

Distance Functions and Geodesics on Points Clouds

Facundo Mémoli*

Guillermo Sapiro†

Abstract

An algorithm for computing intrinsic distance functions and geodesics on sub-manifolds of $\mathbb{R}^d$ given by point clouds is introduced in this paper. The basic idea is that, as shown in this paper, intrinsic distance functions and geodesics on general co-dimension sub-manifolds of $\mathbb{R}^d$ can be accurately approximated by the extrinsic Euclidean ones computed in a thin offset band surrounding the manifold. This permits the use of computationally optimal algorithms for computing distance functions in Cartesian grids. We then use these algorithms, modified to deal with spaces with boundaries, and obtain also for the case of intrinsic distance functions on sub-manifolds of $\mathbb{R}^d$, a computationally optimal approach. For point clouds, the offset band is constructed without the need to explicitly find the underlying manifold, thereby computing intrinsic distance functions and geodesics on point clouds while skipping the manifold reconstruction step. The case of point clouds representing noisy samples of a sub-manifold of Euclidean space is studied as well. All the underlying theoretical results are presented, together with experimental examples, and comparisons to graph-based distance algorithms.

1 Introduction

One of the most popular sources of point clouds are 3D shape acquisition devices, such as laser range scanners, with applications in geoscience, art (e.g., archival), medicine (e.g., prosthodontics), manufacturing (from cars to clothes), and security (e.g., recognition), among other disciplines. These scanners provide in general raw data in the form of unorganized point clouds representing surface samples. With the increasing popularity and very broad applications of this source of data, it is natural and important to work directly with this representation, without having to go to the intermediate step of fitting a surface to it (step that can add computational complexity and introduce errors). See for example [6, 9, 10, 12, 13, 20, 21, 24, 25] for a few of the recent works with this type of data. Note that point clouds can also be used as primitives for visualization, e.g., [7, 13, 26], as well as for editing [34].

Another important field where point clouds are found is in the representation of high-dimensional manifolds by examples (see for example [15, 19, 30]). This type of high-dimensional and general co-dimension data appears in almost all disciplines, from computational biology to image analysis to financial data.

Note that in general a point cloud representation is dimension free, in contrast with other popular representations such as triangular meshes. Some operations, such as the union of point clouds acquired from multiple views, are much easier than when performed on the triangular meshes obtained from them. This paper addresses one of the most fundamental operations in the study and processing of sub-manifolds of Euclidean space, the computation of intrinsic distance functions and geodesics. We show that this can be done directly working with the point cloud, without the need for reconstructing the underlying manifold. We present the corresponding theoretical results, experimental examples, and basic comparisons to mesh-based

*Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455; and Instituto de Ingeniería Eléctrica, Universidad de la República, Montevideo, Uruguay.

†Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455, guille@ece.umn.edu (corresponding author).

Report Documentation Page			Form Approved OMB No. 0704-0188		
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE 2005		2. REPORT TYPE		3. DATES COVERED -	
4. TITLE AND SUBTITLE Distance Functions and Geodesics on Points Clouds				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Office of Naval Research,One Liberty Center,875 North Randolph Street Suite 1425,Arlington,VA,22203-1995				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES The original document contains color images.					
14. ABSTRACT see report					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES 10	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

distance algorithms. The results are valid for general dimensions and co-dimensions, and for manifolds with or without boundary. These results include the analysis of noisy point clouds obtained from sampling the manifold. The theoretical results are presented without proof due to space limitations. The corresponding proofs can be found in our extended report [23].

A number of key building blocks form part of the framework here introduced. The first one is based on the fact that distance functions intrinsic to a given sub-manifold of $\mathbb{R}^d$ can be accurately approximated by Euclidean distance functions computed in a thin offset band that surrounds this manifold. This concept was first introduced in [22], where convergence results are given for co-dimension one hyper-surfaces without boundary. This result is reviewed in §2. In this paper, we first extend this to general co-dimension and to deal with manifolds with or without boundary, §3. We also show that the approximation is true not only for the intrinsic distance function but also for the intrinsic geodesic. This is not a straightforward corollary, since geodesics are based on the gradient of the distance function,¹ which contains singularities at the cut locus [33]. The approximation of intrinsic distance functions (and geodesics) by extrinsic Euclidean ones permits to compute them using computationally optimal algorithms in Cartesian grids. These algorithms are based on the fact that the distance function satisfies a Hamilton-Jacobi partial differential equation (see §2), for which consistent and fast algorithms have been developed in Cartesian grids [14, 28, 29, 32]² (see [16] for extensions to triangular meshes and [31] for other Hamilton-Jacobi equations). That is, due to these results, we can use computationally optimal algorithms in Cartesian grids (with boundaries) also to compute distance functions, and from them geodesics, intrinsic to a given manifold, and in a computationally optimal fashion. Note that in contrast with the popular Dijkstra algorithm, these numerical techniques are consistent, they converge to the true distance when the grid is refined.

