SCIENCE CHINA

January 2012 Vol. 53 No. 1: 1–18
doi:

Skeleton pruning as trade-off between skeleton simplicity and reconstruction error

SHEN Wei[1], BAI Xiang[1]*, YANG XingWei[2] & LATECKI Longin Jan[2]

[1]*Department of Electronics and Information Engineering, Huazhong University of Science and Technology, Wuhan* 430074*, China,*
[2]*Department of Computer and Information Sciences, Temple University, Philadelphia, PA* 19122*, USA*

Received February 9, 2012; accepted May 10, 2012

Abstract Skeletons can be viewed as a compact shape representation in that each shape can be completely reconstructed from its skeleton. However, the usefulness of a skeletal representation is strongly limited by its instability. Skeletons suffer from contour noise in that small contour deformation may lead to large structural changes in the skeleton. A large number of skeleton computation and skeleton pruning approaches has been proposed to address this issue. Our approach differs fundamentally in the fact that we cast skeleton pruning as a trade-off between skeleton simplicity and shape reconstruction error. An ideal skeleton of a given shape should be the skeleton with a simplest possible structure that provides a best possible reconstruction of a given shape. To quantify this trade-off, we propose that the skeleton simplicity corresponds to model simplicity in the Bayesian framework, and the shape reconstruction accuracy is expressed as goodness of fit to the data. We also provide a simple algorithm to approximate the maximum of the Bayesian posterior probability which defines an order for iteratively removing the end branches to obtain the pruned skeleton. Presented experimental results obtained without any parameter tuning clearly demonstrate that the resulting skeletons are stable to boundary deformations and intra class shape variability.

Keywords skeleton, skeleton pruning, shape reconstruction

1 Introduction

The skeleton or *Medial Axis (MA)*, is a compact abstraction of visual shape, which has been widely used for object representation and recognition [1]. Blum's Medial Axis Transform (MAT) has been very influential in computer vision. It computes the centers and radii of maximal disks lying within the object, which touches the boundary with at least two points [2]. There have been many methods proposed for skeleton extraction. However, skeleton extraction is usually troubled by a well-known issue: the skeleton is unstable with respect to small boundary perturbations [3]. Due to the skeleton's sensitivity to boundary deformation and noise, researchers tried to design some stable skeletonization approaches [1,4,5,6,7,8,9] or smart pruning techniques [10,11,12,13,14]. The progress in skeleton extraction not only significantly improved the performance of shape recognition [15,16,17] but also made it possible to utilize skeletons in object detection and contour grouping in natural images [18,19,20]. This demonstrates that it is very important to obtain the skeleton with a clean structure that can represent the visual parts of an object [11].

*Corresponding author (email: xiang.bai@gmail.com)

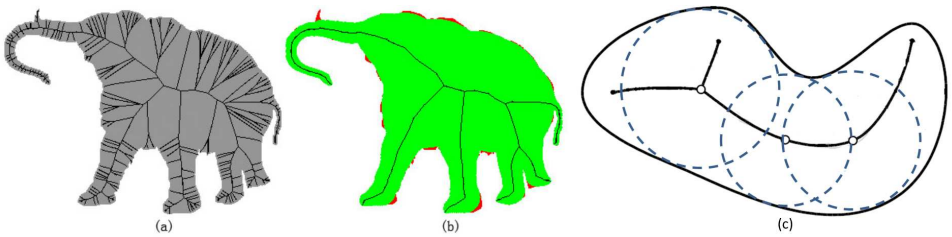

Figure 1: (a) An input shape and its medial axis. (b) The skeleton obtained by the proposed method. Its reconstruction error is marked with the red regions. (c) The medial axis and the maximal disks.

However, existing approaches to skeletonization or skeleton pruning require adjusting thresholds or other parameters for each shape or shape class separately, which usually is a tedious and manual process. Therefore, skeleton based approaches have never been used on large datasets such as the MPEG-7 Shape Dataset [21].

We overcome this problem by casting skeleton pruning as a trade-off between two properties of the skeleton: simplicity and shape reconstruction error. The trade-off is quantified by Bayesian probability, and a single parameter is taken as the weight factor for these two properties which is kept constant in all our experiments. An example of a skeleton obtained by the proposed approach is provided in Fig. 1(b), where the red regions represent the reconstruction error when the shape is reconstructed from the proposed skeleton. All the obtained skeletons are stable in the presence of intra class shape deformations as well as boundary noise as demonstrated by our experimental results.

The proposed approach begins with a connected skeleton composed of all MA points. The MA points are then pruned by iteratively removing skeleton end branches according to a certain order. A sequence of skeletons are generated under this order and we choose the skeleton corresponding to the optimal solution of the target function defined in Section 5 as the final output.

As pointed out by [10], the skeleton of a single connected shape that is useful for skeleton-based recognition should have the following properties:

1. it should preserve the topological information of the original object,
2. the position of the skeleton should be accurate,
3. it should be stable under small deformations,
4. it should contain the centers of maximal disks, which can be used for reconstruction of original object,
5. it should be invariant under Euclidean transformations such as rotations and translations, and
6. it should represent significant visual parts of objects.

The proposed skeletons satisfy all these properties and are computed automatically without any parameter tuning.

2 Related Work

Skeletonizaton approaches can be broadly classified into four categories: 1) thinning/peeling algorithms [22,23,24,25,26], 2) discrete domain algorithms based on the Voronoi diagram [27,28,29], 3) algorithms based on distance transform [30,31,32,33,34], and 4) algorithms based on mathematical morphology [35,36,37,38].

All the obtained skeletons are subjected to the skeleton's sensitivity to boundary deformations and many of them also include pruning methods along with the skeletonization. The earlier works [39,40,41,42] on skeleton-based shape matching cannot often be used in shape recognition for a large dataset due to the instability of skeletons and the limitations of skeletonization techniques. Recent progresses in skeletonization or skeleton pruning have brought a significant improvement for skeleton-based shape recognition. Aslan et al. [15] extract the disconnected skeletons in a coarse level, which are very good

fit for articulated shapes. Bai and Latecki [16] approximately transfer the graph matching process into sequence matching by using skeleton paths between skeleton endpoints as a representation, and their skeleton graphs are built based on the pruned skeletons obtained by [10]. Macrini et al. [17] posit a strategy to build a new kind of skeleton graph named Bone Graph by preprocessing the Shock Graph, which outperforms the Shock Graphs in the shape recognition task [42]. These approaches heavily depend on the quality of obtained skeletons as the object representation.

As an essential part of skeletonization algorithms, skeleton pruning algorithms usually appear in a variety of application-dependent formulations [43]. There are mainly two ways of pruning methods: (1) based on significance measures assigned to skeleton points [28,31,43], and (2) based on boundary smoothing before extracting the skeletons [43,44]. However, boundary smoothing may lead to disruptive effects. In particular, curvature flow boundary smoothing may cause the position of skeletons shift and has difficulty in distinguishing noise from low frequency shape information on the boundary [43]. A different kind of smoothing is proposed in [45]. A great progress has been made in the type (1) of pruning approaches that define a significance measure for skeleton points and remove points whose significance are low. [43] gives a complete analysis and compares such pruning methods. To the common significance measures of skeleton points belong propagation velocity, maximal thickness, radius function, axis arc length, the length of the boundary unfolded. Ogniewicz et al. [28] present a few significance measures for pruning hairy Voronoi skeletons without disconnecting the skeletons. Siddiqi et al. combine a flux measurement with the thinning process to extract a robust and accurate connected skeleton [5]. However, correcting the error in calculating the flux is limited by the pixel resolution and is also proportional to the curvature of the boundary evolution front. This makes the exact location of endpoints difficult. Torsello et al. [7] overcome this problem by taking into account variations of density due to boundary curvature and eliminating the curvature contribution to the error.

