



HAL
open science

Stable Medial-Axis estimation under regularity conditions

C Aaron

► **To cite this version:**

| C Aaron. Stable Medial-Axis estimation under regularity conditions. 2017. hal-01558392v1

HAL Id: hal-01558392

<https://hal.science/hal-01558392v1>

Preprint submitted on 7 Jul 2017 (v1), last revised 4 Jul 2019 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Stable Medial-Axis estimation under regularity conditions

C. Aaron

July 7, 2017

Abstract

The medial axis of a geometrical set $S \in \mathbb{R}^d$ carries a lot of information on the shape of S and is a popular tool in image analysis. The aim of this paper is to stably estimate the medial axis using only a finite sub-sample of S . We are going to prove that, under some regularity conditions, it can be done using only one pruning parameter. A general deterministic result will be given, then applied to three different cases. Firstly there is the usual one where the information is given by a finite sub-sample of ∂S . Secondly, there is the classical one where the set is given by its pixelated image. Finally, we can also apply the general theorem to the case where the information is given by a randomly drawn sample of S . In this more challenging case we also detail the algorithm, the choice of the parameter and give some simulation results as illustration.

Keywords: Manifold Learning, Geometric Inference, Skeleton.

1 Introduction

The medial axis of a set S was introduced in [4] as the set of points that has at least two different projections on ∂S . It was proposed as a tool for biological shape recognition and became popular in image compression and image analysis (see [12] for instance). More recently it has also been applied to wireless networks [20]. In section 3.3 we will see that it can also be applied in statistics for data analysis and regression (see [13] where this kind of application is detailed). In this paper we will mainly focus on the inner medial axis which is the intersection of the medial axis and S which is sufficient for S reconstruction and analysis.

There exist various definitions of the inner medial axis that may differ slightly. Following [21] or [14] we choose to define the medial axis of S as the set of the centers of the maximal balls included in S . More precisely, if $B(x, r)$ denotes the open ball centered in x and of radius r and if $\overset{\circ}{S}$ denotes the interior of S we consider:

$$\mathcal{M}(S) = \{x, \exists r(x) \text{ such that } B(x, r) \subset \overset{\circ}{S} \text{ and } B(x, r(x)) \subsetneq B(x', r') \Rightarrow B(x', r') \not\subset \overset{\circ}{S}\}. \quad (1)$$

Notice that the knowledge of $(x, r(x))$ for all $x \in \mathcal{M}$, known as the “medial axis transformation”, allows one to reconstruct S (see [21]). It is well known that the medial axis is not continuous with respect to the Hausdorff distance d_h (see [14] pages 217 – 238 for topological properties of the medial axis). One of the best illustrations of this phenomenon may be the following: suppose that $d = 2$ and $S = B(O, 1)$ (its medial axis being $\{O\}$) and that S_n is the interior of the convex hull of $\{x_1, \dots, x_n\}$ with $x_i = (\cos(2i\pi/n), \sin(2i\pi/n))$. We clearly have $d_h(S_n, S) \rightarrow 0$ while $\mathcal{M}(S_n) = \cup_i]0, x_i[$ and $d_h(\mathcal{M}(S), \mathcal{M}(S_n)) \rightarrow 1$. This phenomenon implies that estimating the medial axis using a finite sample of points $\mathcal{X}_n = \{X_1, \dots, X_n\}$ is unstable and provides a challenging problem that has been investigated in various papers (see [2] pages 109 – 125 for a

state-of-the-art report). Two different approaches have been investigated, the first one consists of pruning the medial axis of an estimation of S (see [18]); the second one consists of estimating the λ -medial axis instead of the medial axis (see [6],[8]). Our work aims at investigating the pruning option. To our knowledge there are three different ways of pruning the medial axis. The first one investigates how far the reconstructed set is from the initial one (see [5] or [7] for instance). This idea is difficult to exploit here because the original may be unknown, but we will use an idea based on this one when considering the practical problem of the choosing the parameter. The second idea is to prune according to stability parameter as in [3], in which the authors consider two pruning parameters. Recently, in [15], the authors proposed pruning according to a level of estimation of the boundary (to our knowledge this is only valid for $d = 2$). We are also going to prune the medial axis but according to a shape hypothesis (we will see at the end of the introduction that this pruning parameter is linked to the one proposed in [3]). To understand the idea note that if $x \in \mathcal{M}(S)$ then $x \in S$ and there exists $(y, z) \in \partial S$ and $\rho_x > 0$ such that:

$$\begin{cases} \|x - y\| = \|x - z\| = d(x, \partial S) \\ \|y - z\| \geq \rho_x \end{cases}$$

The idea is to restrict to sets \mathcal{S}_ρ smooth enough to have the existence of $\rho > 0$ such that, for all $S \in \mathcal{S}_\rho$ the medial axis $\mathcal{M}(S)$ is the set of points x in S such that there exists $(y, z) \in \mathbb{R}^d$ such that

$$\begin{cases} \|x - y\| = \|x - z\| = d(x, \partial S) \\ \|y - z\| \geq \rho \end{cases} .$$

To plug in an estimator of the medial axis in \mathcal{S}_ρ we first need to extract a subset $\mathcal{Y} = \{Y_1, \dots, Y_N\} \subset \mathcal{X}_n$ of observations “close” to ∂S then we build S_n “close” to S , and estimate the medial axis as follows:

$$\hat{\mathcal{M}}_{\rho, S_n}(\mathcal{Y}) = \{x \in \text{Vor}_{\mathcal{Y}}(y) \cap \text{Vor}_{\mathcal{Y}}(z) \cap S_n, (y, z) \in \mathcal{Y}^2, \|y - z\| \geq \rho\}, \quad (2)$$

where $\text{Vor}_{\mathcal{Y}}(y) = \{z, \|z - y\| \leq d(z, \mathcal{Y})\}$ is the Voronoi cell of y with respect to the set \mathcal{Y} .

Recall that $\hat{\mathcal{M}}_{0, \hat{S}}(\mathcal{Y})$ is just a plug-in estimator of the medial axis (see [16] for the link between Voronoi cells and medial axis). One of the most popular pruning methods (see [7]) consists of considering two “stability” parameters: $r_x = \|x - y_1\|$ and $\theta_x = \max_{i,j} \angle y_i x y_j$ where $\{y_1, \dots, y_k\} = \{Y_i \in \mathcal{Y}, x \in \text{Vor}(Y_i)\}$ and, since empirical observations show that small values for r_x or θ_x are related to points x in an “unstable” branch of the estimated medial axis; it has been proposed to keep only points with $r_x \geq R_0$ and $\theta_x \geq \theta_0$. Later, in [11], it has been proved that such a pruning method allows one to obtain a consistent medial axis estimation when we can access a finite subsample of ∂S . Notice that, even if the initial idea that led us to introduce our pruning parameter is different we have that $\rho_x = 2r_x \sin(\theta_x/2)$, so that our pruning parameter is strongly linked with the usual pruning method. This has two advantages: we have a theoretical proof of the consistence of our estimator in any dimension, and in practice, tuning one parameter is much easier than two.

Section 2 is dedicated to the theoretical study of our medial axis estimator. We give explicit regularity conditions and bound the error in a general (deterministic) theorem. In section 3 we derive different corollaries of the main theorem depending on assumptions on the sample. First, we give the convergence rate for a deterministic sample on ∂S (that is the classical hypothesis in discrete geometry). Then we present the image analysis point of view where the data is discretized on pixels, and, finally we give a way to estimate the medial axis when the points are

randomly drawn on S (to our knowledge this case had only been studied in [13] under stronger assumption and with a slower convergence rate).

Section 4 is dedicated to give more details in this last case. We describe the algorithm and the choice of the parameter, then finally illustrate our method with some simple numerical results.

2 Shape Hypothesis and main result

Throughout all the paper we consider subsets of \mathbb{R}^d . Let us first start with some definitions and notation. Let A be an open set and C a closed set such that $A \cap C = \emptyset$ and introduce $\mathcal{B}(A, C) = \{B(x, r), x \in A, B(x, r) \cap C = \emptyset\}$, $\mathcal{C}^{\max}(A, C)$ the set of the maximal balls (for the inclusion) of $\mathcal{B}(A, C)$ and $\overline{\mathcal{M}}(A, C)$ the set of the centers of the balls of $\mathcal{B}(A, C)$. In the following $S(x, r)$ denotes the sphere of radius r centered at x .

