Title

The Edit Distance for Reeb Graphs of Surfaces

Authors

B. Di Fabio and C. Landi

Abstract

Reeb graphs are structural descriptors that capture shape properties of a topological space from the perspective of a chosen function. In this work we define a combinatorial metric for Reeb graphs of orientable surfaces in terms of the cost necessary to transform one graph into another by edit operations. The main contributions of this paper are the stability property and the optimality of this edit distance. More precisely, the stability result states that changes in the functions, measured by the maximum norm, imply not greater changes in the corresponding Reeb graphs, measured by the edit distance. The optimality result states that our edit distance discriminates Reeb graphs better than any other metric for Reeb graphs of surfaces satisfying the stability property.

Summary

Given a scalar field $(\mathbb{X}, f)$, we can create its Reeb graph $(\mathbb{X}_f, \tilde{f})$. While the latter is a topological space, we define the *combinatorial* Reeb graph Γ_f as being the graph whose vertices are critical points of the scalar field (and are consequentially vertices of $(\mathbb{X}_f, \tilde{f})$.) The edges of Γ_f are exactly the edges of the Reeb graph (contracted cylinders which connect critical points). We define a labeling $l_f : V(\Gamma_f) \to \mathbb{R}$ on Γ_f by stating that $l_f = f|_{C_f}$, where C_f are the critical points of the scalar field. We call a *deformation* of (Γ_f, l_f) to be any finite ordered sequence $T = (T_1, \ldots, T_n)$, where each T_i is an *elementary deformation* (table given below). Furthermore, the elementary deformations have to act on the previously created graph (so an iterative process). We say $\mathcal{T}((\Gamma_f, l_f), (\Gamma_g, l_g)$ is the set of deformations turning (Γ_f, l_f) into (Γ_g, l_g).

Each elementary operation has an associated cost. We take each associated cost and sum them up to get the total cost of the entire transformation. The *Reeb Graph Edit Distance* is the infimum over all possible deformations from (Γ_f, l_f) to (Γ_g, l_g).

The set of deformations $\mathcal{T}$ is guaranteed to be non empty IF the two Reeb graphs are defined on the same domain. Furthermore, it is known to be less than or equal to the maximum function difference ranging over all points $x \in X$.

Let $\mathbb{M}$ be a 2-Dimensional manifold (a surface) and f be a simple Morse function. The term "simple" refers to the fact that each critical level contains only one critical point (so the common definition that we have used). Let $\mathcal{D}(\mathbb{M})$ be the set of *self-diffeomorphisms* on $\mathbb{M}$. We say that f, g are *Right Equivalent* if there exists a diffeomorphism $\zeta \in \mathcal{D}(\mathbb{M}$ such that $f = g \circ \zeta$. It can be shown that two functions defined on the same $\mathbb{M}$ are R-equivalent if and only if their labeled Reeb graphs are isomorphic. We can now define a distance between two equivalence classes of simple morse functions defined on $\mathbb{M}$ (meaning the two associated Reeb graphs are NOT isomorphic) by $d_N([f], [g]) = \inf_{\zeta \in \mathcal{D}(\mathbb{M})} ||f - g \circ \zeta||_{C^0}$. The C^0 refers to the maximum function distance ranging over the set of points in the manifold. This distance is called the *natural pseudo-idstance*. Instead of self-diffeomorphisms, we can have the same exact definition using self-homeomorphisms on $\mathbb{M}$ (not really sure the difference to be honest).

It can then be shown that $d_E((\Gamma_f, l_f), (\Gamma_g, l_g)) = d_N([f], [g])$. Thus, this edit distance is a metric on isomorphism classes of labeled Reeb Graphs. Furthermore, for every pair of simple morse functions f, g defined on $\mathbb{M}$, we have that $d_E((\Gamma_f, l_f), (\Gamma_g, l_g)) \geq d_B(D_f, D_g)$, where d_B is the botteleneck distance and D_f, D_g are the persistence diagrams of f, g, respectively. In addition, $d_E((\Gamma_f, l_f), (\Gamma_g, l_g)) \geq d_F D((\mathbb{M}_f, \tilde{f}), (\mathbb{M}_g, \tilde{g}))$, where $d_F D$ refers to the functional distortion distance between the Reeb graphs $(\mathbb{M}_f, \tilde{f}), (\mathbb{M}_g, \tilde{g})$.

Below are the formal definitions for each elementary deformation.

- For T elementary of type (B), inserting the vertices $u_1, u_2 \in V(\Gamma_g)$, the associated cost is
$$c(T) = \frac{|\ell_g(u_1) - \ell_g(u_2)|}{2}.$$
- For T elementary of type (D), deleting the vertices $u_1, u_2 \in V(\Gamma_f)$, the associated cost is
$$c(T) = \frac{|\ell_f(u_1) - \ell_f(u_2)|}{2}.$$
- For T elementary of type (R), relabeling the vertices $v \in V(\Gamma_f) = V(\Gamma_g)$, the associated cost is
$$c(T) = \max_{v \in V(\Gamma_f)} |\ell_f(v) - \ell_g(v)|.$$
- For T elementary of type (K_i), with $i = 1, 2, 3$, relabeling the vertices $u_1, u_2 \in V(\Gamma_f)$, the associated cost is
$$c(T) = \max\{|\ell_f(u_1) - \ell_g(u_1)|, |\ell_f(u_2) - \ell_g(u_2)|\}.$$
- For $T \in \mathscr{T}((\Gamma_f, \ell_f), (\Gamma_g, \ell_g))$, with $T = (T_1, \ldots, T_r)$, the associated cost is
$$c(T) = \sum_{i=1}^{r} c(T_i).$$

Below are examples for each elementary deformation.

