

An Efficiency Criterion for 2D Shape Model Selection.

Kathryn Leonard

Department of Applied and Computational Mathematics
California Institute of Technology

k.leonard@alumni.brown.edu

Abstract

We propose efficiency of representation as a criterion for evaluating shape models, then apply this criterion to compare the boundary curve representation with the medial axis. We estimate the ε -entropy of two compact classes of curves. We then construct two adaptive encodings for non-compact classes of shapes, one using the boundary curve and the other using the medial axis, and determine precise conditions for when the medial axis is more efficient. Along the way we construct explicit near-optimal boundary-based approximations for compact classes of shapes and an explicit compression scheme for non-compact classes of shapes based on the medial axis. We end with an application of the criterion to shape data.

1. Introduction

The goals of this work are threefold. First, we wish to introduce efficiency of representation as a way to evaluate deterministic shape models, defining for nonlinear shape spaces an analogue of the notion of sparsity in wavelet-type analysis of linear spaces. Our second goal is to use these ideas to analyze the efficiency of two popular shape models: the boundary curve and the medial axis. We demonstrate that the medial axis has a defensible position as a shape model, despite current skepticism. Finally, we are interested in theoretical properties of spaces of shapes. In the course of evaluating the efficiency of boundary and skeletal models, we construct efficient coverings of shape space. The nature and the size of these coverings contributes to the currently growing body of mathematical understanding of these highly nonlinear, infinite-dimensional spaces.

The paper as a whole realizes the first goal, and three theorems realize the second and third. Theorem 13 gives a criterion in terms of medial data for when the medial axis is a more efficient shape model than the boundary curve, and we provide examples of shapes better modeled by each. Theorem 2 gives a tight estimate for the ε -entropy of certain compact classes of curves, and a near-tight estimate for boundary curves, while Theorem 10 gives an indication of what

an optimal covering should look like in the non-compact case. Along the way, we explicitly construct what are, to our knowledge, the first provably optimal coverings of a non-linear infinite-dimensional space. We end with some empirical results that suggest that the medial axis is overwhelmingly more efficient for naturally occurring shapes.

While our work is theoretical in nature, it has strong applications to the larger vision community. Shape spaces have attracted serious attention of late, as researchers work to place shape-related tasks on shape manifolds [8, 2, 16]. Doing so is very difficult, thereby generating a general interest in the structure of shape spaces. In addition, our results support researchers working with the medial axis and other generalized cylinder models. Since Blum's introduction of the medial axis, it has been applied to many shape-related problems, such as recognition [23, 21, 18], animation [20], reconstruction [1] and medical imaging [22]. Finally, the work should interest anyone modeling 2D shape. Regardless of the shape model, knowing that the medial axis or the boundary is more efficient for a particular shape can guide analysis of that model, just as knowing the optimal basis for a linear space guides analysis carried out in other bases.

2. Metric Spaces of Curves

We will identify a shape with the boundary curve of its silhouette. Our spaces of curves will assume a canonical location and orientation for each curve.

Let \mathcal{C} be the collection of all C^1 curves $\gamma(s) \in \mathbb{R}^2$ of bounded curvature and length, arclength parameterized on $D \subset [0, L]$ for $L \geq 0$, with tangent angle function θ_γ , satisfying $\gamma(0) = (0, 0)$, $\theta_\gamma(0) = 0$. Let $\mathcal{J} \subset \mathcal{C}$ be the subset of immersions of S^1 into \mathbb{R}^2 . We are interested in determining the massiveness of these spaces of curves. We do this in two ways: first, by computing the ε -entropy of certain compact subclasses, and second, by analyzing the rate of a deterministic adaptive encoding scheme in the non-compact setting. The compact classes of interest are:

- (a) $\mathcal{C}_K^L \subset \mathcal{C}$, the collection of uniformly Lipschitz curves

of length at most L , so that for $\gamma \in \mathcal{C}_K^L$, $|\theta_\gamma(s_1) - \theta_\gamma(s_2)| \leq K|s_1 - s_2|$.

(b) $\mathcal{J}_K^L = \mathcal{J} \cap \mathcal{C}_K^L$.

We define a C^1 -type metric on \mathcal{C} . For $\gamma_1, \gamma_2 \in \mathcal{C}$ and $\lambda > 0$ a dimension-normalizing constant, define:

$$\rho(\gamma_1, \gamma_2) = \sup_{i=1,2} \left\{ \sup_{\substack{s_j \\ j \neq i}} \inf_{s_i} \frac{1}{\lambda} |\gamma_i(s_i) - \gamma_j(s_j)| + |\theta_i(s_i) - \theta_j(s_j)| \right\}.$$

Note that as $\lambda \rightarrow \infty$, the distance between any two closed curves goes to zero. A suitable choice of λ for \mathcal{C}_K^L , for example, is $\lambda < \frac{1}{K}$. The L^∞ framework is desirable for ε -entropy-type constructions because L^∞ balls stack efficiently, and the inclusion of a first-order term derives from the important role orientation plays in visual perception [14].

3. ε -entropy

The ε -entropy of a totally bounded metric space, defined below, essentially counts the minimum number of ε -balls required to cover the space. There is a direct connection between ε -entropy and fixed length encoding. Suppose one wishes to ε -represent each element in a compact metric space with a fixed number of bits in the most efficient way possible. Given an ε -cover $\{U_i\}_1^n$ for the space, each element can be represented by the ε -ball to which it belongs. Enumerating these balls gives a binary representation for each element of the space, requiring $\lceil \log n \rceil$ bits per ball. If the covering is minimal, the number of bits will be given by the ε -entropy of the space.

