# Optimal Inference for Hierarchical Skeleton Abstraction

**Alexandru Telea**[†] **Cristian Sminchisescu**[‡] **Sven Dickinson**[‡]
[†]Eindhoven University, Department of Mathematics and Computing Science, The Netherlands
[‡]University of Toronto, Department of Computer Science, Canada

## Abstract

*Skeletons are well-known representations that accommodate shape abstraction and qualitative shape matching. However, skeletons are sometimes unstable to compute and sensitive to shape detail, thus making shape abstraction and matching difficult. To address these problems, we propose a principled framework that generates a simplified, abstracted skeleton hierarchy by analyzing the quasi-stable points of a Bayesian-inspired energy function. The resulting model is parameterized by both boundary and internal structure variations corresponding to object scale and abstraction dimensions, and trades-off reconstruction accuracy and representation parsimony. Our experimental results show that the method can produce useful multi-scale skeleton representations at a variety of abstraction levels.*

**Keywords:** *Skeleton abstraction, qualitative shape matching, shock graphs, energy minimization, constrained optimization, minimum description length.*

## 1. Introduction

Recent work on the indexing and matching of shock graphs [12, 10, 5] has shown that similar shapes can lead to structurally dissimilar shock graphs. In a traditional graph matching framework that assumes a one-to-one node correspondence, the induced large distance between the two graphs does not reflect the similarity of their underlying shapes. A possible solution to this problem is to try and compute a more stable set of shock points from which the graph is constructed [4, 11], or to identify unstable shock structure and exclude it from the graph [1]. Still, these approaches commit to a single skeleton representation, derived from a single boundary representation, which can be sensitive to minor boundary perturbations. Such perturbations may be caused by a number of factors, such as small viewpoint changes, in case of view-based 3-D recognition. A small, salient branch in one view of an object may be absent in a very similar view (or object). Depending on the type and location of the corresponding node of that branch in the shock graph, the shock graphs of the two views may still be quite different. This further causes non-trivial problems when matching the two shock graphs.

Another solution to graph matching is to edit one graph so that it is isomorphic to the other. If large structural differences between two graphs representing similar shapes can be assigned a small edit cost, the similarity between the two shapes can be maintained. This approach has been proposed by Sebastian *et al*. [9], and provides effective matching in the presence of such instabilities. However, this approach is very costly, and assumes that a query and target have been identified. Although a hierarchical database partitioning framework has been proposed [8], it still requires a linear search of the prototypes.

A final class of solutions is to construct a multi-scale skeleton representation, such that, among the set of shock graphs, one per scale, one such graph is sufficiently close in structure to its database target to facilitate effective indexing [10]. The concept of multi-scale skeletons is not a new one. Ogniewicz [6] proposed multi-scale skeletons for 2D polygonal shapes. The saliency of a skeleton branch is given by the 'collapsed chord length', or boundary length between the two Voronoi sites generating that branch. The multiple scales are given by upper thresholding the above saliency measure. Similar multi-scale approaches have been proposed for raster objects by Costa *et al* [3] and Telea&Van Wijk [13].

Siddiqi *et al* [11] computed the skeleton by simulating the grassfire flow as a Hamilton-Jacobi equation. Skeleton points are found by upper thresholding the divergence of the object's distance transform gradient. The skeleton points are labeled with the time of shock formation, *i.e*. the local object width, yielding a multi-scale notion. However, this multi-scale is reported to be computationally expensive, and doesn't preserve the object topology [7].

Finally, Borgefors *et al* [2] have computed a hierarchical skeleton by extracting skeletons of successively lower resolutions of a given image. The implied multi-scale is based on the object width, similarly to [11]. Explicit correspondences between skeletons on successive scales are computed to preserve object topology (skeleton connectedness) across scales. For a detailed comparison of the above, see [7].

In this paper, we propose a skeleton simplification method that produces a hierarchy parameterized by *both* boundary *and* internal structure parameters corresponding to object scale *and* abstraction dimensions. In contrast to other approaches, we compute the stable points of our augmented hierarchy using

an algorithm that performs optimal parameter inference. For a given skeleton, the optima are computed under a cost function that trades-off reconstruction accuracy and skeleton simplicity, as suggested by a parsimony principle such as the minimum description length (MDL).