[10] presents a promising skeleton pruning method by dividing the contour into separate segments with the vertices from DCE (Discrete Curve Evolution). Since DCE does not change the topology, the pruned skeleton has the topology of the input skeleton. The idea of using contour partitioning [10] is later extended in skeletonization of wireless sensor networks [46]. Krinidis and Chatzis [47] localize the medial points iteratively starting from a boundary point set based on a physics-based deformable model. Recently, [13,14] introduce an interesting idea for skeleton punning that can prune a group of shapes from the same class in the same pruning process. The merit of their method is that pruning can benefit the shape information from different shapes and find the common structure. However, their method requires that the shapes in a group should be in the same orientation without serious changes of articulations.

3 Basic Skeleton Concepts

Before introducing the proposed approach, we give some definitions. According to Blum's definition of the medial axis (MA) [2], the MA or a skeleton S of a set D is the locus of the centers of maximal disks. A maximal disk of D is a closed disk contained in D that is interiorly tangent to the boundary ∂D and that is not contained in any other disk in D. Each maximal disc must be tangent to the boundary in at least two different points. As illustrated in Fig. 1(c), the exterior solid curve is the boundary of the set D, and the interior solid curve is the MA. Blue dash circles denotes the maximal discs, whose centers are soft dots. With every skeleton point $s \in S$ we also store the radius $r(s)$ of its maximal disk.

In this paper, we use term **medial axis (MA)** to denote the set M of all centers of maximal disks of a given set D while the **skeleton** S to denote any subset of M.

By Theorem 8.2 in [48], the skeleton S is a geometric graph, which means that S can be decomposed into a finite number of connected arcs, called skeleton branches, composed of points of degree two, and the branches meet at skeleton joints (or bifurcation points) that are points of degree three or higher.

The skeleton point having only one adjacent point is an **endpoint (the skeleton endpoint)**; the skeleton point having more than two adjacent points is a **junction point**. If a skeleton point is not an endpoint or a junction point, it is called a **connection point**. (For skeletons in digital images, we assume the curves of the skeleton are one-pixel wide.)

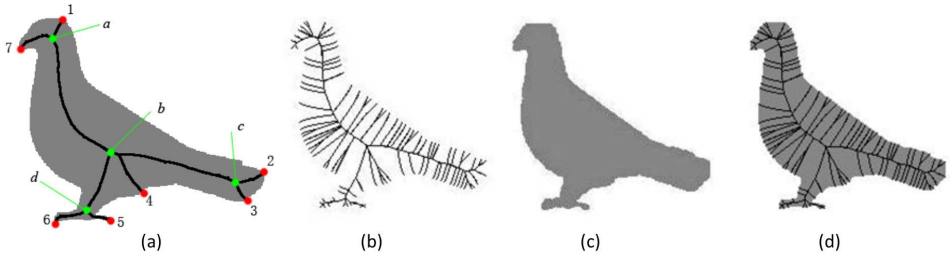

Figure 2: (a) The endpoints (red) and junction points (green) on a bird's skeleton. The reconstruction (c) of the Medial Axis (b) is equal to the original shape in (d).

A skeleton end branch is part of the skeleton between a skeleton endpoint and the closest junction point. Let $l_i (i = 1, 2, \ldots, N)$ be the endpoints of a skeleton S. For each endpoint l_i, $f(l_i)$ denotes the nearest junction point. Formally, an **end branch** $E(l_i, f(l_i))$ is the shortest skeleton path between l_i and $f(l_i)$. For example, in Fig. 2(a), arc from point 1 to point a is a skeleton end branch: $E(1, f(1)) = E(1, a)$. The arc from point a to point b is not an end branch; it is a skeleton (inner) branch. Note that point a is the nearest junction point of two endpoints (1 and 7).

We define **end path** $P(l, o)$ as the shortest skeleton path between an endpoint l and a skeleton point o named seed point. We choose as seed point o a skeleton point whose radius of the corresponding maximal disk is maximal.

In this paper we relax the Blum's definition of the maximal disk. While Blum's definition of maximal disks is restricted only to medial axis points, we extend it to all interior points of the shape D. For each point s in the interior of D, a closed ball $B(s, r)$ centered at s with radius r is a maximal disk of D if $B(s, r) \subseteq D$ and $\partial B(s, r) \cap \partial D \neq \emptyset$, where ∂A yields the boundary of the set A. In particular, we do not require that a maximal disc must be tangent to the boundary in at least two different points. We use Blum's definition of the maximal disk to define the medial axis. Otherwise, we use the relaxed definition in the rest of this paper.

We define a shape space $\mathcal{D}$ as the set of all compact planar sets D such that the boundary ∂D is composed of a finite number of simple closed polygonal curves. Sometimes we call such a D a shape. We define a set of skeletons $\mathcal{S}$ as the set of sets S such that S is a finite union of open polygonal curves that intersect only at their vertices. This implies that every junction point must be a vertex, but we may have additional vertices to allow the skeleton to bend. We observe that we also relax a definition of the skeleton in that a given skeleton S is not required to be linked to any shape D. Of course, in the remainder of this paper we will relate the skeletons to the shapes, but a skeleton S is a finite union of 1D polygonal open curves.

Let a skeleton S be a subset of a shape D. Let $r(s)$ denote the radius of the maximal disk $B(s, r(s))$ centered at a point $s \in S$. The **reconstruction** of D from skeleton S is denoted by $R(S)$

$$R(S) = \bigcup_{s \in S} B(s, r(s)) \tag{1}$$

If a skeleton S is not a subset of D, then $R(S) = \emptyset$. The quality of the reconstruction will allow us to relate the skeletons to shapes in the next section. A special case is illustrated in Fig. 2. If S is the MA of D, then we can exactly reconstruct the original shape D from S, i.e., $R(S) = D$. However, MA of a digital shape in images usually contains spurious branches due to boundary noise and deformation. Our goal is to find a skeleton S that can be seen as a simplification of MA by keeping only skeleton branches that are relevant for the shape representation, while at the same time S should contain enough information to provide a reasonable reconstruction of the shape like the bird skeleton in Fig. 2(a).