In this section, we construct a minimal covering for \mathcal{C}_K^L and a near-minimal covering for \mathcal{J}_K^L . In other words, the boundary curve attains optimal efficiency for fixed-length encodings of shapes.

We begin by introducing the concept of ε -entropy, invented by Kolmogorov as a way of quantifying the massiveness of infinite-dimensional metric spaces. It captures the exponent of the number of balls in a minimal ε -covering of compact subsets of the space [9]. We present one of his theorems below; it will play a key role in our work.

3.0.1 Definitions and preliminaries

Consider a subset $X \subset (M, \rho)$, where M is a metric space with metric ρ . A system of sets $U_\alpha, \alpha \in A$, such that $X \subseteq \bigcup_{\alpha \in A} U_\alpha$ and the diameter d of U_α satisfies $d \leq 2\varepsilon$, is called an ε -cover for X . A set of points $\{x_\alpha\} \subset M$ is an ε -net for X when for any $x \in X$, there exists an x_α so that $\rho(x, x_\alpha) \leq \varepsilon$.

From a compression standpoint, the most desirable ε -cover is one with the fewest balls. This leads to the definition of ε -entropy. Let \mathcal{N}_ε be the cardinality of a minimal ε -cover for a totally bounded set $X \subset (M, \rho)$. Then the ε -entropy of X is $\mathcal{H}_\varepsilon(X, \rho) = \log_2 \mathcal{N}_\varepsilon$.

A set $U \subset X$ is ε -separated when $\rho(x_1, x_2) \geq \varepsilon$ for any $x_1 \neq x_2 \in U$. For \mathcal{M}_ε equal to the maximal number of elements in an ε -separated set $U \subset X$, the ε -capacity of X is $\mathcal{C}_\varepsilon(X, \rho) = \log_2 \mathcal{M}_\varepsilon$. We have the following relationship between these quantities:

$$\mathcal{M}_{2\varepsilon}(X) \leq \mathcal{N}_\varepsilon(X) \leq \mathcal{M}_\varepsilon(X)$$

$$\mathcal{C}_{2\varepsilon}(X) \leq \mathcal{H}_\varepsilon(X) \leq \mathcal{C}_\varepsilon(X).$$

Classically, ε -entropy is estimated by constructing an ε -cover with K_ε elements and a 2ε -separated set with $L_{2\varepsilon}$ elements so that $\lim_{\varepsilon \rightarrow 0} \frac{K_\varepsilon}{L_{2\varepsilon}} = 1$, or at worst is some non-zero constant. We will do the same.

The following theorem of Kolmogorov and Tikhomirov is one of the first estimates of ε -entropy, and is rare because of the tightness of the estimate. Consult [9] for proofs and full generality, and refer to Figure 1 for an idea of what the optimally efficient ε -balls look like.

Theorem 1. [9] For $I = [a, b]$, define $\mathcal{F}_1^I(C)$ to be the collection of functions $f : I \rightarrow \mathbb{R}$ so that $f(a) = 0$ and $|f(x) - f(x')| \leq C|x - x'| \forall x, x' \in I$. Then:

$$\mathcal{H}_\varepsilon(\mathcal{F}_1^I(C), L_\infty) = \begin{cases} \frac{|b-a|C}{\varepsilon} - 1 & \frac{|b-a|C}{\varepsilon} \in \mathbb{Z}^+ \\ \left\lceil \frac{|b-a|C}{\varepsilon} \right\rceil & \text{else.} \end{cases}$$

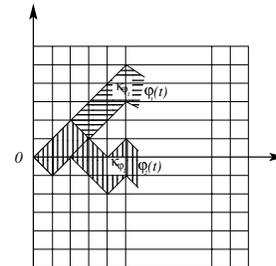


Figure 1. Epsilon corridors [9]. The function $\phi(t)$ defines the upper boundary and the lower boundary is $\phi(t) - 2\varepsilon$.

3.1. ε -entropy for Classes of Curves

This section is devoted to proving the following theorem. Although we have not yet proved it, we believe $C = KL$ in Theorem 2(b).

Theorem 2.

(a) $\mathcal{H}_\varepsilon(\mathcal{C}_K^L, \rho) \sim \frac{KL}{\varepsilon}$.

(b) $\mathcal{H}_\varepsilon(\mathcal{J}_K^L, \rho) \sim \frac{C}{\varepsilon}$ for $(KL - 2\pi) \leq C \leq KL$.

3.1.1 Upper Bound: K_ε

To construct an ε -cover for curves, we will construct an L^∞ covering of the space of tangent angle functions, then lift that covering to a coarse covering in the space of curves. Refining the lifted covering will require only lower order terms. The details are quite technical, but the idea is simple. Given an ε -ball of tangent angle functions, one may integrate to find a primitive curve for each function. For every Δ in arclength, the worst-case curve will have drifted roughly $\Delta\varepsilon$ away from the primitive of the center of the ε -ball. Some correction is therefore needed every so often to keep the balls small in the metric ρ .