Once this basic results are available, we can then proceed and work with point clouds. The basic idea here is to construct the offset band directly from the point cloud and without the intermediate step of manifold reconstruction. This is addressed in §4 and §5 for noise-free points and manifold samples, and in §6 for points considered to be noisy samples of the manifold. In this case, we explicitly compute the probability that the constructed offset band contains the underlying manifold. As we expect, this probability is a function of the number of point samples, the noise level, the size of the offset, and the basic geometric characteristics of the underlying manifold. In the experimental section, §7, we also show the advantages of this framework over mesh-based distance computations.

To conclude this introduction, we should note that to the best of our knowledge, the only additional work explicitly addressing the computation of distance functions and geodesics for point clouds is the one reported in [4].³ The comparison of our framework with the one proposed by this paper is deferred to §8, where we also report the directions our research is taking.

2 Preliminary results and notation

In this section we briefly review the main results in [22], where the idea of approximating intrinsic distances and geodesics by extrinsic ones was first introduced.

We first introduce some basic notation that will be used throughout the article. For a compact and connected set $A \in \mathbb{R}^d$, $d_A(\cdot, \cdot)$ denotes the intrinsic distance between any two points of A , measured

¹Geodesics are the integral curves corresponding to the gradient directions of the intrinsic distance function, and are obtained back-propagating in this gradient direction from the target point to the source one.

²Tsitsiklis first described an optimal-control type of approach to solve the Hamilton-Jacobi equation, while independently Sethian and Helmsen both developed techniques based on upwind numerical schemes.

³In addition to studying the computation of distance functions on point clouds, [4, 30] address the important combination of this with multidimensional scaling for manifold analysis. Prior work on using geodesics and multidimensional scaling can be found in [27].

by paths constrained to be in A . Given a k -dimensional sub-manifold $\mathcal{M}$ of $\mathbb{R}^d$, $\Omega_{\mathcal{M}}^h$ denotes the set $\{x \in \mathbb{R}^d : d(\mathcal{M}, x) \leq h\}$. This is basically an h -offset of $\mathcal{M}$. To state that the sequence of functions $\{f_n(\cdot)\}_{n \in \mathbb{Z}^+}$ uniformly converges to $f(\cdot)$ as $n \uparrow \infty$, we frequently write $f_n \xrightarrow{n} f$. For a given event $\mathcal{E}$, $\mathbf{P}(\mathcal{E})$ stands for its probability of occurring. For a random variable (R.V. from now on) X , its mean value is denoted by $\mathbf{E}(X)$. For a function $f : \Omega \rightarrow \mathbb{R}$, and a subset A of Ω , $f|_A : A \rightarrow \mathbb{R}$ denotes the restriction of f to A .

In [22], we presented a new approach for the computation of weighted intrinsic distance functions over hyper-surfaces. We proved convergence theorems and also addressed the fast, computationally optimal, computation of such approximations. The key starting idea is that distance functions satisfy the (intrinsic) Eikonal equation, a particular case of the general class of Hamilton-Jacobi partial differential equations. Given $p \in \mathcal{S}$ (an hyper-surface in $\mathbb{R}^d$), we want to compute $d_{\mathcal{S}}(p, \cdot) : \mathcal{S} \rightarrow \mathbb{R}^+ \cup \{0\}$, the intrinsic distance function from every point on $\mathcal{S}$ to p . It is well known that the distance function $d_{\mathcal{S}}(p, \cdot)$ satisfies, in the viscosity sense, the equation

$$\begin{cases} \|\nabla_{\mathcal{S}} d_{\mathcal{S}}(p, x)\| = 1 \quad \forall x \in \mathcal{S} \\ d_{\mathcal{S}}(p, p) = 0 \end{cases}$$

where $\nabla_{\mathcal{S}}$ is the intrinsic differentiation (gradient). Instead of solving this intrinsic Eikonal equation on $\mathcal{S}$, we solve the corresponding extrinsic one in the offset band $\Omega_{\mathcal{S}}^h$:

$$\begin{cases} \|\nabla_x d_{\Omega_{\mathcal{S}}^h}(p, x)\| = 1 \quad \forall x \in \Omega_{\mathcal{S}}^h \\ d_{\Omega_{\mathcal{S}}^h}(p, p) = 0 \end{cases}$$

where $d_{\Omega_{\mathcal{S}}^h}(p, \cdot)$ is the Euclidean distance and therefore now the differentiation is the usual one. The main result of [22] is

Theorem 1 ([22]) *Let p and q be any two points on the smooth hyper-surface $\mathcal{S}$, then $|d_{\mathcal{S}}(p, q) - d_{\Omega_{\mathcal{S}}^h}(p, q)| \leq C\sqrt{h}$, for small enough h ,⁴ where C is a constant.*

This simplification of the intrinsic problem into an extrinsic one permits the use of the computationally optimal algorithms mentioned in the introduction. This makes computing intrinsic distances, and from them geodesics, as simple and computationally efficient as computing them in Euclidean spaces. Moreover, as detailed in [22], the approximation of the intrinsic distance $d_{\mathcal{S}}$ by the extrinsic/Euclidean one $d_{\Omega_{\mathcal{S}}^h}$ is never less accurate than the numerical error of these algorithms.