4 Trade-off Quantification

For a given shape D, the ideal skeleton S should: (i) contain no spurious skeleton branches; (ii) provide the best reconstruction of the given shape. In practice, there is a trade-off between (i) and (ii): we desire to seek a skeleton that has as simple structure as possible, whilst the reconstruction error of the skeleton to the given shape is as small as possible. We introduce below intuitive formulas used in Bayesian framework to quantify the trade-off. Feldman and Singh [49] have provided cognitive motivation rooted in human visual perception that the most likely skeleton for a given shape can be obtained by maximizing $p(S|D)$. By Bayes' rule, this probability density function (pdf) can be transformed to

$$p(S|D) = \frac{p(D|S)p(S)}{p(D)} \propto p(D|S)p(S) \tag{2}$$

Hence the most probable skeleton S should explain the shape D in the most plausible way under a generative model given by the conditional pdf $p(D|S)$ and at the same time should be the most likely skeleton according to the prior $p(S)$. The prior $p(S)$ represents the skeleton simplicity. The Bayesian framework allows us for a natural and intuitive combination of likelihood and prior. Consequently, (2) seeks the simplest skeleton that can best explain a given shape D.

For a fixed shape $D \in \mathcal{D}$, we can drop the denominator in (2). It is equal to $p(D) = \int p(D|U)P(U)d(U)$ and stays constant for any S in the numerator, where U varies over the set of all possible skeletons $U \in \mathcal{S}$.

While Feldman and Singh [49] provide a general framework, they do not provide any complete algorithm for skeleton computation according to (2). They leave many computational problems open, because their formulas for prior and conditional pdfs lead to challenging combinatorial issues. In this paper we propose to use exponential family distributions to define both factors in (2).

A key feature of the definition of a skeleton is that the shape can be reconstructed from it [2]. We propose to measure how well a given skeleton S explains shape D as the area difference between D and the shape $R(S)$ reconstructed from S, i.e. $\Lambda(D - R(S))$, where $\Lambda(\cdot)$ denotes the area measured by pixels. We observe that the area of the set difference is very robust to boundary deformations. To be specific, we define

$$\begin{cases} p(D|S) \propto \exp(-\alpha \dfrac{\Lambda(D - R(S))}{\Lambda(D)}), S \subset D \\ p(D|S) = 0, S \nsubseteq D \end{cases}, \tag{3}$$

where α is a normalization factor that will be defined in Section 6. Thus, a given skeleton is more likely, the better is the reconstruction $R(S)$ measured by area difference.

In particular, if $S = M$, where M is the medial axis of D, we have $\Lambda(D - R(M)) = 0$, since $D = R(M)$. However, when D is given by a binary image, we need to modify M to get an ideal S, since M usually contains many spurious branches and may be deformed due to contour noise and digitization artifacts.

The prior pdf $p(S)$ over the skeletons of D is intended to measure the skeleton simplicity. Intuitively, it makes sense that the shorter skeleton S, the simpler is the structure of S. Therefore, we propose

$$p(S) = \exp(-\log(\Gamma(S) + 1)), \tag{4}$$

where $\Gamma(\cdot)$ denotes the normalized curve length and adding the constant one assures that $\log(\Gamma(S)+1) \geqslant 0$. We use log function to scale the measure of curve length, because as it is shown by the popular shape descriptor *Shape Context* [50], scaling the distance with a log factor adds stability to shape representation. In our framework, it is also possible to make the simplicity of a skeleton S dependent on the number of junction point and on the amount of its bending. However, as we will see in our experimental results, expressing the simplicity as the function of length only is sufficient for obtaining high quality skeletons.

By plugging (3) and (4) to (2), we obtain our target function

$$p(S|D) \propto p(D|S)p(S) \propto \exp(-\alpha \frac{\Lambda(D - R(S))}{\Lambda(D)}) \exp(-\log(\Gamma(S) + 1)). \tag{5}$$

Thus, a skeleton S is more likely, the better it explains the shape D, which is measured by normalized area difference, and the simpler S is, which is measured by normalized log length of S. Consequently, by maximizing (5) we obtain the simplest skeleton S that can best explain shape D. The skeleton simplicity corresponds to model simplicity in Bayesian framework, and the shape reconstruction accuracy corresponds to goodness of fit to the data.

5 Pruning Process

In this section, we introduce a simple algorithm to approximate the maximum of (5) which defines an order for iteratively removing the end branches to obtain the pruned skeleton.

Maximizing (5) is equivalent to minimizing

$$-\log p(S|D) \propto \alpha \frac{\Lambda(D - R(S))}{\Lambda(D)} + \log(\Gamma(S) + 1), \tag{6}$$

We denote $\lambda(S|D) = -\log p(S|D)$, the first summand by AR for area ratio and the second by LR for length ratio. Thus, we re-express (6) as

$$\lambda(S|D) \propto \alpha AR(D|S) + LR(S), \tag{7}$$

where $AR(D|S) \geqslant 0$ and $LR(S) \geqslant 0$. α is the weight factor between goodness of fit of $R(S)$ to D measured by $AR(D|S)$ and model simplicity measured by $LR(S)$. Our goal is to find the minimum of (7):

$$S^* = \arg\min \lambda(S|D) \tag{8}$$

However, searching for the optimal solution of (8) is a computationally challenging problem, since there is a huge number of potential skeletons S that need to be examined. We know neither how many vertices the skeleton S should have nor their positions in the shape D. Moreover, the minimization result depends on the normalization factor α. For example, if $\alpha = 0$, then a single point skeleton of length zero is a solution, and the solution is not unique. In fact any point $s \in D$ yields the optimal value $\lambda(S|D) = 0$ for $S = \{s\}$. Hence, we also face the problem of the uniqueness of the solution. Consequently, (8) is an ill-posed problem.

In the following we will introduce additional constraints on the skeleton that will allow us to introduce a well-posed problem that seems to provide a reasonable refinement of (8). First we assume $\alpha > 0$ (the actual value will be determined later), and observe that the larger α is, the more important is the reconstruction accuracy in comparison to the skeleton length. Second, as we stated above if $S = M$, where M is the medial axis of D, $AR(D|S) = 0$. However, this leads to a noisy skeleton with many spurious branches. Our key observation is as follows: for every skeleton U not contained in M, there exists at least one skeleton $S \subseteq M$ with a better reconstruction accuracy and shorter or equal length. The following theorem makes this statement more precise:

Theorem 1. Let M be the MA of shape D. For every skeleton $U \not\subset M$, there exists a skeleton $S \subseteq M$ such that $\Gamma(S) \leqslant \Gamma(U)$ and $AR(D|S) \leqslant AR(D|U)$. Consequently, $\lambda(S|D) \leqslant \lambda(U|D)$.

Proof: The theorem follows from the fact that we can project any point in $u \in U \setminus M$ to the closest point $p(u) \in M$. We observe that $r(u) \leqslant r(p(u))$ by the definition of MA as composed of the centres of maximal disks. Since $r(u) \leqslant r(p(u))$, we obtain $\bigcup_{u \in U \setminus M} B(p(u), r(p(u)))$ has better or equal reconstruction accuracy to $\bigcup_{u \in U \setminus M} B(u, r(u))$.

Let $S = p(U)$. This means that $p(u) = u$ if $u \in M$ and $p(u) \neq u$ if $u \in U \setminus M$. Therefore, S has better or equal reconstruction accuracy than U.

Since p is a function, S cannot have more points (pixels) than U. Hence $\Gamma(S) \leqslant \Gamma(U)$, which completes the proof.