Proposition 3. *There exists a $\left(\delta + \frac{\sqrt{\delta^3}}{\lambda}\right)$ -cover for (\mathcal{C}_K^L, ρ) with no more than:*

$$K_\delta = \left\lceil \frac{4L}{\sqrt{\delta^3}} \right\rceil 2^{\lceil \frac{KL}{\delta} \rceil + \log 3 \lceil \frac{4L}{\sqrt{\delta}} \rceil}$$

elements.

Proof. First, partition the interval $[0, L]$ of possible arclength into subintervals of width $\frac{\sqrt{\delta^3}}{4}$, giving $\left\lceil \frac{4L}{\sqrt{\delta^3}} \right\rceil$ subintervals. Let l_δ be the right endpoint of any such subinterval. We will parameterize γ of length $l \in \left[l_\delta - \frac{\sqrt{\delta^3}}{4}, l_\delta\right]$ by $s \in [0, l_\delta]$ $\frac{d\gamma}{ds} = \frac{l}{l_\delta}$.

Next, construct a δ -cover for the tangent angle functions defined over each particular subinterval. For all curves parameterized over $[0, l_\delta]$, $|d\theta/ds| \leq \frac{l}{l_\delta} K \leq K$. We may therefore apply Theorem 1 to obtain an L^∞ δ -cover for the associated tangent angle functions with at most $2^{\lceil \frac{KL}{\delta} \rceil}$ centers θ_δ . Each θ_δ gives rise to a curve

$$\tilde{\gamma}_\delta(s) = \int_0^s \langle \cos \theta_\delta(t), \sin \theta_\delta(t) \rangle dt.$$

To each γ_δ , we denote the associated neighborhood of curves by $\Gamma_\delta(\gamma_\delta)$. Take $\Delta = \frac{\sqrt{\delta}}{4}$. Every Δ in arclength, refine the cover by adding a new center that jumps a distance $\frac{\sqrt{\delta^3}}{4} + \frac{\delta\Delta}{2}$ in one of 3 evenly spaced directions. A short argument indicates that: (1) such curves give rise to a $\left(\delta + \frac{\sqrt{\delta^3}}{\lambda}\right)$ for $\Gamma'_\delta(\tilde{\gamma}_\delta)$, and (2) the number of balls in the refinement of the cover for $\Gamma_\delta(\gamma_\delta)$ is no more than $2^{\log 3 \lceil \frac{4L}{\delta^{\frac{1}{\varepsilon}-1}} \rceil}$.

Applying this process within each length subinterval gives the result. \square

Taking δ so that $\varepsilon = \delta + \sqrt{\delta^3}$, and applying Proposition 3 proves the following corollary, which gives our upper bound.

Corollary 4. $\mathcal{H}_\varepsilon(\mathcal{C}_K^L, \rho) \preceq \frac{KL}{\varepsilon}$.

3.1.2 Lower Bounds: $L_{2\varepsilon}$

We construct asymptotically 2ε -separated sets in both \mathcal{C}_K^L and \mathcal{J}_K^L . This is not a straightforward as it might appear, as we must guard against pairs of curves where $|\gamma_1(s) - \gamma_2(s)|$ is large but there exists an s' such that $|\gamma_1(s) - \gamma_2(s')|$ is small. To do so, we will restrict how far a curve can wander away from the horizontal axis. In addition, for \mathcal{J}_K^L , we must also ensure that the curves are closed. This is what loses the tightness of the constant in Theorem 13(b).

For both lower bounds, we will construct functions that return to the x -axis relatively often, while maintaining differentiability. To do so, we will concatenate pieces of functions that begin and end on the axis. Counting the number of these smaller functions will require that we count the number of realizations of an n -step symmetric random walk g satisfying $g(0) = g(n) = 0$ and $\int g = 0$, where g is piecewise linear with slope ± 1 . We do so using a probabilistic argument. Let $a(n) = \int_0^n g$, and for fixed n consider the random variable $z = \langle g(n), a(n) \rangle \Sigma^{-1} \langle g(n), a(n) \rangle^T$, where Σ is the covariance matrix for the random vector $\langle g(n), a(n) \rangle$. Using level sets for z , we find a lower bound for the number of walks g so that $\langle g(n), a(n) \rangle$ are ‘‘close enough’’ to the origin, then we add steps to those walks to bring the endpoints and areas back to 0. The details are quite technical, so we refer the interested reader to [11, 12], and instead present the relevant result.

Proposition 5. *Let $I = [a, b]$, and define $\mathcal{F}_{1,0}^I(C)$ to be the collection of functions defined on I so that $f(a) = f(b) = 0$ and $|f(x) - f(x')| \leq C|x - x'|$, $\forall x, x' \in I$. There exists a 2ε -separated set in $(\mathcal{F}_{1,0}^I, L^\infty)$ with $\mathcal{M}'_{2\varepsilon}$ elements, where $\mathcal{M}'_{2\varepsilon} \succeq 2^{\frac{C|b-a|}{\varepsilon}}$.*

3.1.3 Lower Bound for \mathcal{C}_K^L

We are now ready to estimate the lower bound for \mathcal{C}_K^L . We begin by constructing a 2δ -separated set of derivative functions by concatenating smaller pieces, then take the primitives of those functions to find a set of differentiable functions that will be asymptotically 2ε -separated as curves in the metric ρ .