## 2. Inference for Optimal Skeleton Abstraction

We formulate the hierarchical skeleton abstraction as inference over a two parameter family $(b, s)$ representing boundary $b$ and internal structure $s$. Given an initial skeleton $M^d$, the simplified skeleton under the generative transformation set $\mathcal{M}(b, s)$ is an optimal simplification if it is *both* close to the original $M^d$ *but also* has a simple structure in a MDL sense, *e.g.* having as few number of branching points as possible. This trade-off is implemented as a sum of data and MDL energy terms

$$E(b, s|M^d) = E^d(M^d|\mathcal{M}(b, s)) + E^{mdl}(\mathcal{M}(b, s))$$

The low energy points of $E$ can be located by exhaustive search over the resulting $2d$ parameter space. We detail next the model structure including the transformation set, the data likelihood, and the MDL prior.

### 2.1 Skeleton Transformation Set ($\mathcal{M}$)

The simplified skeleton generative transformation set $\mathcal{M}(b, s)$ consists of both boundary $b$ and internal structure $s$ simplifications of the original skeleton $M^d$. While the former performs more classical boundary detail removal, the latter is intended towards shape abstraction.

**Boundary Simplification:** We compute the skeleton and its boundary simplification using the AFMM method [13]. In brief, the AFMM propagates an arc-length boundary parameterization $U$, in normal boundary direction, with constant speed, by solving the Eikonal equation $\nabla T = 1$, with $T = 0$ on the boundary. The simplified skeleton containing the branches corresponding to boundary details longer than $s$ pixels is then given by upper thresholding $\max(\partial U/\partial x, \partial U/\partial y)$ with $s$. We use the AFMM, as it is robust for any value $s \geq 2$ pixels, delivers connected skeletons, and works in near real time.

**Structural Simplification:** The structural simplification $\mathcal{M}(s)$, parameterized by $s$, applies *only* to internal skeleton branches. In the following, denote the neighbors of a point $\mathbf{x}$ of the skeleton $M$ by $n(\mathbf{x}, M) = n(\mathbf{x})$. Denote by $c(\mathbf{x}^s, M^s, M^d) = c(\mathbf{x}^s)$ the corresponding point, in the original skeleton $M^d$, of a point $\mathbf{x}^s$ in the simplified skeleton $M^s$. Finally, denote by $e(M)$ the set of all endpoints of $M$ and by $b_i(M)$ the two endpoints of the $i^{th}$ branch $b_i$ of $M$. $\mathcal{M}(s)$ removes all internal branches shorter than $s$ from $M$. When removing each branch, the two sub-skeletons are reconnected by translating one of them to join the other one. At this point, the structural simplification is, strictly speaking, done. However, this strategy causes large changes in the comparison measure we use for the silhouettes (see §2.2), since all points in the displaced sub-skeleton change their position. In order to preserve the consistency of the skeleton and to minimize the impact of the structural change, we relax (optimize) the points in $M^s$ according to the following cost function

$$F(M^s|M^d) = \sum_{\mathbf{x}^s \in M^s} \sum_{\mathbf{y}^s \in n(\mathbf{x}^s)} (||\mathbf{x}^s - \mathbf{y}^s|| - ||\mathbf{x}^d_{c(\mathbf{x}^s)} - \mathbf{y}^d_{c(\mathbf{y}^s)}||)^2$$

with two hard constraints

$$\mathbf{x}^s_i = \mathbf{x}^d_{c(\mathbf{x}^s_i)}, \forall c(\mathbf{x}^s_i) \in e(M^d)$$

$$\mathbf{x}^s_k = \mathbf{x}^s_p, \quad \text{if } b_i(M^d) = \{c(\mathbf{x}^s_k), c(\mathbf{x}^s_p)\} \textbf{ and } b_i \text{ is eliminated}$$

The first constraint enforces a boundary condition on the skeleton, *i.e.* keeps its endpoints fixed. The second constraint ensures that the elimination of the internal branches is persistent, *i.e.*, these branches don't reappear after the relaxation. Without this constraint, the relaxation tends to stretch back the simplified (removed) branches. To minimize the energy with the above constraints, we use a spring embedder approach. Every neighboring relation $n(\mathbf{x})$ corresponds to an elastic spring. We displace all points $\mathbf{x}$ towards their local energy minimum with respect to their neighbors $n(\mathbf{x})$, by sweeping all skeleton points in breadth first order, starting from the simplification places. This distributes the displacements in gradually decreasing magnitude from the simplified branches to the endpoints. The process converges after a few hundred sweeps.