Let S be a compact set. As defined in (2) we have $\mathcal{M}(S) = \overline{\mathcal{M}}(\mathring{S}, \partial S)$. For any point $x \in \mathcal{M}(S)$ let us introduce $r(x) = d(x, \partial S)$, i.e. $B(x, r(x))$ is a maximal ball of $\mathcal{B}(\mathring{S}, \partial S)$. As noted in [14], the medial axis may have various “bad” properties, and we are going to restrict our study to regular enough sets, namely we will require that balls of radius r_0 roll freely inside and outside S and that S has a K -regular medial axis for some $K < 1$.

Definition 1. *Balls of radius r_0 roll outside and inside S if, for each $x \in \partial S$ there exists O_x^{out} and O_x^{in} such that $B(O_x^{out}, r_0) \subset S^c$ and $B(O_x^{in}, r_0) \subset S$.*

Definition 2. *S has a K -regular medial axis, $\mathcal{M}(S)$, if for all $(x, y) \in \mathcal{M}(S)^2$ we have: $|r(x) - r(y)| \leq K\|x - y\|$.*

Notice that if x and y are two different points of the medial axis such that $B(x, r(x))$ and $B(y, r(y))$ are the associated maximal balls, as $B(x, r(x)) \not\subseteq B(y, r(y))$ and $B(y, r(y)) \not\subseteq B(x, r(x))$ then $|r(x) - r(y)| < \|x - y\|$. Our K -regular condition for the medial axis is thus a bit stronger but, we believe, quite reasonable.

Now recall that the aim is to estimate the medial axis of a set S via an estimation based on a finite number of points. It will be seen later that, regarding the medial axis, the two parts of the Hausdorff distance don't have the same importance. This leads us to define (ε, h) -estimations as follows.

Definition 3. *Let S and \tilde{S} be two sets in \mathbb{R}^d . Then \tilde{S} is an (ε, h) -estimation of S if:*

$$\max_{y \in \tilde{S}} d(y, S) \leq \varepsilon \text{ and } \max_{x \in S} d(x, \tilde{S}) \leq h.$$

The behaviour of the estimated medial axis is made explicit in the following theorem. Note that the smoothness conditions on the boundary are close to the one used in [1] where a similar theorem is obtained for a sufficiently dense sample of the boundary. Nevertheless, here we can just observe points close to the boundary and we avoid the heavy computation of a weighted Delaunay tessellation.

Theorem 1. *Let S be a compact set such that balls of radius r_0 roll freely inside and outside S . Also assume that S has a K -regular medial axis. Let introduce $r_S = \text{diam}(S)/2$. Let α and ε be positive numbers such that $\alpha \leq \varepsilon \leq r_0/4$ and $\Delta_S = r_0^2 - \frac{r_S h^2}{r_0} - 4\varepsilon r_S - \frac{h^2}{2} > 0$. If $\mathcal{Y} = \{y_1, \dots, y_k\} \subset S$ is an (ε, h) -estimation of ∂S and if \hat{S} is a $(0, \alpha)$ -estimation of S , then*

there exist positive constants A, B, C, D, E, F such that, for all e such that:

$$\begin{cases} e \geq 2\varepsilon + \frac{h^2}{r_0} \\ \sqrt{Ah^2 + B\frac{\varepsilon}{e} + C\frac{h^2}{e} + D\frac{h^2\varepsilon}{e^2} + E\frac{h^4}{e^2} + F\frac{\varepsilon^3}{e^2}} \geq 2\sqrt{\left(\frac{h^2}{2r_0} + 2\varepsilon\right)^2 + h^2 + \left(\frac{h^2}{2r_0} + 2\varepsilon\right)\frac{h^2}{r_0}} \\ \sqrt{Ah^2 + B\frac{\varepsilon}{e} + C\frac{h^2}{e} + D\frac{h^2\varepsilon}{e^2} + E\frac{h^4}{e^2} + F\frac{\varepsilon^3}{e^2}} \leq \sqrt{\frac{(1-K)r_0^2\varepsilon}{2(1+K)(2r_0\varepsilon + 3h^2)}} \end{cases}$$

and for all ρ such that $\sqrt{Ah^2 + B\frac{\varepsilon}{e} + C\frac{h^2}{e} + D\frac{h^2\varepsilon}{e^2} + E\frac{h^4}{e^2} + F\frac{\varepsilon^3}{e^2}} \leq \rho \leq \sqrt{\frac{(1-K)r_0^2\varepsilon}{2(1+K)(2r_0\varepsilon + 3h^2)}}$ we have:

$$\hat{\mathcal{M}}_{\rho, \hat{S}}(\mathcal{Y}) \text{ is a } (e, A'\varepsilon + B'h^2)\text{-estimation of } \mathcal{M}.$$

Proof. First note that the inner rolling ball condition implies that for all $x \in \mathcal{M}$, $r(x) \geq r_0$.

For any $x \in \mathcal{M} \cap S$ we have $B(x, r(x) - 2\varepsilon) \subset B(x, r(x) - \varepsilon) \in \mathcal{B}(\hat{S}, \mathcal{Y})$. Indeed $x \in \hat{S}$ because $\alpha < r_0$ and $d(x, \mathcal{Y}) \geq r(x) - \varepsilon$ because \mathcal{Y} is a (ε, h) -estimation of ∂S . Thus there exist $x' \in \hat{S}$ and r' such that:

$$B(x, r(x) - 2\varepsilon) \subset B(x', r') \text{ with } B(x', r') \in \mathcal{C}^{\max}(\hat{S}, \mathcal{Y}).$$

We are now going to prove that for all (x', r') such that $B(x, r(x) - 2\varepsilon) \subset B(x', r')$ and $B(x', r') \in \mathcal{C}^{\max}(\hat{S}, \mathcal{Y})$ we have :

$$\|x - x'\| \leq \frac{1+K}{1-K}(2r_0\varepsilon + 3h^2). \quad (3)$$

Introduce x'^* a point of ∂S such that $d(x', \partial S) = \|x' - x'^*\| = r' - \gamma$ (γ may be negative). Then there exists a point $y_i \in \mathcal{Y}$ such that:

1. $\|y_i - x'^*\|^2 \leq h^2$, because \mathcal{Y} is a (ε, h) -estimation of ∂S ;
2. $y_i \notin \mathcal{B}(O_{x'^*}^{\text{out}}, r_0)$, because $y_i \in S$;
3. $\|x' - y_i\| \geq r'$, since otherwise $B(x', r')$ is not a ball of $\mathcal{B}(\hat{S}, \mathcal{Y})$.

Let us write $y_i = au + bw$ where $u = \frac{x'^* - x'}{\|x'^* - x'\|}$ and w is a unit vector of u^\perp . Notice that $a \geq 0$ and that w can be chosen such that $b \geq 0$ (see Figure 1).