This theorem motivates our constraint to restrict the search for the skeleton S^* to $S \subseteq M$. We can further constrain the solutions of (8) to skeletons with desirable topological and geometric properties. It is important to require S to have the same topology as D. Given that M has the same topology as D, which is guaranteed in the continuous domain, and is true for many discrete medial axis algorithms, e.g., [32], it is sufficient to ensure that S has the same topology as M in order to obtain that S has the same topology as D. We ensure that S has the same topology as M by removing only skeleton end branches from M. We also require the skeleton end branches to terminate on the boundary of D. The above discussion motivates the following three constraints

- $S \subseteq M$.
- S has the same topology as M.
- End branches of S terminate on the boundary of D.

Due to these constraints, solving (8) can be performed by removing skeleton end branches from M. Therefore, our goal is to find a sequence of end branches

$$E_0, \ldots, E_n$$

whose removal minimizes (6). We denote the corresponding sequence of obtained skeletons by

$$S^{(0)} = M \quad \text{and} \quad S^{(i+1)} = S^{(i)} - E_i.$$

We derive now a recursive process that allows us to find such a sequence. A key step is the selection of the end branch E_i for obtaining $S^{(i+1)}$ from $S^{(i)}$. Given a skeleton $S^{(i)}$, we denote all its end branches by

$$E_j^{(i)} (j = 1, 2, ..., m^{(i)}).$$

We define the weight $w_j^{(i)}$ for each $E_j^{(i)}$ as

$$w_j^{(i)} = \alpha \frac{\Lambda(D - R(S^{(i)} - E_j^{(i)}))}{\Lambda(D)} + \log(\Gamma(S^{(i)} - E_j^{(i)}) + 1). \tag{9}$$

The end branch with the smallest weight

$$w_i = \min\{w_j^{(i)} | j = 1, 2, \ldots, m^{(i)}\} \tag{10}$$

is taken as the branch E_i to be removed, and we set $S^{(i+1)} = S^{(i)} - E_i$. In other words, $S^{(i+1)}$ is obtained from $S^{(i)}$ by removing a branch E_i with the smallest weight. We continue this process until all branches are removed from M. In fact, in our pruning process, every step is greedy, and we remove the end branch which makes the value of the target function (7) decrease fastest. Our algorithm defines a pruning order which generates a sequence of skeletons $\{S^{(i)}\}$. Finally, we choose the skeleton $\hat{S}$ from them which satisfies

$$\hat{S} = \arg \min_{S \in \{S^{(i)}\}} \lambda(S|D). \tag{11}$$

Thus, instead solving the ill-posed problem in (8), we propose to solve the well-posed problem stated in (11). In Fig. 3, we show the curves of our target function (7) related to the sequences of skeletons $\{S^{(i)}\}$. The skeletons corresponding to the optimal solution of (11) are shown below the curves.

The pseudo code of the proposed skeleton pruning algorithm is given in the Fig. 4.

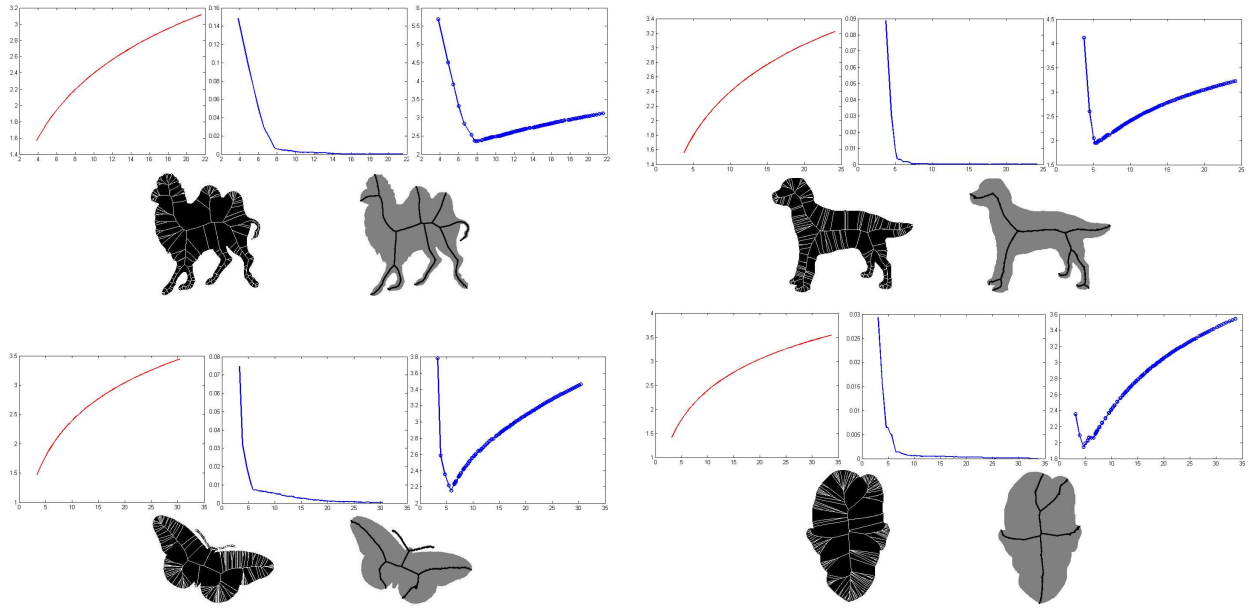

Figure 3: The plots in the first and second column show the curves of $LR(x)$ in red and of $AR(x)$ in blue as functions of the skeleton length x, respectively. The plots in the third column show our target function (7). Original medial axes and skeletons corresponding to the optimal solution of function (11) are shown below each triplet of plots.

Procedure SkeletonPruning(Input M, Output $\hat{S}$)
01. Initialize $S = M$, $\mathcal{S} = \{S\}$.
02. **While** The number of the end branches in S is larger than 2
03. Let L be the set of end points of S and $\{E(L, o)\}$ be the set of end branches of S.
04. **For** all $l \in L$
05. Compute the weights w_l of end branch $E(l, f(l))$ by Eq. (9).
06. **End**
07. Choose the minimum weight $w_{min} = \arg\min_{l \in L} w_l$.
08. Remove the end branch E_{min} with the weight w_{min}: $S = S - E_{min}$
09. $\mathcal{S} = \mathcal{S} \cup S$.
10. **End**
11. $\hat{S} = \arg\min_{S \in \mathcal{S}} \alpha AR(S) + LR(S)$

Figure 4: The pseudo-code for the proposed algorithm.

6 The choice of the normalization factor α

The normalization factor α is defined as

$$\alpha = \beta \log(\frac{\Gamma(M)}{\varpi(M)}), \tag{12}$$

where $\varpi(M)$ denotes the average length of end paths of M emanating from the seed point o and β is a constant. $\frac{\Gamma(M)}{\varpi(M)}$ represents the prior of the shape complexity. Intuitively, the lager $\frac{\Gamma(M)}{\varpi(M)}$ is, the more skeleton branches are contained in M, and the more complex the shape is. The so defined α balances the skeleton simplicity and the shape reconstruction error between shapes with different complexities. In all experiments β is set to 9 which is obtained after a trial and error testing. As our experimental results

demonstrate that it is possible to keep β for all shapes independent of shape complexity and the amount of boundary noise.