Proposition 6. *There exists a 2ε -separated set for (\mathcal{C}_K^L, ρ) with $\mathcal{M}_{2\varepsilon}$ (\mathcal{C}_K^L) elements, where $\mathcal{M}_{2\varepsilon}(\mathcal{C}_K^L) \succeq 2^{\lceil \frac{KL}{\varepsilon} \rceil}$.*

Proof. Choose δ so that $\varepsilon = \frac{\delta}{1 + \frac{K^2\delta}{4}}$, giving $\lim_{\varepsilon \rightarrow 0} \frac{\delta}{\varepsilon} = 1$.

For $L' = \frac{L}{\sqrt{1 + \frac{K^2\delta}{4}}}$, divide the interval $[0, L']$ into subintervals

$I_k = [a_k, a_{k+1}]$ of width $\sqrt{\delta}$, giving $k \leq \left\lceil \frac{L'}{\sqrt{\delta}} \right\rceil$. Within each I_k , apply Proposition 5 with $C = K$ to obtain a collection of 2δ -separated functions $\left\{ \frac{df_j^k}{dx} \right\}$, $j = 1, \dots, n$, where $n \succeq \frac{1}{2} 2^{\frac{KL\sqrt{\delta}}{\varepsilon}}$. The functions satisfy, for $i \neq j$:

$$(i) \frac{df_j^k}{dx}(a_k) = \frac{df_j^k}{dx}(a_{k+1}) = 0$$

$$(ii) \int_{I_k} \frac{df_j^k}{dx} = 0$$

$$(iii) \left\| \frac{df_i^k}{dx} - \frac{df_j^k}{dx} \right\|_{\infty} \geq 2\delta$$

$$(iv) \left\| \frac{df_j^k}{dx} \right\|_{\infty} \leq \frac{K\sqrt{\delta}}{2}$$

Integrating, we obtain a collection of primitives $\{f_j^k\}$ with curvature functions bounded by K , and tangent angle functions $\{\theta_j^k\}$, where $\theta_j^k = \arctan \frac{df_j^k}{dx}$. These primitives satisfy the following properties for $i \neq j$:

$$(i) f_j^k(a_k) = f_j^k(a_{k+1}) = 0$$

$$(ii) \theta_j^k(a_k) = \theta_j^k(a_{k+1}) = 0$$

$$(iii) \|\theta_i^k - \theta_j^k\|_{\infty} \geq \frac{2\delta}{1 + \frac{K^2\delta}{4}} = 2\varepsilon.$$

Construct functions $\{f_i\}$ on $[0, L']$ by concatenating sequences $\{f_i^k\}$. The number of distinct functions will be 2^m , where $m \succeq \left(\frac{L'}{\sqrt{\delta}}\right) \left(\frac{K\sqrt{\delta}}{\delta}\right) = \frac{KL'}{\delta}$, and each will satisfy the properties above. In particular, $\frac{df_i}{dx}(0) = \frac{df_i}{dx}(L') = 0$ and $f_i(0) = f_i(L') = 0$ for every i .

Each function f_i^k has length

$$\begin{aligned} L_i^k &\leq \int_{I_k} \sqrt{1 + (K\frac{\sqrt{\delta}}{2})^2} dx \\ &= \sqrt{\delta} \cdot \sqrt{1 + \frac{K^2\delta}{4}}. \end{aligned}$$

The arclength of f_i is therefore at most $L' \sqrt{1 + \frac{K^2\delta}{4}}$, giving $m \succeq \frac{KL'}{\delta} \succeq \frac{KL}{\varepsilon}$.

Finally, we demonstrate the 2ε -separation of these functions as curves. Recall from the definition of ρ that $\lambda < 1/K$. For $f_i, f_j, i \neq j$, we may assume there exists some subinterval $I = [x_0, x_0 + \frac{\delta}{K}]$ so that $\frac{df_i}{dx}|_I$ has slope $+K$ and $\frac{df_j}{dx}|_I$ has slope $-K$, giving $|\frac{df_i}{dx}(x_0 + \frac{\delta}{K}) - \frac{df_j}{dx}(x_0 + \frac{\delta}{K})| \geq 2\delta$. See Figure 2.

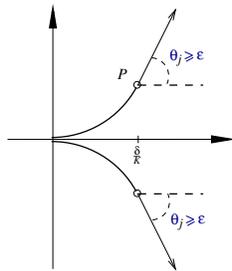


Figure 2. Separation of curves.

Consider the ρ -distance of f_j to the point $P = (\frac{\delta}{K}, f_i(\frac{\delta}{K}))$. A short argument gives:

$$\rho(f_i, f_j) \geq \min_x \frac{1}{\lambda} \left| \frac{\delta}{K} - x \right| + \frac{1}{1 + \frac{K^2\delta}{4}} |\delta + Kx|.$$

Taking derivatives with respect to x , we see that since $\lambda < 1/K$, the minimum of the last expression occurs when $x = \delta/K$, giving $\rho(f_1, f_2) = |\theta_i(\frac{\delta}{K}) - \theta_j(\frac{\delta}{K})| \geq \frac{2\delta}{1 + \frac{K^2\delta}{4}} = 2\varepsilon$ as desired. \square

Corollary 7. $\mathcal{H}_\varepsilon(\mathcal{C}_K^L, \rho) \succeq \frac{KL}{\varepsilon}$.