In [22], the result above was limited to co-dimension one closed hyper-surfaces, and the theory was applied to implicit surfaces, where computing the offset band is straightforward. It is the purpose of the present work to extend the aforementioned Theorem to deal with: (1) sub-manifolds of $\mathbb{R}^d$ of any codimension and with boundary, (2) sub-manifolds of $\mathbb{R}^d$ represented as points cloud, (3) random sampling of sub-manifolds of $\mathbb{R}^d$ in presence or absence of noise, and (4) convergence of geodesic curves in addition to distance functions. We should note that Theorem 1 holds even when the metric is not the one inherited from $\mathbb{R}^d$, obtaining weighted distance functions, see [22]. Although we will not present these new results in such generality, this is a simple extension that will be reported elsewhere.

3 General distance functions and geodesics approximation results

We now generalize Theorem 1 to cover open manifolds and general co-dimension. The convergence of the corresponding geodesic curves, which was not studied in [22], is obtained as well. As mentioned above,

⁴“Small enough h ” means that $h < 1/\max_i \kappa_i(\mathcal{S})$, where $\kappa_i(\mathcal{S})$ is the i -th principal curvature of $\mathcal{S}$. This guarantees having smoothness in $\partial\Omega_{\mathcal{S}}^h$, see [22].

proofs are omitted due to space constraints, and they can be found in [23].

Theorem 2 *Let $\mathcal{S}$ be a compact C^3 sub-manifold of $\mathbb{R}^d$ with smooth boundary $\partial\mathcal{S}$. We then have*

1. *Uniform convergence of the distance functions: $d_{\Omega_S^h}|_{\mathcal{S} \times \mathcal{S}}(\cdot, \cdot) \xrightarrow{h \downarrow 0} d_{\mathcal{S}}(\cdot, \cdot)$*
2. *Convergence of the geodesics: Let x and $y \in \mathcal{S}$ be joined by a unique minimizing geodesic $\gamma_{\mathcal{S}} : [0, 1] \rightarrow \mathcal{S}$ over $\mathcal{S}$, and let $\gamma_h : [0, 1] \rightarrow \Omega_S^h$ be the corresponding minimizing geodesic in the band Ω_S^h . Then $\gamma_h \xrightarrow{h \downarrow 0} \gamma_{\mathcal{S}}$.*

Note that the new setting is wider than the one in Theorem 1 from [22], since the co-dimension of the underlying manifold is not necessarily 1. This is very important for applications such as dimensionality reduction, where the dimension of the underlying manifold is unknown beforehand. The convergence of the distances is uniform following the same hypotheses as in Theorem 1, but we have forfeited the exact rate of convergence estimates. To obtain these rates for open manifolds of general co-dimension, further restrictions are imposed in $\partial\mathcal{S}$, see Theorem 3.

We can immediately obtain the following corollary which will become useful when we need to shape the Ω_S^h band in more convenient ways.

Corollary 1 *Let $\mathcal{S}$ and $\partial\mathcal{S}$ satisfy the hypotheses of Theorem 2. Let $\{\Sigma_i\}_{i \in \mathbb{N}}$ be a family of sets in $\mathbb{R}^d$ such that $\mathcal{S} \subseteq \Sigma_i \forall i \in \mathbb{N}$ and $d_{\mathcal{H}}(\Sigma_i, \mathcal{S}) \xrightarrow{i \uparrow + \infty} 0$ ($d_{\mathcal{H}}$ stands for the Hausdorff distance). Then $d_{\Sigma_i}|_{\mathcal{S} \times \mathcal{S}}(\cdot, \cdot) \xrightarrow{i \uparrow + \infty} d_{\mathcal{S}}(\cdot, \cdot)$.*

We now restrict the manifold in order to obtain explicit rates of convergence for general co-dimension open manifolds.

Definition 1 *We say that the manifold $\mathcal{S}$ with boundary $\partial\mathcal{S}$ is strictly convex if for every pair of points x and y in the interior $\overset{\circ}{\mathcal{S}}$ of $\mathcal{S}$, there exists a unique geodesic joining them that is entirely contained in $\overset{\circ}{\mathcal{S}}$.*

Theorem 3 (Rate of Convergence) *Under the hypotheses of Theorem 2, and assuming $\partial\mathcal{S}$ to be either void or strictly convex, we have that for small enough $h > 0$ and a constant C that does not depend on h , $|d_{\Omega_S^h}(x, y) - d_{\mathcal{S}}(x, y)| \leq C\sqrt{h}d_{\mathcal{S}}(x, y)$, and then $\max_{(x, y) \in \mathcal{S} \times \mathcal{S}} |d_{\Omega_S^h}(x, y) - d_{\mathcal{S}}(x, y)| \leq C\sqrt{h}$.*

Let's point out that we do not know yet whether the convexity condition for $\partial\mathcal{S}$ is really necessary in order to prove this rate of convergence, and advances on this precise issue will be reported elsewhere. Note that still, from Theorem 2, we have uniform convergence, even if the rate is not explicitly computed.