7 Experimental Results

In this section, we evaluate the performance of the proposed method in three parts: 1) stability to shape variances and boundary noise, 2) quality of skeletons measured by reconstruction error and 3) analysis and comparison to other recent methods.

We stress that we use a fixed parameter $\beta = 9$ in all the experiments. All the original skeletons were generated by the algorithm in [32]. It computes skeletons by utilizing the distance transform. The skeletons obtained by [32] can be approximately considered as the medial axes M for each input shape D. In particular, M has the same topology as D and all end branches of M terminate on the contour of D.

7.1 Stability of the proposed method

To demonstrate that our method is stable to shape variance, we apply our method to two well-known shape datasets: Kimia's dataset [40] and Tari's dataset [6], which have also been very popular in skeleton-based shape recognition. Kimia's dataset [40] includes 216 shapes with 18 classes (12 shapes per class), which is a subset of MPEG-7 Shape dataset [21]. There is significant boundary variance among the shapes from the same classes in Kimia's dataset. We show all the results of Kimia's dataset in Fig. 11, and we can observe that the obtained skeletons are not only clean but also stable for each of the 18 shape classes.

We also show our results on Tari's dataset [6] in Fig. 5, which consists of 56 shapes with 14 classes (4 shapes per class). In this dataset, the shapes from the same class may differ by serious articulations, and some shapes are with obvious boundary noise. We still use the same setting: $\beta = 9.0$ for this dataset, and the obtained skeletons demonstrate that our method is stable to articulation changes.

Furthermore, Fig. 6 demonstrates that our skeletons are robust to geometric transformations, such as rotation, scaling, and affine transformations. We use the same setting as before: $\beta = 9.0$.

Our skeleton is robust to boundary noise as demonstrated in Fig. 7. We artificially inserted boundary noise, which is either Gaussian or uniform, to the images in the first column.

7.2 Quantitative Evaluation of Skeleton Quality

The skeleton of high quality should represent significant visual parts of objects. To evaluate this property of pruned skeletons, we compute the reconstruction error ratio (RER) defined by the AR function. The quality of the skeleton is not solely determined by RER, since the reconstruction error of the original medial axis is zero. Therefore, we also compute the skeleton simplicity (SS) defined by the LR function. We report the statistics of RER and SS for our skeletons and skeletons in Bai et al. [10] on Kimia's dataset [40] and Tari's dataset [6] in Table 1. We observe that our method yields simpler skeletons (lower SS values) on both datasets. On Tari's dataset, we also have better RER. Our RER on Kimia's dataset is slightly higher than that of [10], which is needed in order to obtain simpler skeletons. Bai's skeletons are computed by setting a fix threshold at $T = 1$. We also demonstrate several pruned skeletons of the shapes from MPEG-7 dataset in Fig. 8. The RER values of these skeletons as well as of the skeletons in Table 1 are all very small which shows the high quality of the obtained skeletons. This shows that our method yields skeletons with a simple graph structure with only minimal loss in reconstruction quality, which is usually less than one percent of the shape area.

7.3 Analysis and Comparison

In the past, almost all approaches to skeleton pruning require manual parameter tuning for each shape or shape class. However, recent work [47] has tried to obtain stable skeletons with a fixed parameter setting. We compare our method with current state-of-the-art methods [10,47,51].

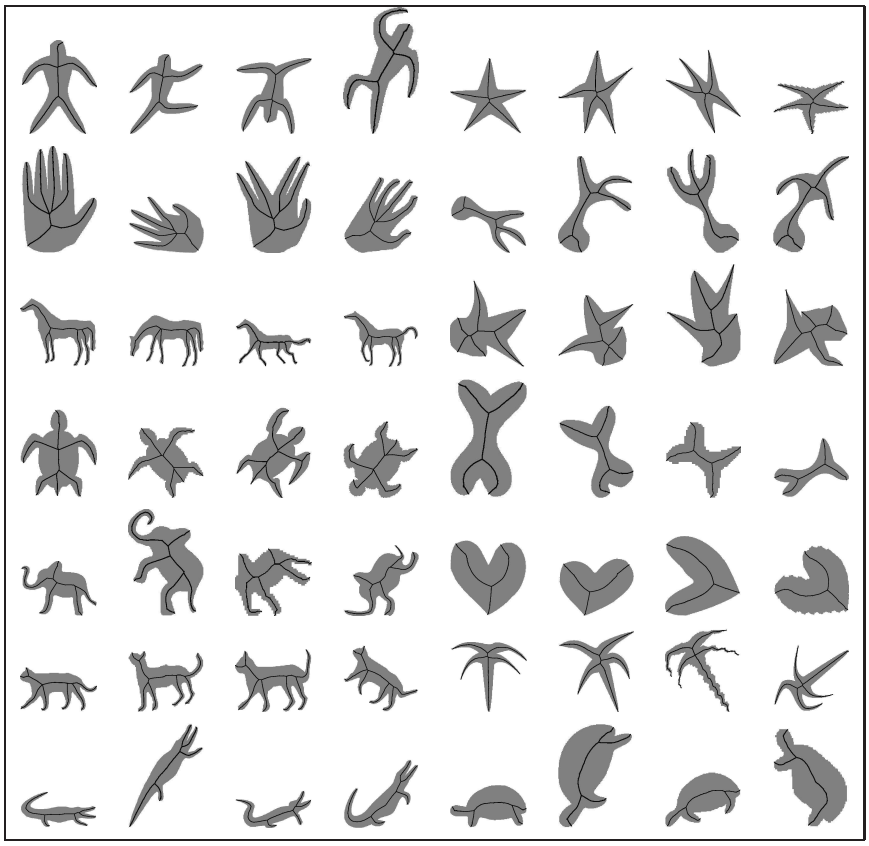

Figure 5: Our skeletons for all shapes in the Tari's dataset [6].

Table 1: The mean and the standard deviation of the reconstruction error ratio and skeleton simplicity on Kimia's dataset [40] and Tari's dataset [6].

		Ours		[10]	
		Mean	Std. dev.	Mean	Std. dev.
Kimia	RER	0.0087	0.0056	0.0068	0.0063
	SS	1.4665	0.2262	1.6452	0.1923
Tari	RER	0.0106	0.0097	0.0193	0.0709
	SS	1.5866	0.1838	1.7071	0.1194

[47] used a fixed parameter to extract stable skeletons of a small dataset shown in Fig. 10(b). For comparison, we test our algorithm on the same data set with the a fixed parameter $\beta = 9.0$. Our skeletons are shown in Fig. 10(a). Both methods yield visually appealing skeletons. However, the skeletons of the method in [47] are not symmetric to the boundary, e.g., see the skeleton branch at the horse's head in Fig. 10(b), and are sensitive to boundary noise, e.g., the skeleton of the human wiggles a lot due to the boundary noise. In contrast, our algorithm starts with the Blum's medial axis, and consequently, is symmetric to the boundary. In several shape matching method [10], the sequence of radii of the maximal disks at the points on the skeleton path is used as a feature to discriminate shapes. If the skeleton is not symmetric to the boundary, the radii of the skeleton points are inaccurate, which could influence the matching result. Therefore, skeletons that are symmetric to the boundary are strongly preferred.