3.1.4 Lower Bound for (\mathcal{J}_K^L, ρ)

Our challenge here is to construct 2ε -separated closed curves. Suppose we have a collection of functions like those constructed above, 2δ -separated in the metric ρ_{C^1} , defined on $[a, b]$ with second derivative bounded by $\pm K$ such that $f(a) = f(b) = 0$. We may select any two such functions and use their graphs to join two halves of a circle of radius $1/K$. See Figure 3. For functions satisfying $f(0) = f(L') = \frac{df}{dx}(0) = \frac{df}{dx}(L') = 0$, and for appropriate choices of $|b - a|$ and δ , these curves will give the desired 2ε -separated set. In other words, we may build upon the construction in Proposition 6 for open curves to obtain a 2ε -separated set of closed curves.

Proposition 8. *There exists a 2ε -separated set for (\mathcal{J}_K^L, ρ) with $\mathcal{M}_{2\varepsilon}(\mathcal{J}_K^L) = 2^{2^m}$ elements where*

$$2m \succeq \frac{KL - 2\pi}{\varepsilon}$$

Proof. Select δ so that $\varepsilon = \frac{\delta}{\sqrt{1 + \frac{K^2\delta}{4}}}$. We will use a slight modification of the collection $\{f_i\}$ constructed in the proof of Proposition 6 to construct a 2ε -separated set of closed curves. Taking any two (not necessarily distinct) functions f_i, f_j , we may join two halves of a circle of radius $1/K$ with the functions, forming a closed C^1 curve with bounded curvature and length. See Figure 3. The number of such curves is 2^{2^m} , where $m \succeq \frac{KL'}{\delta}$. We have from the proof of Proposition 6 that this collection of functions is 2ε -separated as curves in the metric ρ .

Since each function f_i has length $L' \sqrt{1 + \frac{K^2\delta}{4}}$ and two halves of the circle of radius $1/K$ add an additional $2\pi/K$, taking $L' = \frac{L - \frac{2\pi}{K}}{\sqrt{1 + \frac{K^2\delta}{4}}}$, we have $m \succeq \frac{KL - 2\pi}{\delta} \sim \frac{KL - 2\pi}{\varepsilon}$, as claimed. \square

Corollary 9. $\mathcal{H}_\varepsilon(\mathcal{J}_K^L, \rho) \succeq \frac{KL - 2\pi}{\varepsilon}$.

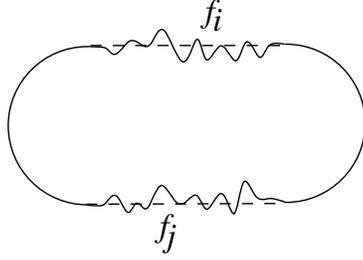


Figure 3. Construction of closed curves for 2ε -separated set.

4. Non-compact Classes and Adaptive Codes

The ε -entropy gives the minimum number of bits required to represent elements in a compact space with a fixed code length. Our computations above indicate that in such a setting, the boundary curve gives an optimally efficient representation. To relax the compactness constraint, we will allow a variable code length, thereby replacing the finite number of codewords of fixed length with a countable number of codewords of varying length in an adaptive coding scheme.

We will construct an adaptive coding scheme for functions in Lemma 10, which will give rise to two adaptive encodings for curves, one based on the boundary curve and the other on the medial axis. We will apply the Rissanen's MDL principle in Section 6 to determine which of those two encodings is most efficient.

4.1. Adaptive Coding for Lipschitz Functions

We carry out an adaptive coding scheme for Lipschitz functions, again motivated by Theorem 1.

Theorem 10. *For every $\varepsilon > 0$, there exists a countable codebook $F_\varepsilon = \{f_1, f_2, \dots\}$, depending only on ε , with the following property. For every Lipschitz function f defined on $[a, b]$ so that $f(a) = 0$ and $f'(x)$ is continuous a.e., there are constants $C(f, \delta)$ such that for all δ , there is a codeword $f_n \in F_\varepsilon$ such that $\|f - f_n\|_\infty \leq \varepsilon$ and f_n has description length:*

$$L(f_n) \leq \left\lceil \frac{\int |f'| + \delta}{\varepsilon} \right\rceil + C(f, \delta).$$

Proof. Since f' is continuous a.e. and bounded, f' is Riemann integrable [5]. Therefore, for any δ , there exists a step function g taking on rational values, with a finite number of jumps at rational points $\{x'_j\}$, so that $|f'| \leq g$ and $\int g \leq \int |f'| + \delta$. This means that on each subinterval $I_j = [x_j, x_{j+1})$ where g is constant, f is Lipschitz with constant $g(x_j)$. Denote the number of jumps by m . Note that $g \equiv C$ gives the result in Theorem 1.

Using g , we may determine a variably spaced finite number of points $\{x_k\}$ so that for any k , $\int_{x_k}^{x_{k+1}} g \leq \varepsilon$. In partic-

ular, on each subinterval I_j , select the points spaced $\frac{\varepsilon}{g(x'_j)}$ apart. There will be at most:

$$\left\lceil \frac{g(x'_j)|I_j|}{\varepsilon} \right\rceil + 1 = \left\lceil \frac{\int_{I_j} g(x)}{\varepsilon} \right\rceil + 1$$

such points. Take $\{x_k\}$ to be the collection of $\{x'_j\}$ together with these equally spaced points.