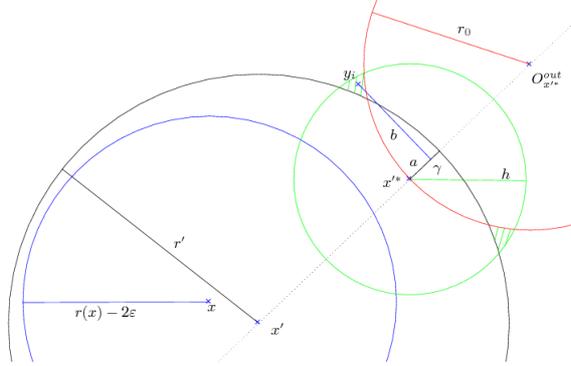


Figure 1:

We have:

$$\begin{cases} a^2 + b^2 \leq h^2 \\ b^2 + (r_0 - a)^2 \geq r_0^2 \\ \|x' - y_i\|^2 = (r' - \gamma + a)^2 + b^2 \end{cases} \Rightarrow \begin{cases} a^2 + b^2 \leq h^2 \\ 2ar_0 \leq h^2 \\ \|x' - y_i\|^2 \leq r'^2 + \gamma^2 - 2(r' + a)\varepsilon + h^2(1 + \frac{r'}{r_0}) \end{cases}$$

The last inequality gives $\gamma \leq r' + a - \sqrt{(r' + a)^2 - h^2(1 + \frac{r'}{r_0})}$ or $\gamma > r' + a + \sqrt{(r' + a)^2 - h^2(1 + \frac{r'}{r_0})}$. This second case is impossible because $\gamma \leq r'$ so we must have:

$$\gamma \leq r' + a - \sqrt{(r' + a)^2 - h^2(1 + \frac{r'}{r_0})} \leq \frac{h^2(1 + \frac{r'}{r_0})}{r' + a} \leq h^2 \left(\frac{1}{r'} + \frac{1}{r_0} \right).$$

Thus, observing now that $r' \geq r(x) - 2\varepsilon \geq r_0/2$, we have $\gamma \leq 3h^2/r_0$. This, in turn guarantees that $B(x', r' - 3h^2/r_0) \subset \mathcal{B}(\hat{S}, \partial S)$, so there exists $B(x'', r(x'')) \in \mathcal{C}^{\max}(\hat{S}, \partial S)$ with $B(x', r' - 3h^2/r_0) \subset B(x'', r(x''))$, that is:

$$\exists x'' \in \mathcal{M} \text{ such that } B(x', r' - 3h^2/r_0) \subset B(x'', r(x'')). \quad (4)$$

Now, as we have $B(x, r(x) - 2\varepsilon) \subset B(x', r')$ and (4) it follows that $B(x, r(x) - 2\varepsilon - 3h^2/r_0) \subset B(x'', r(x''))$. As a consequence, $r(x'') \geq \|x'' - x\| + r(x) - 2\varepsilon - 3h^2/r_0$. Thus $2\varepsilon + 3h^2/r_0 \geq \|x'' - x\| + r(x) - r(x'') \geq (1 - K)\|x - x''\|$ and we finally obtain:

$$\|x - x''\| \leq \frac{2r_0\varepsilon + 3h^2}{1 - K}. \quad (5)$$

Consider again $B(x, r(x) - 2\varepsilon) \subset B(x', r')$. Then $r' \geq \|x - x'\| + r(x) - 2\varepsilon$ and (4) give $r(x'') \geq \|x'' - x'\| + r' - 3h^2/r_0$. Summing the two inequalities (and finally using the K -regularity of the medial axis) we obtain:

$$2\varepsilon + 3h^2/r_0 \geq \|x' - x\| + \|x'' - x'\| + r(x'') - r(x) \geq \|x' - x\| + r(x'') - r(x) \geq (1 - K)\|x'' - x\|.$$

Thus, combining with (5) we get $\|x - x'\| \leq \frac{1+K}{1-K}(2r_0\varepsilon + 3h^2)$. This concludes the proof of (3).

Now we are going to prove that there exists $\bar{x}' \in \mathcal{M}(\hat{S}, \mathcal{Y}) \cap \mathcal{M} \sqrt{\frac{r_0^2(1-K)\varepsilon}{4(K+1)(2r_0\varepsilon+h^2)}} \cdot \hat{S}(\mathcal{Y})$ and r'_x such that $B(x, r(x) - 2\varepsilon) \subset B(\bar{x}', r'_x)$ with $B(\bar{x}', r'_x) \in \mathcal{C}^{\max}(\hat{S}, \mathcal{Y})$. More precisely we will show that this is realized for $r'_x = \max\{r', B(x, r(x) - 2\varepsilon) \subset B(x', r'), B(x', r') \in \mathcal{C}^{\max}(\hat{S}, \mathcal{Y})\}$ and \bar{x}' a point such that $B(x, r(x) - 2\varepsilon) \subset B(\bar{x}', r'_x)$ and $B(\bar{x}', r'_x) \in \mathcal{C}^{\max}(\hat{S}, \mathcal{Y})$. Clearly there exists $y \in \mathcal{S}(\bar{x}', r'_x) \cap \mathcal{Y}$. Suppose that $\max\{\|y - z\|, z \in \mathcal{S}(\bar{x}', r'_x) \cap \mathcal{Y}\} = l \leq r_0/2$. Introduce a point $z_0 \in \mathcal{S}(\bar{x}', r'_x) \cap B(y, l)$, $u = \frac{\bar{x}' - y}{\|\bar{x}' - y\|}$, $x'_t = \bar{x}' + tu$ and $r''_t = \|z_0 - x'_t\|$ (See Figure 2).

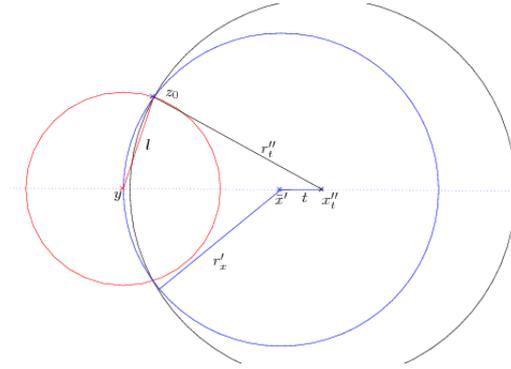


Figure 2:

Note first that $\langle \bar{x}' - z_0, u \rangle = \frac{2(r'_x)^2 - 2l^2}{2r'_x} \geq \frac{1}{2}$ (because $l^2 \leq r_0^2/4 \leq (r'_x)^2$). Thus:

1. For $t > 0$: $r_t'' = \|\bar{x}' - z_0 + tu\| = \sqrt{(r_x')^2 + t^2 + 2t\langle \bar{x}' - z_0, u \rangle} > r_x'$.
2. For $t > 0$: $B(x_t'', r_t'') \cap B^c(\bar{x}', r_x') \cap B(y, l) = \emptyset$, so $d(x_t'', \mathcal{Y} \cap B(y, l)) \geq r_t''$ and, because $d(\bar{x}', \mathcal{Y} \cap B^c(y, l)) > r_x'$, there exists t_1 such that, for all $0 < t < t_0$ we have $B(x_t'', r_t'') \cap \mathcal{Y} = \emptyset$.
Now because \hat{S} is open there exists t_0 such that, for all $t < t_0$ $x_t'' \in \hat{S}$ and $B(x_t'', r_t'') \cap \mathcal{Y} = \emptyset$.

As $B(\bar{x}', r_x')$ is a ball of $\mathcal{C}^{\max}(\hat{S}, \mathcal{Y})$ with maximal radius we have that for all $0 < t < t_0$, $B(x_t'', r_t'') \notin \mathcal{C}^{\max}(\hat{S}, \mathcal{Y})$ and, because of previous observations there exists $y_t \in B(x_t'', r_t'') \cap B^c(x, r(x) - 2\varepsilon)$. Now, because for $t > 0$,

$$B(\bar{x}', r_x') \cap \{z, \langle z - \bar{x}', u \rangle \geq \langle \bar{x}' - z_0, u \rangle\} \subset B(x_t'', r_t'') \cap \{z, \langle z - \bar{x}', u \rangle \geq \langle \bar{x}' - z_0, u \rangle\}$$

and because $B(x, r(x) - 2\varepsilon) \subset B(x', r_x')$ we have that for all $0 < t < t_0$ there exists:

$$y_t \in B(x_t'', r_t'') \cap B^c(x, r(x) - 2\varepsilon) \cap \{z, \langle z - \bar{x}', u \rangle \leq \langle \bar{x}' - z_0, u \rangle\}.$$

So, passing to the limit $t \rightarrow 0$ we obtain that there exists

$$y' \in \mathcal{S}(\bar{x}', r') \cap B^c(x, r(x) - 2\varepsilon) \cap \{z, \langle z - \bar{x}', u \rangle \leq \langle \bar{x}' - z_0, u \rangle\}.$$

This finally implies that there exists $y' \in \mathcal{S}(\bar{x}', r') \cap B^c(x, r(x) - 2\varepsilon)$ with $\|y' - y\| \leq l$. Introduce $\gamma = \|x - y\| - (r(x) - 2\varepsilon)$ (notice that $\gamma > \varepsilon$ because \mathcal{Y} is a (ε, h) -estimation of ∂S) and $\theta = \angle y\bar{x}'y'$ (see Figure 3).