To summarize this section, the extrinsic Euclidean distance and geodesic computed in a thin offset band around the manifold $\mathcal{S}$ uniformly converges to the intrinsic ones on $\mathcal{S}$, for general co-dimensions and both closed and open manifolds. Under special cases, the rate of convergence can be explicitly computed. With these new results, we are now ready to proceed with manifolds given by point clouds.

4 Point clouds

Let $\mathcal{P}_n \triangleq \{p_1, \dots, p_n\}$ be points on the sub-manifold $\mathcal{S}$ and define $\Omega_{\mathcal{P}_n}^h \triangleq \bigcup_{i=1}^n B(p_i, h)$. Let also h and $\mathcal{P}_n$ be such that $\mathcal{S} \subseteq \Omega_{\mathcal{P}_n}^h$.⁵ Based on this, we then have $\mathcal{S} \subseteq \Omega_{\mathcal{P}_n}^h \subseteq \Omega_S^h$. Now the argument carries over easily, for if p and q are any two points in $\mathcal{S}$, one has $d_{\Omega_S^h}(p, q) \leq d_{\Omega_{\mathcal{P}_n}^h}(p, q) \leq d_{\mathcal{S}}(p, q)$, so $0 \leq d_{\mathcal{S}}(p, q) - d_{\Omega_{\mathcal{P}_n}^h}(p, q) \leq d_{\mathcal{S}}(p, q) - d_{\Omega_S^h}(p, q)$, and the rightmost quantity can be bounded by $C h^{1/2}$ (see [22] and §3). We therefore obtain the following basic results for point clouds:

⁵We should point out that results such as those in [1, 2, 3] provide basic theory to guarantee that the offset band $\Omega_{\mathcal{P}_n}^h$, while covering $\mathcal{S}$, has the correct topology.

Theorem 4 Assume that $\mathcal{S} \subseteq \Omega_{\mathcal{P}_n}^h$ then, under the hypotheses of Theorem 3, we have that for a constant C , $\max_{(p,q) \in \mathcal{S} \times \mathcal{S}} |d_{\mathcal{S}}(p, q) - d_{\Omega_{\mathcal{P}_n}^h}(p, q)| \leq C\sqrt{h}$.

Corollary 2 Let $\{h_n\}_{n \in \mathbb{N}}$ be such that $h_n \xrightarrow{n \uparrow +\infty} 0$ and $\mathcal{S} \subseteq \Omega_{\mathcal{P}_n}^{h_n}$ for all n , then $d_{\Omega_{\mathcal{P}_n}^{h_n}}|_{\mathcal{S} \times \mathcal{S}}$ converges uniformly to $d_{\mathcal{S}}$ as $n \uparrow +\infty$.

These basic results indicate that in order to compute distance functions, and from them geodesics, on manifolds represented by point clouds, all what needs to be done is to construct the union of balls $\Omega_{\mathcal{P}_n}^h$, and then inside this band in $\mathbb{R}^d$ use classical Euclidean algorithms. Note that the explicit reconstruction of the underlying manifold is not needed.

In general, different radii for each ball defining the set $\Omega_{\mathcal{P}_n}^h$ can be considered. In this case let $\mathcal{P}_n \mathcal{R}_n$ be given by $\{(p_1, r_1), \dots, (p_n, r_n)\}$, that is the set of pairs (p, r) where p are points sampled on the manifold. Define $\Omega_{\mathcal{P}_n}^{\mathcal{R}_n} \triangleq \bigcup_{i=1}^n B(p_i, r_i)$. Then, if we ensure that $\mathcal{S} \subseteq \Omega_{\mathcal{P}_n}^{\mathcal{R}_n}$ and that $R \triangleq \max_{1 \leq i \leq n} r_i$ is small enough, we can prove results similar to Theorem 4 and Corollary 2. In general, all what is needed is to guarantee that the union of balls includes the underlying manifold $\mathcal{S}$ (this coverage principle will be addressed in more detail in future sections). Moreover, we are not restricted to use just balls, and for example hyper-ellipsoids can be used as well to define the band. In particular, we can use hyper-ellipsoids that are adapted to the local geometry and density of the point cloud. This can be obtained by iterative local principal component analysis [8]. This scenario may be desirable for example when more resources are spent on areas of the manifold that are “harder” (high curvatures) than on those geometrically simpler parts (low curvatures).

5 Distance functions on randomly sampled manifolds

In the previous section, we dealt with a deterministic set of points in $\mathbb{R}^d$. We now consider a probabilistic framework, with points randomly sampled from a manifold $\mathcal{S}$ in $\mathbb{R}^d$. As one can guess from §4, we first have to study the crucial event $\{\mathcal{S} \subseteq \Omega_{\mathcal{P}_n}^h\}$, when we are given that the points in $\mathcal{P}_n$ are randomly sampled from $\mathcal{S}$. To illustrate our results, we assume a uniform distribution of the points over the manifold, $\mathcal{P}_n$ is an *i.i.d.* sequence of points uniformly sampled over $\mathcal{S}$.⁶ This means that for any $A \subseteq \mathcal{S}$, $\mathbf{P}(p_i \in A) = \frac{\mu(A)}{\mu(\mathcal{S})}$, for all $i \in \{1, 2, \dots, n\}$. Here, $\mu(A)$ denotes the volume measure of the subset A of the manifold $\mathcal{S}$. The results below can be extended for other sampling strategies.⁷

Theorem 5 Let $\mathcal{S}$ be a smooth compact k -dimensional sub-manifold of $\mathbb{R}^d$. Let $n \uparrow +\infty$ and $h_n \downarrow 0$ be such that both $\lim_n \frac{\ln h_n}{nh_n^k} = 0$ and $\lim_n nh_n^k = +\infty$ hold, then $\mathbf{P}\left(\mathcal{S} \subseteq \Omega_{\mathcal{P}_n}^{h_n}\right) \xrightarrow{n \uparrow +\infty} 1$.