In order to demonstrate the accuracy of the skeleton quantitatively, we use the error measure of skeletons proposed in [47] in our evaluation. We apply the skeletonization approach to a simple shape

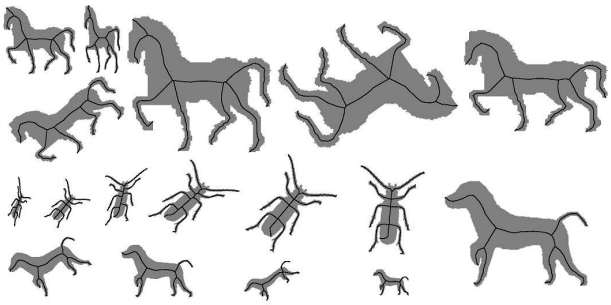

Figure 6: Our skeletons are stable to affine transformations.

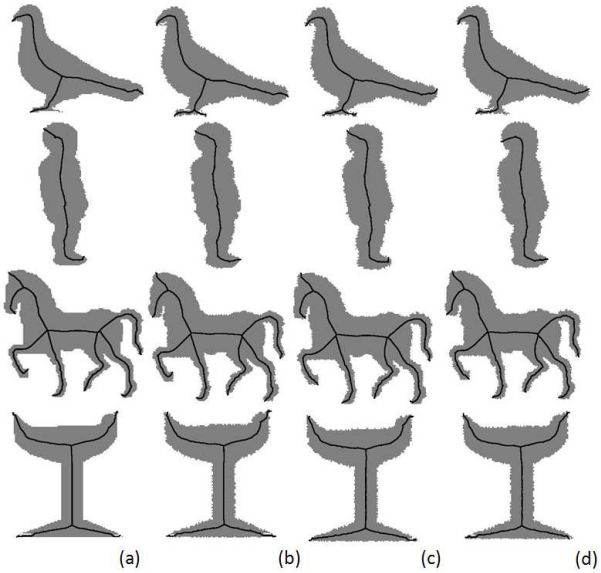

Figure 7: Our skeletons on shapes with noisy boundaries. (a) Initial images. (b),(c) Images with Gaussian boundary noise of (0,1) and (0, 1.5). (d) Images with uniform noise (1.2) inserted to original boundaries.

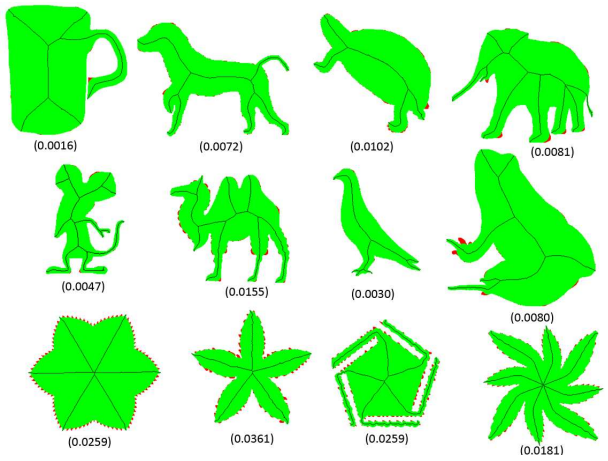

Figure 8: Skeletons obtained by proposed method. The shapes are green and the reconstruction errors are marked red. The value of the reconstruction error ratio is given in the brackets below each shape.

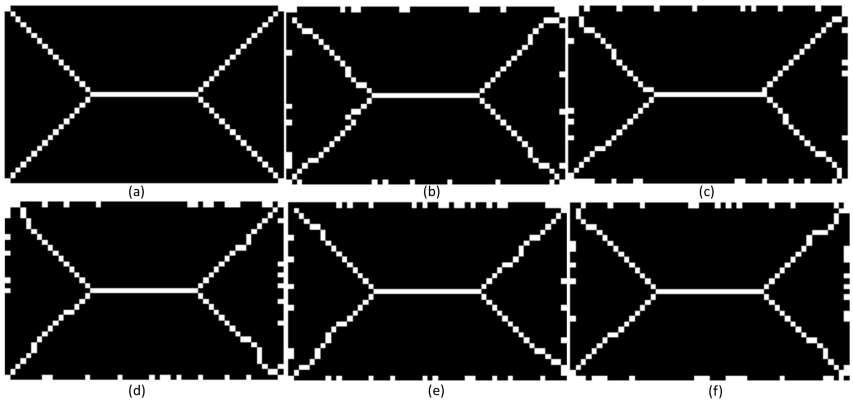

Figure 9: Skeletons of rectangles. (a) the original rectangle, (b) the rectangle with Gaussian noise (0, 1.0) on its boundary, (c) with Gaussian noise (0, 1.5), (d) with uniform noise (1.0), (e) with uniform noise (1.2), (f) with uniform noise (1.5).

Table 2: Skeletonization error for rectangles presented in Fig. 9

	Obtained by method in [47]	Obtained by method in [10]	Obtained by method in [51]	Obtained by our method
Average error pixel	0.31	0.28	0.25	**0.24**

(a rectangular region) with known skeleton (Fig. 9a). Then, we compare the skeletonization results of this shape for a variable amount of boundary noise (noisy skeletons)(Fig. 9b-f). For fairness, we test the same rectangles with the noise of the same intensity as in [47]. The error between a noisy skeleton S and the ground-truth skeleton G is computed as the mean square error between their points:

$$Err(S,G) = \frac{1}{N}\sum_{i=1}^{N}\sqrt{(S_x(i)-G_x(i))^2+(S_y(i)-G_x(i))^2}, \tag{13}$$

where N is the number of skeleton points in S, $(S_x(i), S_y(i))$ are the coordinates of the ith skeleton point in S and $(G_x(i), G_y(i))$ is the closest point in G to the point $S_x(i), S_y(i)$. The average error of the five types of noisy skeletons compared to the ground-truth skeleton in Fig. 9 is shown in Table 2. This evaluation demonstrates that our skeletons are more accurate and more robust to noise than the methods in [10,47,51].

We give another comparison to the method in [10] by using a quantitative way to evaluate the quality of the pruned skeletons by measuring the performance of shape retrieval. We use a recent skeleton matching method [16] to perform shape retrieval based on the skeletons extracted by [10] and the proposed algorithm, respectively. For the method in [10], we use a fixed stop threshold $T = 1$ to compute all the skeletons of the shapes in Kimia dataset. The retrieval results based on the two different skeleton pruning methods are shown in Table 3. For each query shape, we check whether the 11 closest matches are in the same category as the query. In Table 3, the retrieval result is summarized as the number of 11 closest matches that fall into the correct category. Therefore, the best possible number for each column is 216. We observe that the retrieval results based on the skeletons obtained by the proposed method are better, which demonstrates a superior quality of the obtained skeletons.

7.4 Experimental Details

Our pruning algorithm has high time cost, since it computes the reconstruction in every iterative step. Its average time for obtaining a pruned skeleton from a 512×512 image is about 7 minutes on a standard

Table 3: Retrieval results on Kimia's dataset with Path Similarity [16] based on the skeletons pruned by two different methods.