### 2.2 Likelihood Term ($E^d$)

We first define the distance between two silhouettes $S_1$ and $S_2$ as

$$\mathcal{D}(S_1, S_2) = \frac{1}{\sum_{\mathbf{p} \in S_1} T(\mathbf{p}, S_1)} \sum_{\mathbf{p} \in S_1 - S_2} T(\mathbf{p}, S_1) \quad (1)$$

where $T(\mathbf{p}, S) = \min_{\mathbf{s} \in S} \parallel \mathbf{p} - \mathbf{s} \parallel$ is the solution of the Eikonal equation computed by the AFMM, *i.e.*, the distance transform of $S$ (Sec. 2.1). The denominator in (1) normalizes $\mathcal{D}$ between zero and one. Next, we define the likelihood term as a symmetric distance sum between the silhouette induced by the simplified skeleton $S^s = \mathcal{S}(\mathcal{M}(b, s))$ and the original silhouette $S^d = \mathcal{S}(M^d)$

$$E^d(M^d|\mathcal{M}(b, s)) = \mathcal{D}(S^s, S^d) + \mathcal{D}(S^d, S^s) \quad (2)$$

### 2.3 MDL Prior ($E^{mdl}$)

The MDL prior we chose encourages shapes that have a simple skeleton. The simplest shape is the circle, whose skeleton is a point. We thus use a prior term that measures the eccentricity of a proposed simplification (other discrete measures based, for example, on the number of branches, may be possible)

$$E^{mdl}(\mathcal{M}(b, s)) = \left| \frac{\sqrt{Area(S^s)}}{Per(S^s)} - \frac{1}{2\sqrt{\pi}} \right| \quad (3)$$

where $Area(S)$ is the area of a silhouette $S$ and $Per(S)$ is its perimeter. The term $\frac{1}{2\sqrt{\pi}}$ in (3) is the eccentricity of a circle, so $E^{mdl}$ is zero for circular shapes. In practice, we also normalize $E^{mdl}$ between zero and one by dividing it by the eccentricity of the initial shape $S^d$.

## 3. Experiments

We have tested the proposed algorithm on a variety of silhouette images. In fig. 2, we show results from a larger set of experiments involving biological shapes, here a bird, a horse and a human hand. The first three rows show, from left to right, the original and simplified silhouettes, as well as the corresponding skeletons and the distance transform of the simplified skeleton for an arbitrary choice of $(b, s)$ parameter values. Notice how the structural skeleton changes in column IV lead to representations that better reflect our perceptual abstraction, *e.g.* for the horse by removing the branch that connects the back two legs to the spine, similarly the three leftmost bottom branches of the bird and the upper branches corresponding to the fingers of the palm. Notice that the structural changes do not affect the overall aspect ratio of the figure (skeleton endpoints remain unchanged, see §2.1 for details). See also fig. 1 for a more quantitative evolution of the energy function and its individual components, and note the trade-off between reconstruction and structural parsimony terms. Interesting regions of the plot are the ones before sample 150, where the energy approaches a flat plateau and then starts increasing gently.