Of independent interest is the following Random Sampling Asymptotic Theorem for sub-manifolds of $\mathbb{R}^d$, which can be proven using the preceding Theorem:⁸

Theorem 6 Under the hypotheses of Theorem 5, we have $\mathbf{p} \lim_n d_{\mathcal{H}}(\mathcal{S}, \Omega_{\mathcal{P}_n}^{h_n}) = 0$.

We now proceed with the convergence results for the corresponding distance functions. Being p and q points on $\mathcal{S}$, our main goal is to prove some kind of random convergence towards $d_{\mathcal{S}}(p, q)$ for the random

⁶We will also write $p_i \sim \mathbf{U}[\mathcal{S}]$.

⁷Note that we are considering a sequence of offsets that goes to 0. This is not necessary, but it is in the spirit of our work, since we need $h_n \rightarrow 0$ in order to achieve convergence of the distances.

⁸Here, for the sequence of R.V. X_n , and the R.V. X , $\mathbf{p} \lim_n X_n = X$ means that for any $\delta > 0$, $\lim_n \mathbf{P}(|X_n - X| > \delta) = 0$.

variables $d_{\Omega_{\mathcal{P}_n}^h}(p, q)$. Let's assume from now on that h (or h_n) is small enough. In the same spirit as in the deterministic case, it is meaningful to consider $\max_{(x,y) \in \mathcal{S} \times \mathcal{S}} |d_{\Omega_{\mathcal{P}_n}^h}(x, y) - d_{\mathcal{S}}(x, y)|$. However, to this end, we must be able to evaluate the function $d_{\Omega_{\mathcal{P}_n}^h}(\cdot, \cdot)$ at all points on $\mathcal{S} \times \mathcal{S}$. This can be guaranteed only if $\mathcal{S} \subseteq \Omega_{\mathcal{P}_n}^h$. We are then led to consider the event in the following Theorem:

Theorem 7 *Under the hypotheses of Theorem 5, and for $\varepsilon > 0$,*

$$\lim_{n \rightarrow \infty} \mathbf{P} \left(\left\{ \mathcal{S} \subseteq \Omega_{\mathcal{P}_n}^{h_n} \right\} \cap \left\{ \max_{(x,y) \in \mathcal{S} \times \mathcal{S}} |d_{\Omega_{\mathcal{P}_n}^{h_n}}(x, y) - d_{\mathcal{S}}(x, y)| \leq \varepsilon \right\} \right) = 1.$$

Details about how the intrinsic characteristics of the manifold play a role in the above convergence results are reported in [23].

To summarize, in this section we have addressed the distance convergence results for randomly sampled manifolds. We have presented results regarding the probability that the manifold of interest is covered by the union of balls centered at the random samples and that the error between the corresponding intrinsic and extrinsic distance functions is bounded. We now proceed to add noise to the samples.

6 Distance functions on noisily sampled manifolds

Since manifold samples are expected to be noisy, it is important to add this to the results presented above. Although a particular model of noise is assumed in this section, the results can be easily extended to other models.

Assume that $\{p_1, \dots, p_n\}$ is a set of *i.i.d.* random points such that each $p_i \sim \mathbf{U}[\Omega_{\mathcal{S}}^{\Delta}]$. The main point here is to notice that, whenever $\{\mathcal{S} \subseteq \Omega_{\mathcal{P}_n}^h\}$ holds, we have:

$$\max_{(x,y) \in \mathcal{S} \times \mathcal{S}} |d_{\Omega_{\mathcal{P}_n}^h}(x, y) - d_{\mathcal{S}}(x, y)| \leq C\sqrt{h + \Delta}. \quad (1)$$

In fact, as a consequence of the presence of noise in the points, we have that $\max_{p \in \mathcal{P}_n} d(\mathcal{S}, p) \leq h + \Delta$. Therefore, $\Omega_{\mathcal{P}_n}^h \subseteq \Omega_{\mathcal{S}}^{h+\Delta}$, and the claim follows from Theorem 4. That is, our bound regarding the absolute convergence of the distance in the cloud of points is worse than the one we can guarantee in absence of noise. Moreover, we have bounded this “worsening” using the parameter describing the noise. Up to first order, this worsening can be approximated by $\frac{C\Delta}{2\sqrt{h}}$.