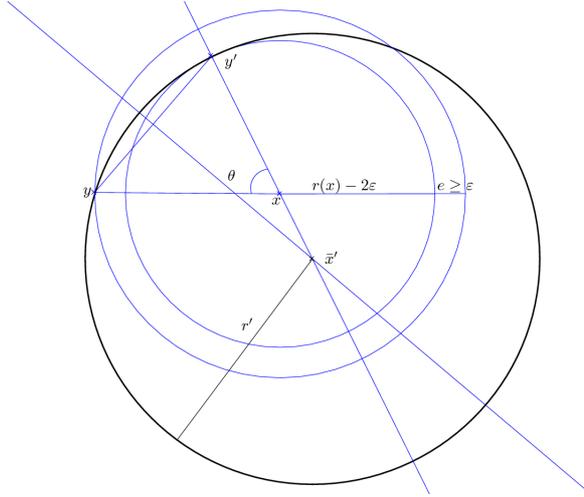


Figure 3:

By $\|y - y'\|^2 = \|y - x\|^2 + \|x - y'\|^2 + 2\langle y - x, x - y' \rangle$ it follows that $\|y - y'\|^2 = (r(x) - 2\varepsilon)^2 + (r(x) - 2\varepsilon + \gamma)^2 - 2(r(x) - 2\varepsilon)(r(x) - 2\varepsilon + \gamma) \cos(\theta)$. Thus we have:

$$\cos(\theta) = \frac{(r(x) - 2\varepsilon)^2 + (r(x) - 2\varepsilon + \gamma)^2 - \|y - y'\|^2}{2(r(x) - 2\varepsilon)(r(x) - 2\varepsilon + \gamma)}. \quad (6)$$

We also have $\bar{x}' - x = \|\bar{x}' - x\|(\cos(\theta), -\sin(\theta))$. Thus

$$\|\bar{x}' - y\|^2 = (r(x) - 2\varepsilon + \gamma + \|\bar{x}' - x\| \cos(\theta))^2 + (\|\bar{x}' - x\| \sin(\theta))^2. \quad (7)$$

But we also have $\|\bar{x}' - y\|^2 = \|\bar{x}' - y'\|^2 = (r')^2 = (r(x) - 2\varepsilon + \|\bar{x}' - x\|)^2$. Thus, with (7) we finally obtain:

$$(r(x) - 2\varepsilon + \gamma)^2 + 2\|\bar{x}' - x\|(r(x) - 2\varepsilon + \gamma) \cos(\theta) + \|\bar{x}' - x\|^2 = (r(x) - 2\varepsilon)^2 + 2\|\bar{x}' - x\|(r(x) - 2\varepsilon) + \|\bar{x}' - x\|^2$$

and thus :

$$\|\bar{x}' - x\| = \frac{1}{2} \frac{2\gamma(r(x) - 2\varepsilon) + \gamma^2}{(r(x) - 2\varepsilon) - (r(x) - 2\varepsilon + \gamma) \cos(\theta)}.$$

This and (6) gives

$$\|\bar{x}' - x\| = \frac{2\gamma(r(x) - 2\varepsilon)^2 + \gamma^2(r(x) - 2\varepsilon)}{\|y - y'\|^2 - 2\gamma(r(x) - 2\varepsilon) - \gamma^2}. \quad (8)$$

Finally, recall that $r(x) - 2\varepsilon \geq r_0/2$ and $\gamma \geq \varepsilon$, and observe that because of (5) we have $\|\bar{x}' - x\| \leq \frac{K+1}{1-K}(2r_0\varepsilon + 3h^2)$. Therefore, from (8) we obtain

$$\|y - y'\|^2 \geq \frac{(r_0\varepsilon + \varepsilon^2)(r_0/2 + \frac{K+1}{1-K}(2r_0\varepsilon + 3h^2))}{\frac{K+1}{1-K}(2r_0\varepsilon + 3h^2)} \geq \frac{(1-K)r_0^2\varepsilon}{2(1+K)(2r_0\varepsilon + 3h^2)}.$$

Introduce $C_0 = \min\{\sqrt{\frac{(1-K)r_0^2\varepsilon}{2(1+K)(2r_0\varepsilon + 3h^2)}}, \frac{r_0}{2}\}$. Recall that $\|y - y'\| \leq l$ so that $l = \max\{\|y - z\|, z \in S(\bar{x}', r'_x) \cap \mathcal{Y}\} \geq C_0$. Combining with (3) means that:

$$\text{For all } x \in \mathcal{M}, \text{ there exists } x' \in \hat{\mathcal{M}}_{C_0, \hat{S}} \text{ with } \|x - x'\| \leq \frac{1+K}{1-K}(2r_0\varepsilon + 3h^2). \quad (9)$$

Now take $x \in S$ and consider its distance to \mathcal{M} : $e = d(x, \mathcal{M}) > 0$. In this last part of the proof we put $l = d(x, \partial S)$. Consider now a point $x'^* \in \partial S$ such that $\|x - x'^*\| = l$. As $B(x, l) \subset \overset{\circ}{S}$ one can introduce $B(x', r(x'))$ a ball of $\mathcal{C}^{\max}(\overset{\circ}{S}, \partial S)$ containing $B(x, l)$. Recall that the regularity condition on S allows the existence of $O_{x'^*}^{\text{out}}$ such that $B(O_{x'^*}^{\text{out}}, r_0) \subset S^c$ and observe that x, x', x'^* and $O_{x'^*}^{\text{out}}$ are on the same line, and that $r(x') = \|x' - x\| + l$ with $e' = \|x' - x\| \geq e$.

Because \mathcal{Y} is a (εh) -estimation of ∂S there exists $y \in \mathcal{Y}$ such that $\|x'^* - y\| \leq h$. Obviously, as $d(y, \partial S) \leq \varepsilon$, we have $\|y - x'\| \geq r(x') - \varepsilon$ and since $y \in S$ we have $\|y - O_{x'^*}^{\text{out}}\| \geq r_0$. See Figure 4 for the construction.

Since $\|x'^* - y\| \leq h$ and $\|y - O_{x'^*}^{\text{out}}\| \geq r_0$, a short calculation shows that $\|x - y\|^2 \leq (r(x') - e')^2 + h^2 + (r(x') - e')h^2/r_0$. Thus, for all $y_i \in \mathcal{Y}$ such that $x \in \text{Vor}(y_i)$, we have $y_i \in B(x, \sqrt{(r(x') - e')^2 + h^2 + (r(x') - e')\frac{h^2}{r_0}}) \cap Bc(x', r(x') - \varepsilon) = E_x$. Writing $y_i = x + au + bw$, where $u = \frac{x - x'}{\|x - x'\|}$ and $w \in u^\perp$ we have

$$\begin{cases} b^2 + a^2 \leq (r(x') - e')^2 + h^2 + (r(x') - e')h^2/r_0 \\ b^2 + (e' + a)^2 \geq (r(x') - \varepsilon)^2 \end{cases}$$

After calculation we obtain that

$$\begin{cases} \frac{-e'^2 + e'(r(x') + \frac{h^2}{2r_0}) + \frac{\varepsilon^2}{2} - \varepsilon r(x') - \frac{h^2}{2} - \frac{r(x')h^2}{2r_0}}{e'} \leq a \leq l^2 + h^2 + lh^2/r_0 \\ b^2 \leq (r(x') - e')^2 + h^2 + (r(x') - e')h^2/r_0 - a^2. \end{cases}$$

Introduce $a_{\min}(e') = \frac{-e'^2 + e'(r(x') + \frac{h^2}{2r_0}) + \frac{\varepsilon^2}{2} - \varepsilon r(x') - \frac{h^2}{2} - \frac{r(x')h^2}{2r_0}}{2e'}$. Note that