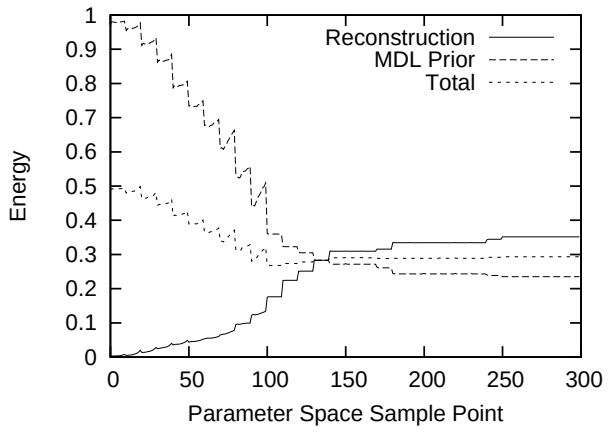


**Figure 1. Different components of the energy function for the horse in fig. 2, sampled at 300 points in parameter space over the range b=(15,800) with 30 samples and s=(0,90) with 10 samples, unfolded on the x axis.**

The last two rows of fig. 2 show an in-depth evolution of the skeleton under the change of the $(s, b)$ parameter pair. We found that the configurations given by $(s, b)$=$(0, 20)$, $(0, 50)$, $(50, 50)$ and $(50, 170)$ (not shown) are quasi-stable energy points. Notice also that these stable points correspond to intuitive abstractions, *e.g.*, when all the four finger skeleton branches have a common root. These will in turn lead to shock-graph based representations with fewer spurious nodes, therefore significantly simpler to match.

## 4. Conclusions

In this paper, we have presented a framework for multi-scale skeleton simplification and abstraction. In order to addresses the instability of 2d skeletons to 3d viewpoint and minor shape deformation, we propose a framework that produces a simplified, abstracted skeleton hierarchy by searching the quasi-stable points of a Bayesian-inspired energy function, parameterized by boundary and internal structure variations and trading-off reconstruction accuracy and representation parsimony using an MDL principle. We give experimental results that show the method can extract useful multi-scale skeleton representations at various abstraction levels. Future and ongoing work is exploring extensions to abstraction based on multiple input skeletons, alternative MDL priors and skeleton clustering methods, as well as multi-scale skeleton matching.

## References


[1] J. August, K. Siddiqi, and S. Zucker. Ligature instabilities in the perceptual organization of shape. *CVIU*, 76(3):231–243, 1999.

[2] G. Borgefors, G. Ramella, and G. S. di Baja. Hierarchical decomposition of multiscale skeletons. *PAMI*, 23(11):1296–1312, 2001.

[3] L. Costa and R. Cesar. *Shape Analysis and Classification: Theory and Practice*. CRC Press, 2001.

[4] P. Dimitrov, C. Phillips, and K. Siddiqi. Robust and efficient skeletal graphs. In *CVPR*, 2000.

[5] D. Macrini, A. Shokoufandeh, S. Dickinson, K. Siddiqi, and S. Zucker. View-based 3-d object recognition using shock graphs. 2002.

[6] R. Ogniewicz. Skeleton-space: a multiscale shape description combining region and boundary information. In *CVPR*, pages 746–751, 1994.

[7] S. Pizer, K. Siddiqi, G. Szekely, S. Damon, and S. Zucker. Multiscale medial loci and their properties. *IJCV*, 55(2-3):155–179, 2003.

[8] T. Sebastian, P. Klein, and B. Kimia. Shock-based indexing into large shape databases. In *ECCV*, pages 731–746, 2002.

[9] T. Sebastian, P. Klien, and B. Kimia. Recognition of shapes by editing their shock graphs. *PAMI*, 2004.

[10] A. Shokoufandeh, S. Dickinson, K. Siddiqi, and S. Zucker. Indexing using a spectral encoding of topological structure. In *CVPR*, pages 491–497, 1999.

[11] K. Siddiqi, S. Bouix, A. Tannenbaum, and S. Zucker. Hamilton-jacobi skeletons. *IJCV*, 48(3):215–231, 2002.

[12] K. Siddiqi, A. Shokoufandeh, S. Dickinson, and S. Zucker. Shock graphs and shape matching. *IJCV*, 30:1–24, 1999.

[13] A. Telea and J. van Wijk. An augmented fast marching method for computing skeletons and centerlines. In *IEEE VisSym*, pages 251–260, 2002.