To conclude, we wish to estimate the probability of having $\mathcal{S} \subseteq \bigcup_{i=1}^n B(p_i, h)$. It is easy to see that as a first “reality compliant” condition one should have $h > \Delta$. Starting from the setting for Theorem 7, we can prove the following:

Theorem 8 *Let $\mathcal{S}$ be a k -dimensional sub-manifold of $\mathbb{R}^d$ satisfying the usual hypotheses. Let $\Delta_n < h_n \rightarrow 0$ as $n \uparrow +\infty$, such that both $n\Delta_n^k \rightarrow \infty$ and $\frac{\ln \Delta_n}{n\Delta_n^k} \rightarrow 0$ as $n \uparrow +\infty$ hold, then for any $\varepsilon > 0$*

$$\lim_{n \rightarrow \infty} \mathbf{P} \left(\left\{ \mathcal{S} \subseteq \Omega_{\mathcal{P}_n}^{h_n} \right\} \cap \left\{ \max_{(x,y) \in \mathcal{S} \times \mathcal{S}} |d_{\Omega_{\mathcal{P}_n}^{h_n}}(x, y) - d_{\mathcal{S}}(x, y)| \leq \varepsilon \right\} \right) = 1.$$

7 Examples

We now present examples of distance functions and geodesics for point clouds, Figure 1; use these computations to find intrinsic Voronoi diagrams, Figure 2 (see also [17, 18]); and compare the results with those obtained with mesh-based techniques, Figure 3.⁹ The theoretical results presented in previous sections show

⁹All the figures in this paper are in color. VRML files corresponding to these examples can be found at mountains.ece.umn.edu/~guille/pc.htm.

that the intrinsic distance and geodesics can be approximated by the Euclidean ones computed in the band defined by the union of balls centered at the points of the cloud. The problem is then simplified to first computing this band, and then use computationally optimal techniques to compute the distances and geodesics inside this band, exactly as done in [22] for implicit surfaces. The band itself can be computed in several ways, and for the examples below we have used constant radii. Locally adaptive radii can be obtained as mentioned in §4, or can be based on diameters obtained for example from minimal spanning trees.

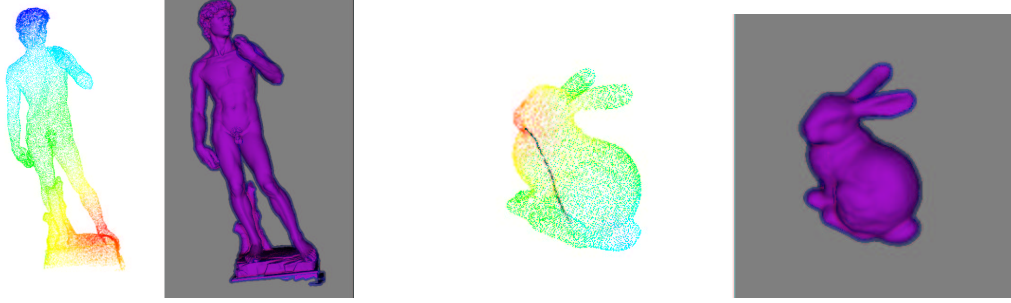


Figure 1: *Left: Intrinsic distance function for a point cloud. A point is selected in the head of the David, and the intrinsic distance is computed following the framework here introduced. The point cloud is colored according to their intrinsic distance to the selected point, going from dark blue (close) to bright red (far). The offset band, given by the union of balls, is shown next to the distance figure. Right: Same as before, with a geodesic curve between two selected points.*



Figure 2: *Voronoi diagram for point clouds. Four points (left) and two points (right) are selected on the cloud, and the point cloud is divided (colored) according to their geodesic distance to these four points. Note that this is a surface Voronoi, based on geodesics computed with our proposed framework, not an Euclidean one.*

We now make some very basic comparisons between our approach to geodesic distance computations and those based on graph approximations to the manifold, such as the one in Isomap [30].¹⁰ The goal is to show that such graph-based techniques are more sensitive to noise in the point cloud sample. This is expected, since the geodesic in such techniques goes through the noisy samples, while in our approach, they just go through the union of balls.

In the case of N noisy points (where the noise is as in the hypotheses of Section §6, with $\Delta \ll 1$), sampled on a rectilinear segment of length d_0 in $\mathbb{R}^2$, an explicit construction shows that the mean value error for the graph approximation d_G^Δ is approximately given by $\mathbf{E}(|d_G^\Delta - d_0|) \simeq \frac{\Delta^2}{3d_0} N^2$. On the other hand, for our approximation, d_h^Δ , if the segment is contained in the union of the balls centered at the sampling points,

¹⁰Isomap builds a mesh by locally connecting the (noisy) samples.

$\mathbf{E}(|d_h^\Delta - d_0|) = 0$. The probability of covering the segment by the band can be made arbitrarily close to 1 by increasing N . More precisely, one can prove that if p stands for the value of the probability of *not* covering the segment, then $p \leq k \frac{d_0}{\Delta} (1 - k' \frac{\Delta}{d_0})^N$, for some positive constants k and k' . We then note that for a fixed noise level Δ , by increasing N we actually worsen the graph approximation, whereas we are making our approximation better.