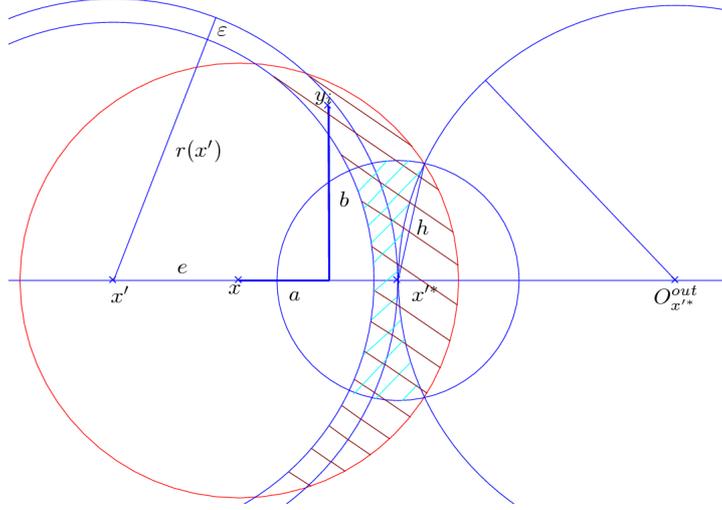


Figure 4: As there exists an observation in the blue crosshatched domain, all the y_i such that $x \in \text{Vor}_y(y_i)$ are in E_x , the brown crosshatched domain.

i. If $a_{\min}(e') \geq 0$ then $\text{diam}(E_x) = 2\sqrt{(r(x') - e')^2 + h^2 + (r(x') - e')h^2/r_0 - a_{\min}(e')^2}$

ii. For all e' : $\text{diam}(E_x) \leq 2\sqrt{(r(x') - e')^2 + h^2 + (r(x') - e')h^2/r_0}$

Introduce now $\Delta(x') = r(x')^2 - \frac{r(x')h^2}{r_0} - 4\epsilon r(x') - \frac{h^2}{2} + 2\epsilon^2 + \frac{h^4}{4r_0^2}$ and notice that $\Delta(x') \geq \Delta_S \geq 0$ and that $\sqrt{\Delta(x')} \geq r(x') - \frac{h^2}{r_0} - 4\epsilon - \frac{h^2}{2r(x')} \geq r(x') - \frac{3h^2}{2r_0} - 4\epsilon$ thus:

i. If $\frac{h^2}{r_0} + 2\epsilon \leq e' \leq r(x') - \frac{h^2}{2r_0} - 2\epsilon$ we have $a_{\min}(e') \geq 0$ and there exists A, B, C, D, E and

$$F \text{ such that: } \text{diam}(E_x) \leq \sqrt{Ah^2 + B\frac{\epsilon}{e'} + C\frac{h^2}{e'} + D\frac{h^2\epsilon}{e'^2} + E\frac{h^4}{e'^2} + F\frac{\epsilon^3}{e'^2}}$$

ii. If $e' \geq r(x') - \frac{h^2}{2r_0} - 2\epsilon$ then $\text{diam}(E_x) \leq 2\sqrt{(\frac{h^2}{2r_0} + 2\epsilon)^2 + h^2 + (\frac{h^2}{2r_0} + 2\epsilon)\frac{h^2}{r_0}}$

Take now e large enough to have

$$e \geq \frac{h^2}{r_0} + 2\epsilon \quad (10)$$

and small enough to have

$$2\sqrt{(\frac{h^2}{2r_0} + 2\epsilon)^2 + h^2 + (\frac{h^2}{2r_0} + 2\epsilon)\frac{h^2}{r_0}} \leq \sqrt{Ah^2 + B\frac{\epsilon}{e} + C\frac{h^2}{e} + D\frac{h^2\epsilon}{e^2} + E\frac{h^4}{e^2} + F\frac{\epsilon^3}{e^2}} \quad (11)$$

We have:

$$\hat{\mathcal{M}}_{\sqrt{Ah^2 + B\frac{\epsilon}{e} + C\frac{h^2}{e} + D\frac{h^2\epsilon}{e^2} + E\frac{h^4}{e^2} + F\frac{\epsilon^3}{e^2}}, \hat{s}(\mathcal{Y}) \subset \mathcal{M} \oplus eB$$

This, combined with (9) and the natural inclusion $\rho \leq \rho' \Rightarrow \hat{\mathcal{M}}_{\rho', \hat{s}}(\mathcal{Y}) \subset \hat{\mathcal{M}}_{\rho, \hat{s}}(\mathcal{Y})$, ensures that, since e satisfies the conditions (10), (11) and

$$\sqrt{Ah^2 + B\frac{\epsilon}{e} + C\frac{h^2}{e} + D\frac{h^2\epsilon}{e^2} + E\frac{h^4}{e^2} + F\frac{\epsilon^3}{e^2}} \leq \sqrt{\frac{(1-K)r_0^2\epsilon}{2(1+K)(2r_0\epsilon + 3h^2)}}, \quad (12)$$

then for any ρ such that $\sqrt{Ah^2 + B\frac{\varepsilon}{e} + C\frac{h^2}{e} + D\frac{h^2\varepsilon}{e^2} + E\frac{h^4}{e^2} + F\frac{\varepsilon^3}{e^2}} \leq \rho \leq \sqrt{\frac{(1-K)r_0^2\varepsilon}{2(1+K)(2r_0\varepsilon + 3h^2)}}$ we have: $\hat{\mathcal{M}}_{\rho, \hat{S}}$ is a $\left(e, \frac{1+K}{1-K}(2r_0\varepsilon + 3h^2)\right)$ -estimation of \mathcal{M} . This concludes the proof. \square

3 Applications for different models

Now we can apply Theorem 1 under different sample hypotheses. More precisely we give convergence rates of the estimated medial axis when the information is given at n points. In the first case we have n points located on the boundary, in the second we consider an n -pixel size image and in the third case we consider a sample of n points randomly drawn on S . The two first ones are classical in medial axis estimation but the last one, illustrated with simulations in the following section, is more challenging.

3.1 Deterministic sample of ∂S

Usually the medial axis estimation is done under the following model: S is “known” and we want to, computationally plot the medial axis using a finite number of points located on ∂S . Notice that the rolling balls condition implies that ∂S is a \mathcal{C}_1^1 $(d-1)$ -dimensional manifold (see [19]). Therefore there exists λ_S such that for all n there exists $\{x_1, \dots, x_n\} \subset \partial S$ such that $\partial S \subset \bigcup_i B\left(x_i, \lambda_S n^{-\frac{1}{d-1}}\right)$. We are considering in this first corollary this kind of sets of points.

Corollary 1. *Assume that there exists λ such that \mathcal{X}_n is a $(0, \lambda n^{-1/(d-1)})$ -estimation of ∂S , the boundary of a compact set S that has a K -regular medial axis and such that balls of radius r_0 roll inside and outside S . There exist constants ρ_0 and ρ_1 such that, for all $\rho_0 < \rho < \rho_1$ there exists C_ρ such that, for n large enough:*

$$d_h(\mathcal{M}, \hat{\mathcal{M}}_{\rho, S}(\mathcal{X}_n)) \leq C_\rho n^{-\frac{2}{d-1}}.$$

Proof. We are under the hypotheses of Theorem 1 with $\alpha_n = 0$ and $h_n = \lambda n^{-1/(d-1)}$ and $\varepsilon_n = \frac{\lambda^2}{r_0} n^{-2/(d-1)}$ (indeed, as \mathcal{X}_n is a $(0, h_n)$ -estimation of ∂S it is also a (ε_n, h_n) -estimation of ∂S). Chose $e_n = e_0 n^{-2/(d+1)}$ with e_0 large enough to have

$$e_0 \geq \frac{2\lambda^2 + 1}{r_0} \text{ and } \sqrt{B\frac{\lambda^2}{r_0 e_0} + C\frac{\lambda^2}{e_0} + D\frac{\lambda^4}{r_0 e_0^2} + E\frac{\lambda^4}{e_0^2}} < \sqrt{\frac{(1-K)r_0}{10(1+K)}}.$$

Observe that for n large enough all the conditions of Theorem 1 are satisfied and for all ρ with $\sqrt{B\frac{\lambda^2}{r_0 e_0} + C\frac{\lambda^2}{e_0} + D\frac{\lambda^4}{r_0 e_0^2} + E\frac{\lambda^4}{e_0^2}} < \rho < \sqrt{\frac{(1-K)r_0}{10(1+K)}}$ we have:

$$d_h(\mathcal{M}, \hat{\mathcal{M}}_{\rho, S}(\mathcal{X}_n)) \leq \max\left(e_0, \left(\frac{A'}{r_0} + B'\right)\right) \lambda^2 n^{-\frac{2}{d-1}}.$$

\square

3.2 Pixelization

In this subsection only we consider $S \subset [0, 1]^d$ and $[0, 1]^d$ is divided into n “cubic” pixels of size $c_n = n^{-1/d}$ (here, n cannot be an arbitrary integer since we are constrained to have $n = p^d$, but we choose to express the results with regard to this parameter n that represents the amount of information given). As in the previous application, we consider here that S is known then pixelized (that is, \hat{S}_n is a union of pixels) to draw the medial axis. The pixelization method can be one of the following

1. Gaussian: a pixel belongs to \hat{S}_n if its center belongs to S ;
2. Random Type 1: if a pixel is included in S then it is in \hat{S}_n ; if a pixel is included in S^c then it is in \hat{S}_n^c ; otherwise randomly choose between \hat{S}_n or \hat{S}_n^c with a probability 0.5;
3. Random Type 2: a pixel belongs to \hat{S}_n with a probability proportional to the volume of the intersection between S and the pixel...