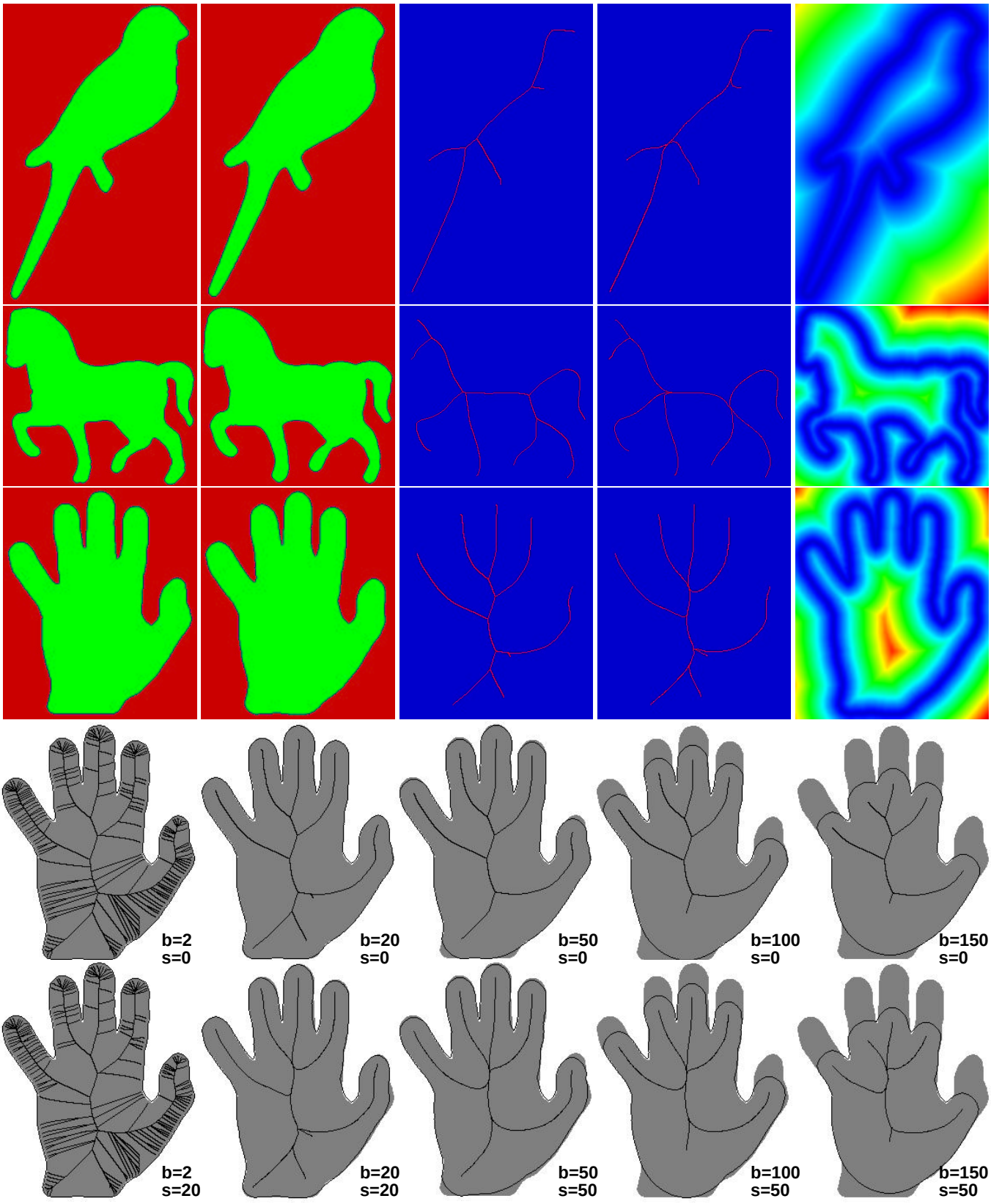


**Figure 2. Skeleton simplifications for different objects. First three rows, left to right: initial silhouette, simplified silhouette, skeleton after boundary simplification, skeleton after internal structure simplification, and distance transform of the simplified silhouette. Last two rows: initial silhouette (gray) and skeleton and simplified silhouette and skeleton (black) for different $s$ and $b$ values.**