We now construct an approximation f_n for f . We claim there exists a piecewise linear function ϕ_n , with slope $\pm g(x_k)$ on the interval $J_k = [x_k, x_{k+1})$, vanishing at a , so that f is contained in an ε -corridor $K(\phi_n)$ with upper boundary ϕ_n as depicted in Figure 1. On J_1 , take $\phi_n(x) = g(0)x$. Certainly, since f is Lipschitz with constant $g(0)$ on I_1 , $f \subset K(\phi_n)$. Inductively, assume ϕ_n has been constructed so that $f \subset K(\phi_n)$ for $x \leq x_k$. We wish to select ϕ_n so that f remains in $K(\phi_n)$ on I_k . Since f is Lipschitz with constant $g(x_k)$ on I_k and $f \subset K(\phi_n)$ for $x \leq x_k$, one of the following is true:

- (a) $f(x_{k+1}) \in [\phi_n(x_k) - g(x_k)(x - x_k), \phi_n(x_k) + g(x_k)(x - x_k)] \subset [\phi_n(x_k) - \varepsilon, \phi_n(x_k) + \varepsilon]$,
- (b) $f(x_{k+1}) \in [\phi_n(x_k) - g(x_k)(x - x_k) - 2\varepsilon, \phi_n(x_k) + g(x_k)(x - x_k) - 2\varepsilon] \subset [\phi_n(x_k) - 3\varepsilon, \phi_n(x_k) - \varepsilon]$.

If (a) is true, then defining ϕ_n to have positive slope on I_k gives $f \subset K(\phi_n)$. If (b) is true, then defining ϕ_n to have negative slope on I_k gives the desired result. Taking f_n to be the center of the corridor $K(\phi_n)$, we have $\|f - f_n\|_\infty \leq \varepsilon$.

Encode f by encoding f_n , or equivalently, ϕ_n . To do so requires describing g , which in turn requires describing the collection of points $\{x_k\}$ as outlined above. Since g has rational jumps at rational values, encoding g requires a fixed and finite number of bits depending only on f and δ , yielding the constant $C(f, \delta)$. We must also describe the sequence of signs \pm to assign to the slopes $g(x'_j)$ at each of the points $\{x_k\}$, which requires a single bit at each of the $\{x_k\}$. This requires at most:

$$\sum_j \left(\left\lceil \frac{\int_{I_j} g}{\varepsilon} \right\rceil + 1 \right) + m \leq \left\lceil \frac{\int_{[a,b]} g}{\varepsilon} \right\rceil + 2m$$

bits. Then, absorbing $2m$ into $C(f, \delta)$, the total number of bits required to describe f_n satisfies:

$$L(f_n) \leq \frac{\int g}{\varepsilon} + C(f, \delta) \leq \frac{\int |f'| + \delta}{\varepsilon} + C(f, \delta),$$

as claimed. □

4.2. Adaptive Coding for Curves

In the compact setting, the leading term in the bit rate comes from approximating the tangent angle function; correcting for location requires only lower order terms. The same argument applies in the adaptive setting. Theorem 10 describes a method for adaptively encoding functions in a non-compact space by adaptively approximating each function's derivative, and we may use it to encode curves by applying it to the associated tangent angle function. Doing so gives an encoding with leading term for the bit rate of:

$$\frac{\int |\kappa_\gamma| ds + \delta}{\varepsilon},$$

for some fixed δ .

5. Medial Axis

Before we describe our encoding scheme for the medial axis representation and return to the discussion of efficient representation, we briefly review the medial axis construction, introduced by Blum as an alternative description for closed planar regions [3].

The **medial axis pair** (\mathbf{m}, r) of a closed region in \mathbb{R}^2 consists of $\mathbf{m}(t)$, the closure of the locus of centers of maximal bitangent circles contained within the region, and $r(t)$, the function of associated radii. We will consider medial axis of regions bounded by a simple, closed, C^2 curve, a setting where many beautiful results may be proved [6, 11].

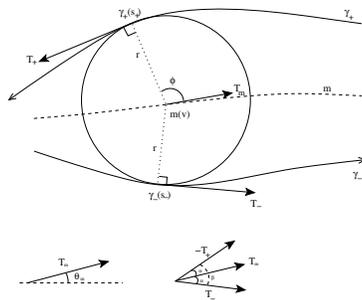


Figure 4. Notation and relationships for the medial axis.

First, we introduce some notation, applicable locally within a parameterized branch $m(t)$ of the medial curve \mathbf{m} for a closed curve γ . Refer to Figure 4. Subscripts m refer to quantities associated to the medial branch. Following $m(t)$ in the direction of its orientation, subscripts \pm refer to the portions of the boundary curve lying to the right or the left of m , so $\gamma_+(t)$ is the boundary point lying to the left of $m(t)$ and $\gamma_-(t)$ is the boundary point lying to the right. The standard orientation for γ_\pm is therefore the same as the one inherited from m for γ_- and opposite for γ_+ . Denote by ϕ the smaller angle between \mathbf{T}_m and $-\mathbf{N}_+$, the outward pointing normal to γ_+ . The symbols s , s_+ , s_- , and v

will always indicate the arclength parameter for γ , γ_+ , γ_- , and m , respectively, and the symbol ' will be reserved for derivatives with respect to v .

Theorem 11.