Let us say that a pixel is a boundary pixel (a pixel in \mathcal{Y}) if it is in \hat{S}_n and has a neighbor in \hat{S}_n^c or if it is in \hat{S}_n and has a neighbor in \hat{S}_n^c . Observe that for this proposed method we have the existence of λ such that $\varepsilon_n = h_n = \lambda c_n$. Thus, by Theorem 1, there exist suitable values for ρ such that $d_h(\mathcal{M}, \hat{\mathcal{M}}_{\rho, \hat{S}_n}(\mathcal{Y})) = O(n^{-\frac{2}{d}})$.

3.3 Sample randomly drawn on S

In the introduction we claimed that the medial axis may be an important tool for statistics, data analysis and regression. In [13] a regression model that can be solved with the medial axis is detailed but we are convinced that this object has other possible applications in statistics. Indeed suppose that the data (in \mathbb{R}^d) are drawn on M a d' -dimensional sub-manifold of \mathbb{R}^d but that, due to a noise, we can only access to observations on $M \oplus rB$, when r (the noise amplitude) is smaller than ρ_0 , the reach of M . Then M is the medial axis of $M \oplus rB$ and the estimation of the medial axis allows us to obtain information of the unknown M . In this case the medial axis is exactly M and it can be estimated. We can also apply a dimension estimation method on \hat{M} to estimate the dimension of M which is an important information on the number of “indendant non-linear components”. Finally, let us observe that the mean distance to M divided by the mean distance in M can be interpreted as a non-linear correlation coefficient.

In the case where the points are randomly drawn on S satisfying the regularity conditions of Theorem 1 and assuming additional hypothesis on the probability distribution, the two following Corollaries shows two ways to practically estimate the inner medial axis. For this we use $C_r(\mathcal{X}_n)$ the r -convex hull estimator of S (see [17]) to estimate S and to identify the subset \mathcal{Y} of sample points located close to the boundary.

Corollary 2. *Let $\mathcal{X}_n = \{X_1 \dots X_n\}$ be an iid sample of points, drawn on a compact set S that has a K -regular medial axis and such that balls of radius r_0 roll inside and outside S . Assume that the density f of the sample satisfies $f(x) \geq f_0 > 0$. For all $x \in S$. for all $r < r_0$ denote by $\hat{S}_r(\mathcal{X}_n)$ the r -convex hull of \mathcal{X}_n and put $\mathcal{Y}_r(\mathcal{X}_n) = \partial \hat{S}_r(\mathcal{X}_n) \cap \mathcal{X}_n$. There exist constants ρ_0 and ρ_1 such that, for all $\rho_0 < \rho < \rho_1$ there exists C_ρ such that*

$$d_h(\mathcal{M}, \hat{\mathcal{M}}_{\rho, \hat{S}_r(\mathcal{X}_n)}(\mathcal{Y}_r(\mathcal{X}_n))) \leq C_\rho \left(\frac{\ln n}{n} \right)^{\frac{2}{d+1}} \quad e.a.s.$$

Proof. We refer the reader to [17] to be convinced that there exists constants a_r, b_r and c_r such that, eventually almost surely:

1. $\hat{S}_r(\mathcal{X}_n)$ is (eventually almost surely) a $(0, a_r(\ln n/n)^{\frac{2}{d+1}})$ -estimation of S
2. $\mathcal{Y}_r(\mathcal{X}_n)$ is (eventually almost surely) a $(a_r(\ln n/n)^{\frac{2}{d+1}}, b_r(\ln n/n)^{\frac{1}{d+1}})$ -estimation of S .

Now taking n large enough to have $a_r(\ln n/n)^{\frac{2}{d+1}} \leq r_0/2$, we are under the hypothesis of Theorem 1. Chose $e_n = e_0(\ln n/n)^{2/d+1}$ and introduce $\lambda_0(e_0) = \frac{Ba_r}{e_0} + \frac{Cb_r^2}{e_0} + \frac{Db_r^2 a_r}{e_0^2} + \frac{Eb_r^4}{e_0^2}$, $\lambda_1 = \frac{(1-K)r_0^2 a_r}{2(1+K)(2r_0 a_r + 3b_r^2)}$, $\lambda_2 = \frac{1+K}{1-K}(2r_0 a_r + 3b_r^2)$ and $\mu_0 = \frac{r_0^2 a_r}{2\lambda_2}$. Take e_0 large enough to have

$\lambda_0(e_0) < \lambda_1$ so as to ensure that, for n large enough, Conditions (10),(11) and (12) are satisfied. For all ρ such that $\sqrt{\lambda_0} < \rho < \sqrt{\lambda_1}$ we have that $\hat{\mathcal{M}}_{\rho, \hat{S}}$ is a $(e_0(\ln n/n)^{2/(d+1)}, \lambda_2(\ln n/n)^{2/(d+1)})$ -estimation of \mathcal{M} for n large enough. That concludes the proof \square

From a computational point of view it is easier to restrict the estimated medial axis to a Devroye-Wise (that is, only $\cup B(X_i, r_n)$; see [10]) estimator of the support instead of to the r -convex hull. The following corollary establishes a result on this.

Corollary 3. *Let $\mathcal{X}_n = \{X_1 \dots X_n\}$ be an iid sample of points, drawn on a compact set S that has a K -regular medial axis and such that balls of radius r_0 roll inside and outside S . Assume that the density f of the sample satisfies $f(x) \geq f_0 > 0$ for all $x \in S$. For all $r < r_0$ denote by $\hat{S}_r(\mathcal{X}_n)$ the r -convex hull of \mathcal{X}_n and put $\mathcal{Y}_r(\mathcal{X}_n) = \partial \hat{S}_r(\mathcal{X}_n) \cap \mathcal{X}_n$. Let ν_n be a sequence such that $\nu_n \rightarrow 0$ and $S \subset \bigcup_i \bar{B}(X_i, \nu_n)$. Then, for any given ρ :*

$$\hat{\mathcal{M}}_{\rho, \hat{S}_r(\mathcal{X}_n)}(\mathcal{Y}_r(\mathcal{X}_n)) = \hat{\mathcal{M}}_{\rho, \cup_i \bar{B}(X_i, \nu_n)}(\mathcal{Y}_r(\mathcal{X}_n)) \text{ e.a.s.}$$