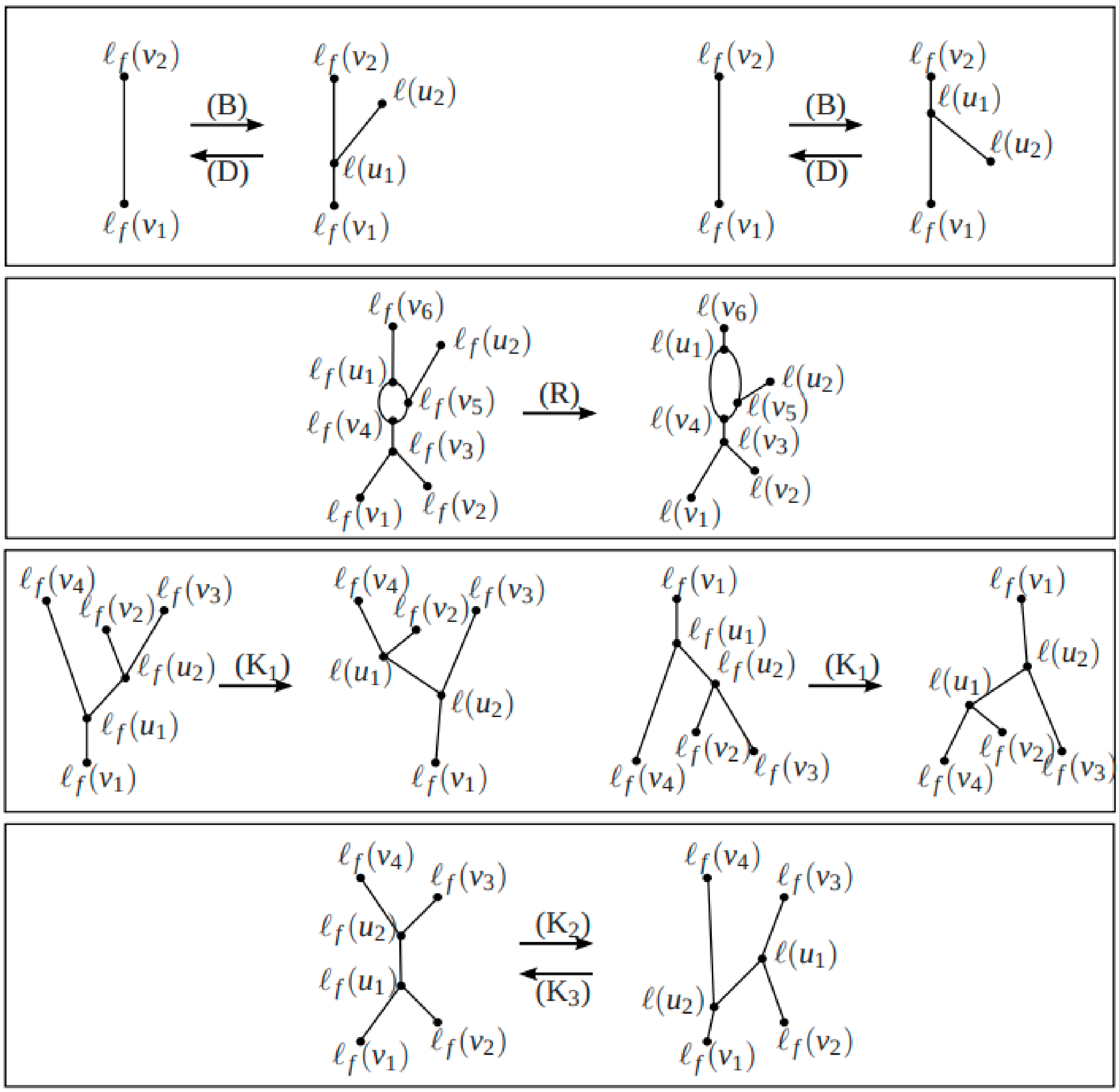

TABLE 1. A schematization of the elementary deformations of a labeled Reeb graph provided by Definition 2.1.

Questions

- Is a two dimensional topological space equipped with a function, which turns out to be a scalar field, equivalent to just a 2-Dimensional Manifold equipped with a height function?

Theoretical Contributions

Edit distance can be shown to be more discriminative than both the bottleneck distance of persistence diagrams and functional distortion distance between Reeb graphs. Because of this, I believe it is automatically more discriminative than the interleaving distance between Reeb graphs. Check the "strong equivalence of functional distortion distance and interleaving distance" paper to check .

Computational Contributions

None