For a smooth branch of the medial curve, the following relationships hold:

1. $r' = -\cos \phi$.
2. $\theta_\pm = \theta_m \pm \phi + \frac{\pi}{2}$.
3. $\mathbf{N}_\pm = r' \mathbf{T}_m \mp \sqrt{1 - r'^2} \mathbf{N}_m = -\cos \phi \mp \sin \phi \mathbf{N}_m$.

6. Model Selection

We return to the question of shape model selection, choosing between the boundary curve and the medial axis. In the compact setting, ε -entropy provides a benchmark for optimal description length. No such benchmark exists in the non-compact setting, but we can make reasonable comparisons using a deterministic implementation of Rissanen's Minimum Description Length (MDL) principle [17]. Behind this formulation is the idea that the best compression scheme will exploit regularities in the space, thereby resulting in the smallest expected code length. This gives a method for evaluating coding schemes: the encoding resulting in the shortest code length is the one best capturing the salient structures.

While the boundary curve gives an optimal fixed-length bit rate, a variable-length adaptive code allows the medial axis to display its strengths. In both fixed- and adaptive-length encoding, the high-order term in the bit length comes from encoding the tangent angle function to the boundary. We may therefore exploit the way in which the medial axis encodes the geometry of the boundary curve. Because $\theta_\pm = \theta_m \pm \phi + \pi/2$, to construct an ε -approximation to θ_γ requires approximations to θ_m and ϕ with error summing to ε . For $0 < \eta < \varepsilon$, if we η -encode θ_m and $(\varepsilon - \eta)$ -encode ϕ , we obtain an encoding length for θ_γ with leading term:

$$\frac{\int |\kappa_m| dv + \delta_1}{\eta} + \frac{\int |\phi| dv + \delta_2}{\varepsilon - \eta}.$$

Note that this also explains why the medial axis will be less efficient in the fixed length setting: to construct a uniform covering with ε -balls, both θ_m and ϕ must be accurate within $\varepsilon/2$. For most curves, such a high level of accuracy in both angle functions is unnecessary.

To relate the number of bits in this medial axis encoding to the number of bits in a boundary encoding, we require the following lemma, which is a consequence of the properties listed in Theorem 11.

Lemma 12. Let m be a branch of a medial curve of length l for $\gamma \in \mathcal{E}$. Let s be an arclength parameter for γ so that $\gamma(s)|_{s \in D} = \gamma_+ \cup \gamma_-$. Then:

$$\int_D |\kappa| ds = \int_{[0,l]} |\kappa_m| + |\phi'| + ||\kappa_m| - |\phi'||| dv.$$

To select the optimal value for the error tolerance η within a fixed region of a branch of \mathbf{m} , we minimize the expression:

$$\frac{\int_{[0,l]} |\kappa_m| dv}{\eta} + \frac{\int_{[0,l]} |\phi'| dv}{\varepsilon - \eta}$$

with respect to η for $0 \leq \eta \leq \varepsilon$. Note that since the δ_i can be made arbitrarily small, we have dropped reference for ease of computation. Set $a = \int_{[0,l]} |\kappa_m| dv$ and $b = \int_{[0,l]} |\phi'| dv$. The minimization then gives $\eta = \varepsilon/2$ for $a = b$ and $\eta = \frac{\varepsilon}{b-a} (\sqrt{ab} - a)$ for $a \neq b$. We will adaptively select regions of the medial curve and associated local optimal values for η to give the best bit rate for the medial axis within a particular branch.

Theorem 13. Let m be an arclength parameterized medial branch defined on a closed interval I with corresponding boundary segments γ_{\pm} . Then encoding γ_{\pm} via the medial axis is more efficient than directly adaptively encoding θ_{γ} whenever I can be partitioned into a finite number of subintervals I_j where for each j :

$$\frac{\sup_{I_j} |\kappa_m|}{\sup_{I_j} |\phi'|} > 2 + \sqrt{3} \text{ or}$$

$$\frac{\sup_{I_j} |\phi'|}{\sup_{I_j} |\kappa_m|} > 2 + \sqrt{3}.$$

Proof. Following the proof of Theorem 10, select g_{κ} and δ_1 satisfying $|\kappa_m| \leq g_{\kappa}$ and $\int g_{\kappa} \leq \int |\kappa_m| + \delta_1$, and select g_{ϕ} and δ_2 satisfying $|\phi'| \leq g_{\phi}$ and $\int g_{\phi} \leq \int |\phi'| + \delta_2$. Construct an encoding scheme as follows. Partition I , the domain of m , into maximal subdomains $\{I_i\}$ on which both g_{κ} and g_{ϕ} are constant. Since both functions are piecewise constant with a finite number of jumps, the number of such subdomains will be finite. On each subdomain, compute the minimizing η_i . Then the medial axis will be more efficient when:

$$g_{\kappa} + g_{\phi} + 2\sqrt{g_{\kappa}g_{\phi'}} < 2 \max\{g_{\kappa}, g_{\phi'}\}.$$

For $g_{\kappa} \geq g_{\phi}$, this gives:

$$\frac{g_{\kappa}}{g_{\phi'}} > 2 + \sqrt{3},$$

otherwise take the reciprocal of the left side of the inequality.