Proof. With the condition on ν_n we clearly have $\hat{\mathcal{M}}_{\rho, \hat{S}_r(\mathcal{X}_n)}(\mathcal{Y}_r(\mathcal{X}_n)) \subset \hat{\mathcal{M}}_{\rho, \cup_i \bar{B}(X_i, \nu_n)}(\mathcal{Y}_r(\mathcal{X}_n))$. Let us now prove the reverse inclusion by contradiction. Introduce $\nu'_n = \max(\nu_n, d_h(\hat{S}_r(\mathcal{X}_n), S))$ and consider a point in $x \in \hat{\mathcal{M}}_{\rho, \cup_i \bar{B}(X_i, \nu_n)}(\mathcal{Y}_r(\mathcal{X}_n)) \setminus \hat{\mathcal{M}}_{\rho, \hat{S}_r(\mathcal{X}_n)}(\mathcal{Y}_r(\mathcal{X}_n))$ then $x \in \cup_i \bar{B}(X_i, \nu_n) \setminus \hat{S}_r(\mathcal{X}_n)$. The regularity condition then implies that $x \in \partial S \oplus \nu'_n B$ and there exists $x^* \in \partial S$ with $\|x - x^*\| \leq \nu'_n$. Now, eventually almost surely there exists a point $Y_i \in \mathcal{Y}_r$ such that $\|Y_i - x^*\| \leq b_r(\ln n/n)^{1/(d+1)}$ and $\|x - Y_i\| \leq b_r(\ln n/n)^{1/(d+1)} + \nu'_n$ so that if $x \in \text{Vor}(Y_i) \cap \text{Vor}(Y_j)$ then Y_i and Y_j are in $B(x, b_r(\ln n/n)^{1/(d+1)} + \nu'_n)$. Thus $\|Y_i - Y_j\| \leq 2, b_r(\ln n/n)^{1/(d+1)} + \nu'_n$. But, as $x \in \hat{\mathcal{M}}_{\rho, \cup_i \bar{B}(X_i, \nu_n)}(\mathcal{Y}_r(\mathcal{X}_n))$, we have the existence of i and j such that $x \in \text{Vor}(Y_i) \cap \text{Vor}(Y_j)$ and $\|Y_i - Y_j\| \leq \rho$ this is impossible for n large enough. \square

4 Simulations

4.1 Algorithm

Step 1: Detection of boundary points. First we have to identify the set $\mathcal{Y}_r(\mathcal{X}_n) = \partial \hat{S}_r(\mathcal{X}_n) \cap \mathcal{X}_n$. This is an easy task since:

Proposition. $X_i \in \mathcal{Y}_r(\mathcal{X}_n) \Leftrightarrow \max\{\|y - X_i\|, y \in (\text{Vor}_{\mathcal{X}_n}(X_i))\} \geq r$.

Proof. Observe that $\max\{\|y - X_i\|, y \in (\text{Vor}_{\mathcal{X}_n}(X_i))\} \geq r$. This implies that there exists x such that $B(x, r) \cap \mathcal{X}_n = \emptyset$. Thus, by definition of S_r ; $B(x, r) \cap S_r(\mathcal{X}_n) = \emptyset$ and $x_n = (1/n)x + (r - 1/n)X_i \rightarrow X_i \in S_r^c(\mathcal{X}_n)$ with $X_i \in S_r(\mathcal{X}_n)$, so $X_i \in \partial S_r(\mathcal{X}_n)$. Conversely, if $X_i \in \partial S_r(\mathcal{X}_n)$ there exist two sequences x_n and $y_n \in S_r^c(\mathcal{X}_n)$ such that $x_n \rightarrow X_i$, $x_n \in B(y_n, r)$ and $B(y_n, r) \cap \mathcal{X}_n = \emptyset$. $y_n \in S \oplus rB$ which is compact. Thus, after an extraction we can suppose that $y_n \rightarrow y$. As $r < \|y_n - X_i\| \leq r + \|x_n - X_i\|$ we have at the limit $\|y - X_i\| \geq r$. Moreover, since for all n , $B(y_n, r) \cap \mathcal{X}_n = \emptyset$, we have $B(y, r) \cap \mathcal{X}_n = \emptyset$ and therefore $y \in \text{Vor}_{\mathcal{X}_n}(X_i)$. \square

Step 2: Choice of ν_n and computation of some points of $\hat{\mathcal{M}}_{\rho, \cup_i \bar{B}(X_i, \nu_n)}(\mathcal{Y}_r(\mathcal{X}_n))$ and attribution of a structure. We propose to choose $\nu_n = 2 \max_i(\max_j(\|X_i - X_j\|))$ as suggested in [9]. Now one can compute \mathcal{Y}_r and then obtain

$$\mathcal{Z}_{\rho, \nu_n} = \{z \in \bigcap_{i_1 < i_2 < \dots < i_{d+1}} \text{Vor}_{\mathcal{Y}_r}(Y_{i_j}), \min_i \|z - X_i\| \leq \nu_n, \exists(k, l) \|Y_{i_k} - Y_{i_l}\| \geq \rho\}$$

This is a clearly a finite subset $\{z_1, \dots, z_{N_\rho}\}$ of $\hat{\mathcal{M}}_{\rho, \cup_i \bar{B}(X_i, \nu_n)}(\mathcal{Y}_r(\mathcal{X}_n))$ on which we propose to compute the graph G_ρ such that $[z_i, z_j]$ is an edge of G_ρ if there exists $Y_k \in \mathcal{Y}_r$ such that z_i and z_j belongs to $\text{Vory}(Y_k)$.

4.2 Choice of the parameter

We propose to choose the parameter by considering two functions of ρ . The first, named $\text{dis}(\rho)$ represents a distance from \mathcal{X}_n to the reconstructed support based on the medial axis computed with the parameter ρ . More precisely we compute: $\text{dis}_1(\rho) = \frac{1}{n} \sum_i 1_{X_i \in \tilde{S}_\rho}$ where $\tilde{S}_\rho = \cup_{z \in \mathcal{Z}_{\rho, \nu_n}} \bar{B}(z, d(z, \mathcal{Y}))$. We also compute $\text{dis}_2(\rho) = \frac{1}{n} \sum_i d(X_i, \tilde{S}_\rho)$. Then dis_1 is the percentage of observations that are not in \tilde{S}_ρ while dis_2 is $d(X, \tilde{S}_\rho)$. A “good” value for ρ is a value such that $\text{dis}(\rho)$ is large enough to reconstruct correctly the set.

We also expect the medial axis to be regular and we propose to measure the regularity with the function $\text{ncc}(\rho)$ that is the number of connected components of G_ρ . A suitable value for ρ should be a large enough value such that $\text{ncc}(\rho)$ and $\text{dis}(\rho)$ are small enough.

4.3 Some results

The 2–dimensional case. Since the medial axis was originally proposed for biological image analysis we propose to first test our method on “biological” image. We present some results on a leaf image and on a sub-image of da Vinci’s Vitruvian Man (see Figure 5 which presents the original image and the “working” version in black and white). We then test the proposed

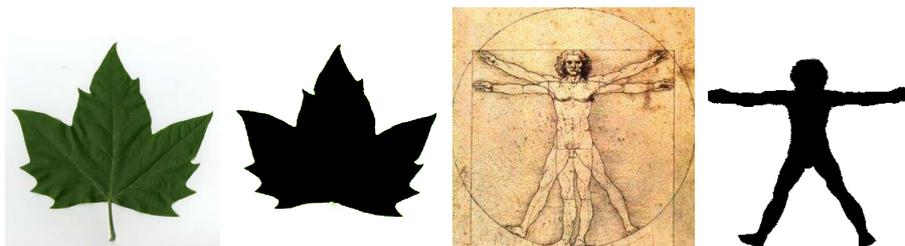


Figure 5: Original image and there “working” version

algorithm and choice for the parameters; see results in Figure 6. First notice that the choice of a suitable parameter is easier in the case of the Vitruvian Man (where the ncc function brings some information).

The 3–dimensional case. In this section we draw points on $M \oplus rB$ and observe how M is recognized via this medial axis approach. We propose two different M : first a Moebius strip then a trefoil knot. In figure 7 we present the simulation result for $n = 10000$ points. As previously, we show the evolution of the functions ncc , dis and the choice of the parameter. Then we present the observations (the sample on $M \oplus rB$ (here $r = 0.3$) on which our algorithm is applied). Finally we present the original sample on M (the unknown manifold we want to estimate) and the graph G_ρ .

