_1 \) and \( \pi_2 \) have the same profile and so does the terrain in between, but all base planes will measure volumes of earth and/or air. If there is a “natural morph” from \( \pi_1 \) to \( \pi_2 \) which happens to follow the surface of \( T \), then we prefer not to measure any volume. We hence consider more general base surfaces, which contain the paths \( \pi_1 \) and \( \pi_2 \).

There are many ways to define a surface between the paths \( \pi_1 \) and \( \pi_2 \). For example, we could use a minimal surface or a surface obtained from a constrained Delaunay triangulation between \( \pi_1 \) and \( \pi_2 \) [9, 10]. Unfortunately, the constrained Delaunay triangulation gives a base surface that is not robust: a minor change in the location of one vertex of \( \pi_1 \) can cause an edge flip and therefore a vastly different base surface, causing a dramatic change in measured volume. However, a minimal area surface (MAS) between the paths \( \pi_1 \) and \( \pi_2 \) is a natural choice. When the paths lie in a plane, the minimal area surface is that plane, and hence for the terrains in Figure 2(a)–(c), the results are the same as for RP. For terrains such as the one in Figure 3 we will measure some positive volume of earth and air.

The last two base surfaces that we consider are motivated by geomorphological processes, specifically, water flow on terrains. In particular, we are interested in the minimum amount of earth that must be removed for a stream to move from course \( \pi_1 \) to course \( \pi_2 \). This scenario implies that we are interested in an asymmetric distance measure, because it is “easier” for a path (stream) higher up to change its flow to a lower path than vice versa.

We can view the change from path \( \pi_1 \) to path \( \pi_2 \) as a morph. Imagine the projections \( \pi_1' \) and \( \pi_2' \) of \( \pi_1 \) and \( \pi_2 \) onto the \((x, y)\)-plane and let \( D \) denote the part of the terrain between \( \pi_1' \) and \( \pi_2' \). A morph that transforms \( \pi_1' \) to \( \pi_2' \) for which all intermediate paths are simple and where no part of \( D \) is covered more than once is called a monotone isotopy (see Fig. 4(a)). It is a smooth transition that can only move forward from \( \pi_1' \) towards \( \pi_2' \) and must stay inside \( D \). There are many such monotone isotopies. In 3D, the morph from \( \pi_1 \) to \( \pi_2 \) also includes an elevation for every point on every intermediate path of the morph. Since the morph is smooth and forward only, it defines a function over \( D \). If we choose the elevations such that this function is continuous over \( D \), its image is a surface which can be used as a base surface.

Each individual path in a morph can be seen as a mapping from the interval \([0, 1]\) to the path from \( s \) to \( t \). For all paths in the morph and any parameter \( r \in [0, 1] \) we then get a smooth sequence of locations by tracing the points at parameter \( r \) over all intermediate paths of the isotopy (see Fig. 4(b)). These traces are curves that are transverse to the intermediate curves; they are referred to as matching curves. They connect one location on \( \pi_1 \) to one location on \( \pi_2 \) with a smooth curve.

![Figure 2: (a) A flat terrain, (b) a hill, and (c) a sloped hill with longitudinal paths.](image)

![Figure 3: A valley with transverse paths (left), cut by a horizontal base plane (right). Any base plane that stays close to the input paths will cut off earth or air.](image)

![Figure 4: (a) A monotone isotopy in \( D \) illustrated by intermediate paths of the morph between \( \pi_1' \) and \( \pi_2' \). (b) Two matching curves of this isotopy. The matching curve at parameter 0 is simply \( s' \) and the matching curve at parameter 1 is \( t' \), the projections of \( s \) and \( t \).](image)
To let (the stream) $\pi_1$ change its course to $\pi_2$, all matching curves should be monotonically decreasing in elevation. Therefore, we want the base surface to be a surface on or below $T$, for which a corresponding monotone isotopy exists whose matching curves are all monotonically decreasing. Among these, we pick the one that measures the smallest volume of earth. We call the resulting surface the \textit{water flow surface} (WFS).

To understand the WFS and how it differs from the RP, even if the paths $\pi_1$ and $\pi_2$ lie in a plane, consider the example in Figure 5(a). Here the RP distance is the volume of the entire bump since it is above the RP (see Fig. 5(b)), whereas the WFS distance from the upper to the lower path is the volume of the bump above the saddle point (see Fig. 5(c)). Similarly, if the terrain between $\pi_1$ and $\pi_2$ would contain unevennesses but no local maxima, then the WFS distance would be $0$ and the RP distance would be some measurable value. The WFS distance will also be $0$ for the example in Figure 3 (left), but for the example in Figure 6 it gives the volume between $T$ and a horizontal plane through $s$ and $t$ on one side of the valley (restricted to the domain $D$).

Another way of understanding the WFS distance is that it is the volume of all points on or under $T$ for which any path over the terrain towards $\pi_1$ must descend at some point. Using this observation we can efficiently compute the WFS on a TIN-based terrain. We do this by computing, for each vertex $v \in T$, the highest path $\pi_1$ towards $\pi_1$. This is defined as the path with the highest minimum-height point among all paths from $v$ to $\pi_1$. That is, it descends only when it must. To efficiently compute all highest paths, we construct a \textit{highest path tree} that contains all highest paths towards $\pi_1$. The required algorithms are described in [5, 15].

The WFS is asymmetric: swapping $\pi_1$ and $\pi_2$ results in a different base surface. The last base surface we consider is the symmetric version of the WFS, and hence we refer to it as the \textit{symmetric flow surface} (SFS). It is the volume of all points on or below $T$ for which any path over the terrain to $\pi_1 \cup \pi_2$ must descend. Since we have a wider choice of destinations for these paths, we will never measure more volume with the SFS distance than with the WFS distance. The SFS measures a volume of $0$ in Figure 6, unlike the other base surfaces we described.

ACKNOWLEDGMENTS

W. Sonke, M. van Kreveld, T. Ophelders, B. Speckmann and K. Verbeek are supported by the Netherlands Organisation for Scientific Research (NWO) under project no. 639.023.208 (W.S., T.O., and B.S.), no. 612.001.651 (M.v.K.), and no. 639.021.541 (K.V.).

REFERENCES