	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	11th
Pruned skeletons obtained by [10]	216	216	215	216	212	210	210	205	204	189	173
Proposed method	216	216	215	216	213	210	210	207	205	191	178

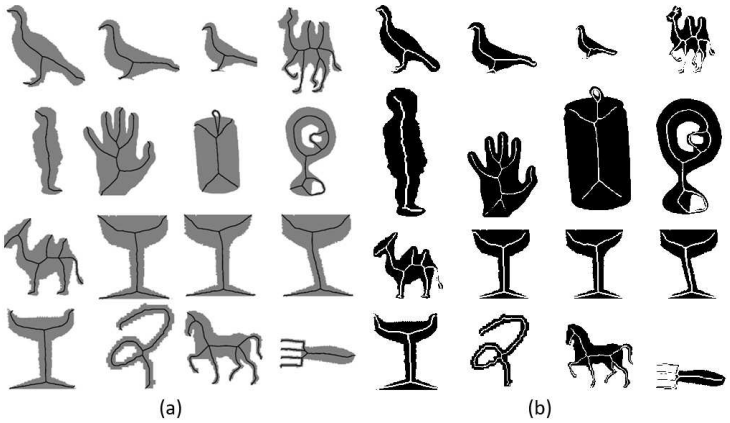

Figure 10: (a) Skeletons obtained by the proposed algorithm, and (b) skeletons obtained by [47].

PC. Our method is much slower than Bai's method [10] whose time cost is only about 28 seconds under the same conditions. Therefore, to obtain the experimental results presented in subsection 7.1 and 7.3, we used the following strategy to increase speed: First, we pruned the skeletons with the pruning algorithm in [10] to a coarse level; then, we continued with the proposed algorithm. In this way, the average time cost has been reduced to about 34 seconds, and the obtained skeletons are virtually indistinguishable.

8 Conclusion

In this paper, we establish a Bayesian framework for computing skeletons with a trade-off between reconstruction accuracy and skeleton simplicity. Such skeletons are obtained by iteratively removing the skeleton end branches with the smallest weight related to the trade-off quantification. Our algorithm requires only one parameter that balances the reconstruction accuracy and the skeleton simplicity. The parameter was fixed in all our experiments, which was remarkable, since we applied our algorithm to a large number of shapes. The results on three different datasets, Kimia's dataset [40], Tari's dataset [6], and Krinidis' dataset [47], demonstrate that our method is very stable to shape deformation and boundary noise. Consequently, it can be widely used in various applications, including skeleton-based shape recognition on large datasets.

Acknowledgements

This work was supported by the National Natural Science Foundation of China No. 60903096, National Natural Science Foundation of China 61222308, and Fundamental Research Funds for the Central Universities HUST 2011TS110)

References

1 Siddiqi K, Pizer S. Medial representations: Mathematics, algorithms and applications. Springer, 2008.
2 Blum H. Biological shape and visual science (part I). Journal of Theoretical Biology, 1973, 38:205–287
3 August J, Siddiqi K, Zucker S. Ligature instabilities and the perceptual organization of shape. Computer Vision and Image Understanding, 1999, 76(3):231–243
4 Dimitrov P, Phillips C, Siddiqi K. Robust and efficient skeletal graphs. In: CVPR, 2000.
5 Siddiqi K, Bouix S, Tannenbaum A, et al. Hamilton-Jacobi skeletons. International Journal of Computer Vision, 2002, 48(3):215–231
6 Aslan C, Tari S. An axis-based representation for recognition. In: ICCV, 2005.
7 Torsello A, Hancock E. Correcting curvature-density effects in the Hamilton-Jacobi skeleton. IEEE Trans. Image Processing, 2006, 15(4):887–891
8 Borgefors G, Ramella G, di Baja GS. Hierarchical decomposition of multiscale skeleton. IEEE Trans. Pattern Analysis and Machine Intelligence, 2001, 13(11):1296–1312
9 Katz R, Pizer S. Untangling the blum medial axis transform. Int Journal Computer Vision, 2003, 55:139–153
10 Bai X, Latecki L, Liu W Y. Skeleton pruning by contour partitioning with discrete curve evolution. IEEE Trans. Pattern Analysis and Machine Intelligence, 2007, 29:449–462
11 Bai X, Latecki L. Discrete skeleton evolution. In: EMMCVPR, 2007.
12 Eede M, Macrini D, Telea A, et al. Canonical skeletons for shape matching. In: ICPR, 2006.
13 Ward A, Hamarneh G. Gmat: The groupwise medial axis transform for fuzzy skeletonization and intelligent pruning. In: Technical Report, School of Computing Science, Simon Fraser University, 2008.
14 Ward A, Hamarneh G. The groupwise medial axis transform for fuzzy skeletonization and pruning. IEEE Trans. Pattern Analysis and Machine Intelligence, 2010, 32(6):1084–1096
15 Aslan C, Erdem A, Erdem E, et al. Disconnected skeleton: Shape at its absolute scale. IEEE Trans. Pattern Analysis and Machine Intelligence, 2008, 30:2188–2203
16 Bai X, Latecki L. Path similarity skeleton graph matching. IEEE Trans. Pattern Analysis and Machine Intelligence, 2008, 30:1282–1292
17 Macrini D, Siddiqi K, Dickinson S. From skeletons to bone graphs: Medial abstraction for object recognition. In: CVPR, 2008.
18 Bai X, Wang X G, Latecki L, et al. Active skeleton for non-rigid object detection. In: ICCV, 2009.
19 Levinshtein A, Sminchisescu C, Dickinson S. Multiscale symmetric part detection and grouping. In: ICCV, 2009.
20 Adluru N, Latecki LJ, R. Lakämper R, et al. Gross A. Contour grouping based on local symmetry. In: ICCV, 2007.
21 Latecki L, Lakämper R, Eckhardt U. Shape descriptors for non-rigid shapes with a single closed contour. In: CVPR, 2000.
22 Arcelli C, di Baja GS. A width-independent fast thinning algorithm. IEEE Trans. Pattern Analysis and Machine Intelligence, 1985, 7:463–474
23 Pudney C. Distance-ordered homotopic thinning: A skeletonization algorithm for 3d digital images. Computer Vision and Image Understanding, 1998, 72(3):404–413
24 Leymarie F, Levine M. Simulating the grassfire transaction form using an active contour model. IEEE Trans. Pattern Analysis and Machine Intelligence, 1992, 14(1):56–75
25 Golland P, Grimson E. Fixed topology skeletons. In: CVPR, 2000.
26 Tang Y, You X. Skeletonization of ribbon-like shapes based on a new wavelet function. IEEE Trans. Pattern Analysis and Machine Intelligence, 2003, 25(9):1118–1133
27 Brandt J, Algazi V. Continuous skeleton computation by voronoi diagram. CVGIP: Image Understanding, 1992, 55:329–338
28 Ogniewicz R, Kübler O. Hierarchic voronoi skeletons. Pattern Recognition, 1995, 28(3):343–359
29 Mayya N, Rajan V. Voronoi diagrams of polygons: a framework for shape representation. In: CVPR, 1994.
30 Arcelli C, di Baja GS. Euclidean skeleton via center-of-maximal-disc extraction. Image and Vision Computing, 1993, 11:163–173
31 Malandain G, Fernandez-Vidal S. Euclidean skeletons. Image and Vision Computing, 1998, 16:317–327
32 Choi WP, Lam KM, Siu WC. Extraction of the euclidean skeleton based on a connectivity criterion. Pattern Recognition, 2003, 36:721–729
33 Kimmel R, Shaked D, Kiryati N, Bruckstein A. Skeletonization via distance maps and level sets. Computer Vision and Image Understanding, 1995, 3:382–391
34 Ge Y, Fitzpatrick J. On the generation of skeletons from discrete euclidean distance maps. IEEE Trans. Pattern Analysis and Machine Intelligence, 1996, 18(11):1055–1066
35 Meyer F. Skeletons and perceptual graphs. Signal Processing, 1989, 16:335–363
36 Maragos P, Schafer R W. Morphological skeleton representation and coding of binary images. IEEE Trans. Acoustics,