Recalling the construction of the functions g_{κ} and g_{ϕ} , the result is proved. \square

An interpretation of Theorem 13 is that the medial axis decouples the curvature of the boundary curve into the portion coming from the curvature of the local axis of symmetry, i.e., the medial curve, and the portion coming from variation around that axis. When the boundary curvature comes primarily from one source or the other, the medial axis is more efficient. When the curvature of the boundary relies heavily on both sources (the most extreme case of which is when $a = b$), the boundary curve is more efficient.

6.1. Empirical data

In theory, Theorem 13 partitions shapes silhouettes into two classes: those best modeled by the boundary and those best modeled by a skeleton. Application of the criterion to the Kimia database of 1003 shapes [19], however, indicates the the medial axis is most often the more efficient model for naturally occurring shapes. See Figure 5, where the medial axis is more efficient for every shape. A related paper applies the criterion to three shape databases and offers detailed analysis of the results [13].

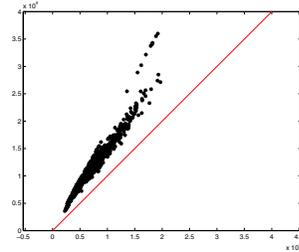


Figure 5. Medial axis bit rate (y-axis) against boundary bit rate (x-axis) for $\varepsilon = 0.01$.

7. Discussion

We have presented a novel criteria for shape model selection and offered a preliminary application to real data. Our results clearly indicate that the medial axis is an efficient shape representation, which gives insight into the inherent structure of shape space and potentially strengthens all shape models. Skeletal preferences can be built into priors on feature detectors, reduce the search for large wavelet coefficients, guide construction of constellations, as well as many other applications. We hope our work will influence these tasks, as well as inspiring further investigations into model efficiency.

References

- [1] N. Amenta, S. Choi, R. Kolluri. The power crust, union of balls, and the medial axis transform. *Computational Geometry: Theory and Applications*, **19**(2-3): 127-153, 2001.
- [2] M. F. Beg, M. I. Miller, A. Trounev, L. Younes. Computing metrics via geodesics on flows of diffeomorphisms. *International Journal of Computer Vision*, July 2003.
- [3] H. Blum. Biological shape and visual science. *J. Theor. Biol.* **38**:205-287, 1973.
- [4] J. N. Damon. Determining the geometry of boundaries of objects from medial data. *IJCV*, **63**(1): 45-64, 2005.
- [5] G. Folland. *Real analysis*, John Wiley and Sons, 1984, p. 55.
- [6] P. Giblin and B. Kimia. On the intrinsic reconstruction of shape from its symmetries. Preprint, 2002.
- [7] R. A. Katz, S. M. Pizer. Untangling the Blum medial axis transform. *IJCV*, **55**(2-3):139-153, 2003.
- [8] E. Klassen, A. Srivastava, W. Mio, S. Joshi. Analysis of planar shapes using geodesic paths in shape space. *IEEE Trans. Pattern Analysis and Machine Intelligence*, **26**:372-383, 2004.
- [9] A.N. Kolmogorov and V.M. Tikhomirov. ε -entropy and ε -capacity. *American Mathematical Society Translations, Series 2*, **17**:277-364, 1959.
- [10] A. Latto, D. Mumford, J. Shah. The representation of shape. *IEEE 1984 Proceedings of the Workshop on Computer Vision Representation and Control*.
- [11] K. Leonard. Measuring Shape Space: ε -entropy, adaptive coding and 2-dimensional shape. PhD Thesis, Brown University, 2004.
- [12] K. Leonard. Efficient Shape Modeling: ε -entropy, Adaptive Coding, and Boundary Curves -vs- Blum's Medial Axis. Preprint, 2005.
- [13] K. Leonard. 2D shape model selection via efficiency measures. Preprint, 2005.
- [14] W. Li, G. Westheimer. Human discrimination of the implicit orientation of simple symmetrical patterns. *Vision Research*, **37**(5):565-572, 1997.
- [15] D. Macrini, A. Shokoufandeh, S. Dickinson, K. Siddiqi, S. Zucker. View-based 3-D object recognition using shock graphs. Technical report, 2001.
- [16] P. Michor, D. Mumford. Riemannian geometries on the space of plane curves. Preprint, 2003.
- [17] J. Rissanen. *Stochastic complexity in statistical inquiry*. World Scientific Press, 1989.
- [18] T. Sebastian, P. Klein, B. Kimia. Recognition of shapes by editing shock graphs. *Proceedings of 8th International Conference of Computer Vision, Vancouver*, IEEE Computer Society Press, July, 2001, p. 755-762.
- [19] D. Sharvit, J. Chan, B.B. Kimia. Symmetry-based indexing of image databases. In *Content-Based Access of Image and Video Libraries*, 1998.
- [20] J. Shen, D. Thalmann. Fast realistic human body deformation for animation and VR applications. *Applications of Computer Graphics International*, Pohang, 1996, p. 166-174.
- [21] K. Siddiqi, B. Kimia. A shock grammar for recognition. *Proceedings of the Conference on Computer Vision and Pattern Recognition*, 1996, p. 507-513.
- [22] P. Yushkevich, P. Fletcher, S. Joshi, A. Thall, S. Pizer. Continuous medial representations for geometric object modeling in 2-D and 3-D. *Proceedings of the 1st Generative Model-based Vision Workshop*, Copenhagen, DK, June, 2002.
- [23] S. C. Zhu, A. L. Yuille. FORMS: A flexible object recognition and modeling system. *International Journal of Computer Vision* **20**:187-212, 1996.