In the table below we present results of simulations carried out for the SwissRoll dataset [30], see Figure 3. We used 10000 points to define the manifold. We then generated 10000 noise vectors, being each component uniform with power one and zero mean. Then, we generated noisy datasets from the noiseless SwissRoll dataset by adding the noise vector times a constant n_k to each vector of the noiseless initial dataset. We then chose 1000 corresponding points in each dataset and computed the intrinsic pairwise distance approximation obtaining the matrices $\{(D_{ij}^{g,n_k})\}$ and $\{(D_{ij}^{h,n_k})\}$ for the graph-based and our approach respectively, where $k = 1, 2, \dots, 5$, $i, j \in [1, 1000]$, and n_k denotes the noise level. We then computed the values of $\max_{ij} |D_{ij}^{g,n_k} - D_{ij}^{g,0}|$ and $\max_{ij} |D_{ij}^{h,n_k} - D_{ij}^{h,0}|$ for each k , where $D_{ij}^{g,0}$ and $D_{ij}^{h,0}$ stand for noiseless intrinsic distance approximations. In the table, h indicates the radii and k the size of the neighborhood for Isomap. The graph approximation shows less robustness to noise than our method, as was argued above. This is also true for the sensitivity,¹¹ where our approach outperforms the graph-based one by at least one order of magnitude. Note that the sensitivity for our approach can be formally studied from Theorem 3.

Noise Power (n_k^2)	$\max_{ij} D_{ij}^{g,n_k} - D_{ij}^{g,0} $	k	$\max_{ij} D_{ij}^{h,n_k} - D_{ij}^{h,0} $	h
0.0001	2.5222	7	0.5266	1.8
0.01	4.6409	7	0.9430	1.8
0.04	5.1737	7	1.2489	1.8
0.09	5.3292	7	1.4682	1.8
0.16	5.4651	7	1.7965	1.8

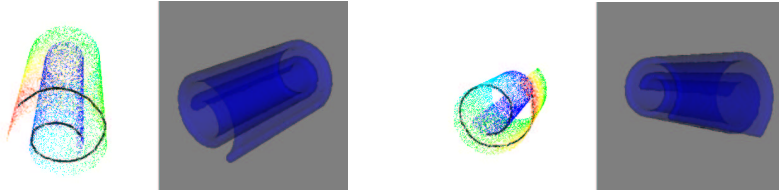


Figure 3: Surface used to compare graph-based approaches with our framework.

8 Concluding remarks

In this paper, we have shown how to compute distance functions and geodesics intrinsic to a generic manifold defined by a point cloud, without the intermediate step of manifold reconstruction. The basic idea is to use computationally optimal algorithms for computing Euclidean distances in an offset band surrounding the manifold, and use these to approximate the intrinsic distance. The underlying theoretical results were complemented by experimental illustrations.

As mentioned in the introduction, an alternative technique to compute geodesic distances was introduced in [4, 30]. In contrast with our work, the effects of noise were not addressed in [4]. Moreover, as we have explicitly shown, our framework is more robust to noise. We should also add that the theoretical results reported in [4] impose stronger restrictions on the manifolds than the ones we need. On the other hand, we

¹¹Sensitivity is defined as $\left|1 - \frac{\text{distance for noisy points}}{\text{distance for clean points}}\right|$.

should note that the memory requirements of our framework are larger, and this needs to be addressed for very high dimensions [23].

We are currently working on the use of this framework to create multiresolution representations of point clouds (in collaboration with N. Dyn, see also [6, 9, 10, 25]), to perform object recognition (see [11] for early work on the use of geodesic distances for this), and to compute basic geometric characteristics of the underlying manifold, all this of course without reconstructing the manifold. Applications of our framework for high dimensional data are also currently being addressed, and preliminary results are reported in [23].

Acknowledgments

We acknowledge useful conversations on the topic of this paper with L. Aspirot, P. Bermolén, R. Coifman, N. Dyn, O. Gil, R. Kimmel, and A. Pardo. We thank M. Levoy and the Digital Michelangelo Project for data provided for this project. This work was supported by a grant from the Office of Naval Research ONR-N00014-97-1-0509, the Presidential Early Career Award for Scientists and Engineers (PECASE), and a National Science Foundation CAREER Award.