Speech, and Signal Processing, 1986, 34(5), 1228–1244
37 Goutsias J, Schonfeld D. Morphological representation of discrete and binary images. IEEE Trans. Pattern Analysis and Machine Intelligence, 1991, 39(6): 1369–1379
38 Kresch R, Malah D. Morphological reduction of skeleton redundancy. Signal Processing, 1994, 38: 143–151 (1994)
39 Zhu S, Yuille A. Forms: A Flexible object recognition and modeling system. Int Journal Computer Vision, 1996, 20(3):187–212
40 Sebastian T, Klein P, Kimia B. Recognition of shapes by editing their shock graphs. IEEE Trans. Pattern Analysis and Machine Intelligence, 2004, 26(5):550–571
41 Liu T, Geiger D, Kohn R. Representation and self-similarity of shapes. In: ICCV, 1998.
42 Siddiqi K, Shokoufandeh A, Dickinson S, et al. Shock graphs and shape matching. International Journal of Computer Vision, 1999, 35(1):13–32
43 Shaked D, Bruckstein A. Pruning medial axes. Computer Vision and Image Understanding, 1998, 69(2):156–169
44 Mokhtarian F, Mackworth A. A theory of multiscale, curvature-based shape representation for planar curves. IEEE Trans Pattern Analysis and Machine Intelligence, 1992, 14:789–805
45 Gold C, Thibault D, Liu Z. Map generalization by skeleton retraction. In: ICA Workshop on Map Generalization, 1999.
46 Jiang H B, Liu W P, Wang D, et al. Case: Connectivity-based skeleton extraction in wireless sensor network. In: INFOCOM, 2009.
47 Krinidis S, Chatzis V. A skeleton family generator via physics-based deformable models. IEEE Trans. Image Processing, 2009, 18(1):1–11
48 Choi H, Choi S, Moon H. Mathematical theory of medial axis transform. Pacific Journal of Math, 1997, 181(1):57–88
49 Feldman J, Singh M. Bayesian estimation of the shape skeleton. Proceeding of National Academy of Sciences of the United States of America, 2006, 103(47):18,014–18,019
50 Belongie S, Malik J, Puzicha J. Shape matching and object recognition using shape contexts. IEEE Trans Pattern Analysis and Machine Intelligence, 2002, 24(4):509–522
51 Shen W, Bai X, Hu R, Wang H, Latecki L. Skeleton growing and pruning with bending potential ratio. Pattern Recognition, 2011, 44(2):196–209

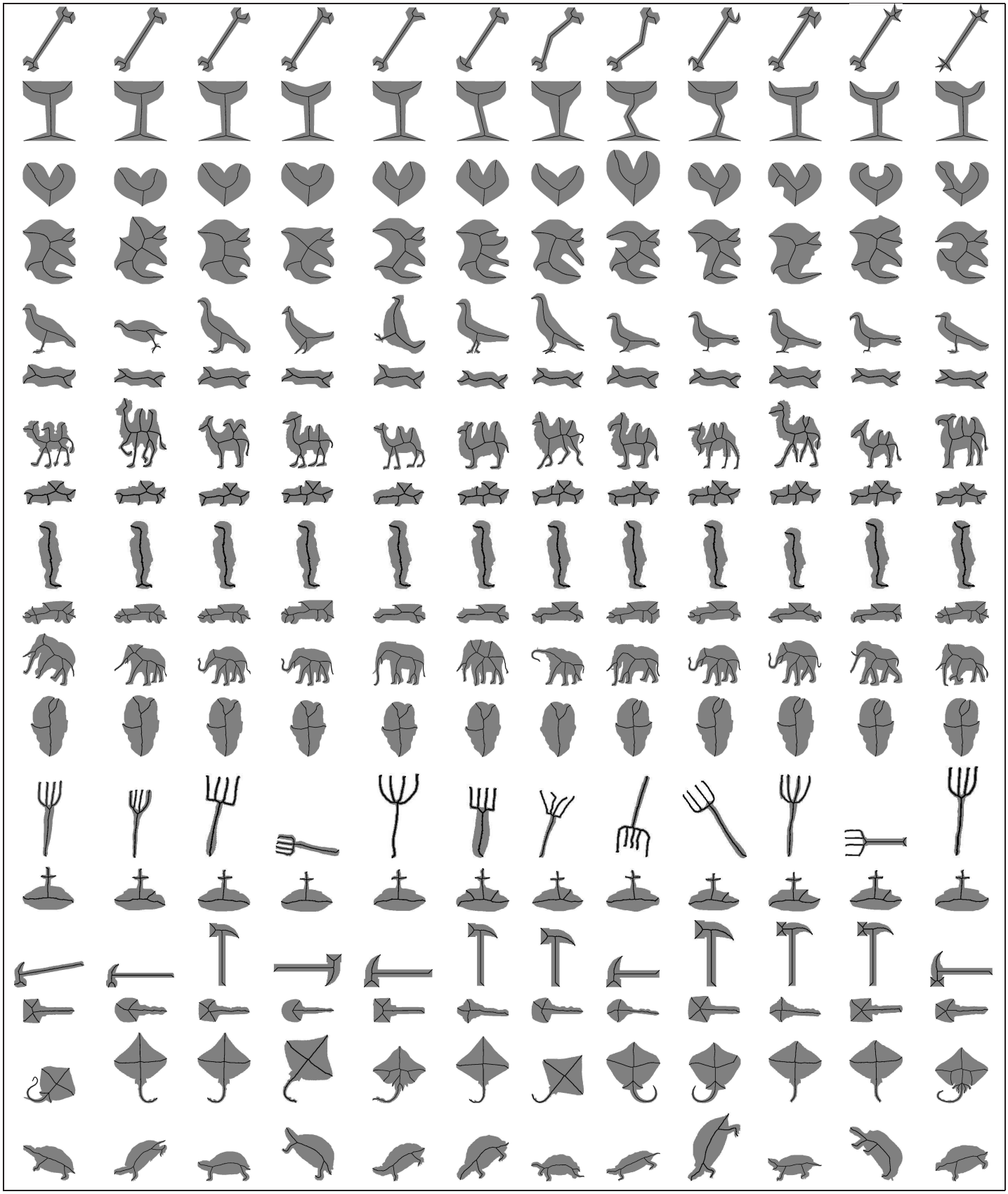

Figure 11: Our skeletons for all shapes in Kimia's dataset [40].