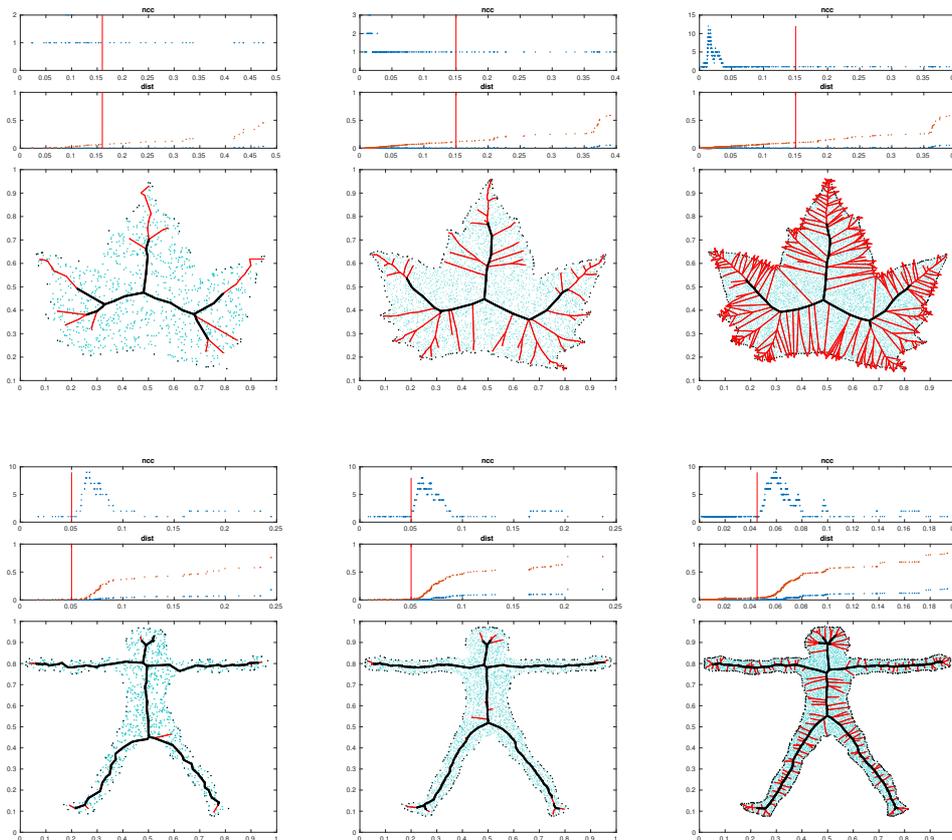


Figure 6: Estimated medial axis, sample sizes $n = 1000$, $n = 5000$ and $n = 20000$. Blue dots the sample, black dots \mathcal{Y} , thin red graph is G_0 and bold black graph is G_ρ , a suitable value for ρ being chosen according to ncc and dis functions (plotted above) and represented by the red line

References

- [1] N. Amenta, S. Choi, and R. Kolluri. The power crust, unions of balls, and the medial axis transform. *Computational Geometry*, 19:127–153, 2001.
- [2] D. Attali, J. Boissonnat, and E. Edelsbrunner. *Mathematical Foundations of Scientific Visualization, Computer Graphics, and Massive Data Exploration*. Spr, 2009.
- [3] D. Attali and A. Montanvert. Modeling noise for a better simplification of skeletons. In *Proc. Internat. Conf. Image Process.*, 1996.
- [4] H. Blum. A transformation for extracting new descriptors of shape. *Models for the Perception of Speech and Visual Form*, Issue 5:362–380, 1967.
- [5] J. W. Brandt and V. R. Algazi. Continuous skeleton computation by voronoi diagram. *CVGIP: Image Understanding*, 1992.

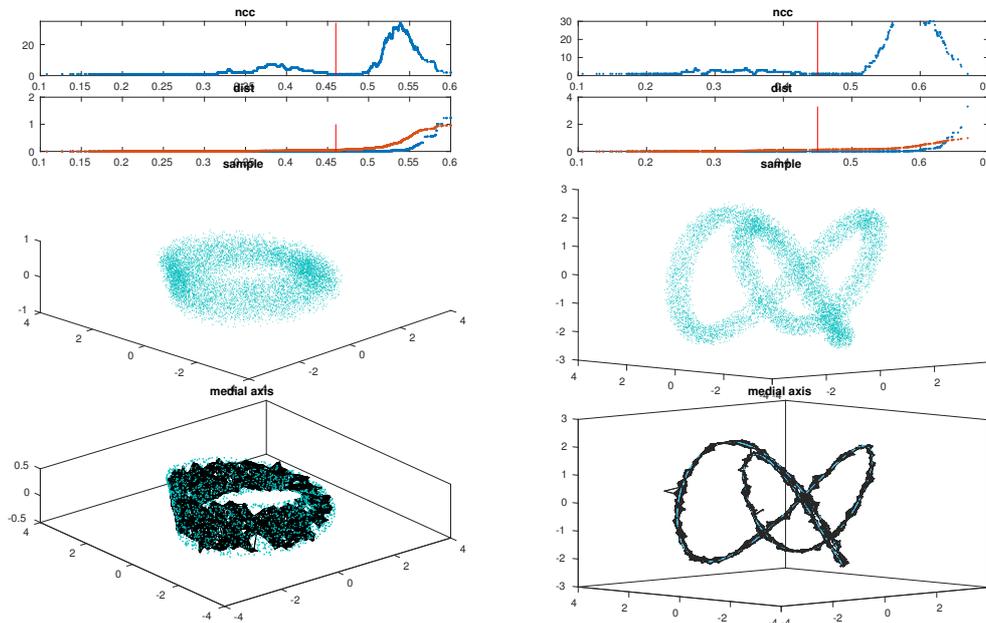


Figure 7: Estimated medial axis for a noised Moebus-strip and a noised trefoil knot, sample sizes $n = 10000$.

- [6] F. Chazal and A. Lieuter. The " λ -medial axis". *Graphical Models*, 67(4):304–331, July 2005.
- [7] S. W. Choi and H.-P. Seidel. Linear one-sided stability of mat for weakly injective 3d domain. In *Proc. 7th ACM Sympos. Solid Modeling Appl.*, pages 344–355, 2002.
- [8] A. Cuevas, P. Llop, and B. Pateiro-Lopez. On the estimation of the medial axis and inner parallel body. *Journal of Multivariate Analysis*, 129:171–185, August 2014.
- [9] A. Cuevas and A. Rodriguez-Casal. On boundary estimation. *Advances in Applied Probability*, 36(2):340–354, 2004.
- [10] L. Devroye and G.L. Wise. Detection of abnormal behavior via nonparametric estimation of the support. *SIAM Journal of Applied Mathematics*, 38(3):480–488, 1980.
- [11] T. Dey and W. Zhao. Approximating the medial axis from the voronoi diagram with a convergence guarantee. *Algorithmica*, 38:179–200, 2003.
- [12] D. Fritsch, S. Pizer, B. Morse, D. Eberly, and A. Liu. The multiscale medial axis and its applications in image registration. *Pattern Recognition Letter*, 15:445–452, May 1994.
- [13] C. Genovese, M. Perone-Pacifico, I. Verdinelli, and L. Wasserman. The geometry of non-parametric filament estimation. *Journal of the American Statistical Association*, 107(498), 2012.

- [14] G. Matheron. *Image Analysis and Mathematical Morphology, Volume 2: Theoretical Advances*. Academic Press, London,, 1988.
- [15] A. Montero and J. Lang. Skeleton pruning by contour approximation and the integer medial axis transform. *Computer and Graphics*, 36:477–487, 2012.
- [16] R. Ramamurthy and R. T. Farouki. Voronoi diagram and medial axis algorithm for planar domains with curved boundaries, i and ii. *Journal of Computational and Applied Mathematics*, 102:119–141, 1999.
- [17] A. Rodríguez-Casal. Set estimation under convexity type assumptions. *Annales de l’Institut Henri Poincaré (B): Probability and Statistics*, 43(6):763–774, 2007.
- [18] D. Shaked and A. Bruckstein. Pruning medial axes. *Computer Vision and Image Understanding*, 69(2):156–169, February 1998.
- [19] G. Walther. On a generalization of Blaschke’s rolling theorem and the smoothing of surfaces. *Mathematical Methods in the Applied Sciences*, 22(4):301–316, 1999.
- [20] S. Xia, N. Ding, M. Jin, H. Wu, and Y. Yang. Medial axis construction and applications in 3d. In *Proc. IEE INFOCOM*, pages 305–309, 2013.
- [21] L. Younes. *Shapes and Diffeomorphisms*. Springer-Verlag Berlin Heidelberg, 2010.