References

- [1] N. Amenta, S. Choi, and R. Kolluri, “The power crust, unions of balls, and the medial axis transform,” *Computational Geometry: Theory and Applications* **19**, pp. 127-153, 2001.
- [2] N. Amenta, S. Choi, and R. Kolluri, “The power crust,” *Proceedings of 6th ACM Symposium on Solid Modeling*, pp. 249-260, 2001.
- [3] N. Amenta and R. Kolluri, “Accurate and efficient unions of balls,” *ACM Symposium on Computational Geometry*, pp. 119-128, 2000.
- [4] M. Bernstein, V. de Silva, J. Langford, and J. Tennenbaum, “Graph approximations to geodesics on embedded manifolds,” <http://isomap.stanford.edu/BdSLT.ps>
- [5] **Blitz++** website: www.oonumerics.org/blitz
- [6] J-D. Boissonnat and F. Cazals, “Coarse-to-fine surface simplification with geometric guarantees,” in A. Chalmers and T-M. Rhyne, Editors, *EUROGRAPHICS '01*, Manchester, 2001.
- [7] M. Botsch, A. Wiratanaya, and L. Kobbelt, “Efficient high quality rendering of point sampled geometry,” *EUROGRAPHICS Workshop on Rendering*, 2002.
- [8] R. Coifman, personal communication.
- [9] T. K. Dey, J. Giesen, and J. Hudson, “Decimating samples for mesh simplification,” *Proc. 13th Canadian Conference on Computational Geometry*, pp. 85-88, 2001.
- [10] N. Dyn, M. S. Floater, and A. Iske, “Adaptive thinning for bivariate scattered data,” *Journal of Computational and Applied Mathematics* **145(2)**, pp. 505-517, 2002.
- [11] A. Elad (Elbaz) and R. Kimmel, “Bending invariant representations for surfaces,” *In Proc. of CVPR'01*, Hawaii, Dec. 2001.
- [12] M. S. Floater and A. Iske, “Thinning algorithms for scattered data interpolation,” *BIT Numerical Mathematics* **38**, pp. 705-720, 1998.
- [13] M. Gross et al, *Point Based Computer Graphics*, EUROGRAPHICS Lecture Notes, 2002 (graphics.stanford.edu/~niloy/research/papers/ETH/PointBasedComputerGraphics_TutorialNotes.pdf).

- [14] J. Helmsen, E. G. Puckett, P. Collela, and M. Dorr, "Two new methods for simulating photolithography development in 3D," *Proc. SPIE Microlithography IX*, pp. 253, 1996.
- [15] P. W. Jones, "Rectifiable sets and the traveling salesman problem," *Invent. Math.* **102**, pp. 1-15, 1990.
- [16] R. Kimmel and J. A. Sethian, "Computing geodesic paths on manifolds," *Proc. National Academy of Sciences* **95:15**, pp. 8431-8435, 1998.
- [17] R. Kunze, F. E. Wolter, and T. Rausch, "Geodesic Voronoi diagrams on parametric surfaces," *IEEE*, 1997.
- [18] G. Leibon and D. Letscher, "Delaunay triangulations and Voronoi diagrams for Riemannian manifolds," *Computational Geometry 2000*, Hong Kong, 2000.
- [19] G. Lerman, "How to partition a low-dimensional data set into disjoint clusters of different geometric structure," preprint, 2000.
- [20] L. Linsen, "Point cloud representation," *CS Technical Report*, University of Karlsruhe, 2001.
- [21] L. Linsen and H. Prautzsch, "Local versus global triangulations," *EUROGRAPHICS '01*, 2001.
- [22] F. Mémoli and G. Sapiro, "Fast computation of weighted distance functions and geodesics on implicit hyper-surfaces," *Journal of Computational Physics* **173:2**, pp. 730-764, 2001.
- [23] F. Mémoli and G. Sapiro, "Distance functions and geodesics on surfaces and point clouds," January 2003 (mountains.ece.umn.edu/~guille/pc.htm).
- [24] M. Pauly and M. Gross, "Spectral processing of point-sampled geometry," *ACM SIGGRAPH*, pp. 379-386, 2001.
- [25] M. Pauly, M. Gross, and L. Kobbelt, "Efficient simplification of point-sampled surfaces," *IEEE Visualization 2002*.
- [26] S. Rusinkiewicz and M. Levoy, "QSplat: A multiresolution point rendering system for large meshes," *Computer Graphics (SIGGRAPH 2000 Proceedings)*, 2000.
- [27] E. Schwartz, A. Shaw, and E. Wolfson, "A numerical solution to the generalized mapmaker's problem: Flattening nonconvex polyhedral surfaces," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **11:9**, pp. 1005-1008, 1989.
- [28] J. Sethian, "Fast marching level set methods for three-dimensional photolithography development," *Proc. SPIE International Symposium on Microlithography*, Santa Clara, California, March, 1996.
- [29] J. A. Sethian, "A fast marching level-set method for monotonically advancing fronts," *Proc. Nat. Acad. Sci.* **93:4**, pp. 1591-1595, 1996.
- [30] J. B. Tenenbaum, V. de Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, pp. 2319-2323, December 2000.
- [31] R. Tsai, L.-T. Cheng, S. Osher, and H.-K. Zhao, "Fast sweeping algorithms for a class of Hamilton-Jacobi equations," *UCLA-CAM Report*, October 2001 (<http://www.math.ucla.edu/applied/cam/index.html>).
- [32] J. N. Tsitsiklis, "Efficient algorithms for globally optimal trajectories," *IEEE Transactions on Automatic Control* **40** pp. 1528-1538, 1995.
- [33] F. E. Wolter, "Cut loci in bordered and unbordered Riemannian manifolds," *Doctoral Dissertation, Technische Universität Berlin*, 1985.
- [34] M. Zwicker, M. Pauly, O. Knoll, and M. Gross, "PointShop 3D: An interactive system for point-based surface editing," *SIGGRAPH*, 2002.