Planar shape decomposition made simple

Nikos Papanelopoulos
papanelo@image.ntua.gr
National Technical University of Athens
Athens, Greece

Yannis Avrithis
http://image.ntua.gr/iva/iavr/
University of Athens
Athens, Greece

Abstract

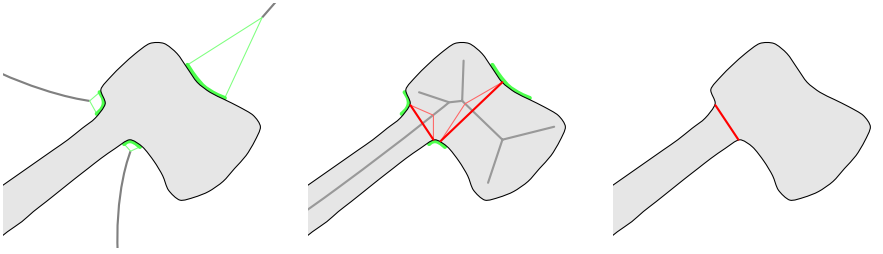
We present a very simple computational model for planar shape decomposition that naturally captures most of the rules and salience measures suggested by psychophysical studies, including the minima and short-cut rules, convexity, and symmetry. It is based on a medial axis representation in ways that have not been explored before and sheds more light into the connection between existing rules like minima and convexity. In particular, vertices of the exterior medial axis directly provide the position and extent of negative minima of curvature, while a traversal of the interior medial axis directly provides a small set of candidate endpoints for part-cuts. The final selection follows a simple local convexity rule that can incorporate arbitrary salience measures. Neither global optimization nor differentiation is involved. We provide qualitative and quantitative evaluation and comparisons on ground-truth data from psychophysical experiments.

1 Introduction

THE psychophysical, ecological, and computational aspects of planar shape decomposition into parts have been studied for more than five decades [25]. Although a complete theory of object recognition remains an impossibility, it is believed that our ability to recognize objects by their silhouette alone is related to simple rules by which the visual system decomposes shapes into parts [8]. In computer vision, object detection and recognition has deviated from such studies, but understanding visual perception towards learning better representations is always relevant [29].

Recent work on the subject has introduced ever more complex computational models relying on combinatorial optimization [16, 18, 23]. The main focus of such models is *convexity*, although the support from psychophysical studies is limited or absent [11, 24]. The most recognized rules underpinning shape decomposition are the *minima rule* [8] and the *short-cut rule* [27], along with the definition of *part-cuts* [26]. However, attempts to reflect these rules into simple computational models still resort to optimization and new ad-hoc rules [17]. Although the medial axis has been one of the first representations used even before the formulation of these rules [1, 3], it is not frequently used today. On the other hand, quantitative evaluation has been practically impossible until recently [5, 12].

Contribution. In this work, we revisit the problem assuming the medial axis representation and introduce a new computational model referred to as *medial axis decomposition* (MAD). Contrary to common belief [17], we argue that this representation is both efficient and robust, at least as far as decomposition is concerned, and as long as a part hierarchy [25] is not



(a) exterior (b) interior (c) cuts

Figure 1: Main elements of our method. (a) Exterior medial axis and concave corners (in green) as boundary arcs that are each the projection of one medial axis end vertex (minima rule). (b) Interior medial axis and candidate cuts (in red) whose endpoints are contained in corners and are projection points of the same medial axis point; only one such cut is selected per corner and medial axis branch. (c) Final cuts according to short-cut and convexity rules: the shortest cuts are selected for each corner such that each shape part is locally convex at the corner, roughly forming an interior angle less than π (up to tolerance).

sought. We show that it is possible to incorporate all rules suggested by psychophysical studies into a computational model that is so simple that one nearly "reads off" part-cuts from the medial axis. In doing so, we suggest a stronger definition of part-cuts concerning local symmetry such that the list of candidate cuts is linear in the number of minima. We also shed more light into the relation of minima to convexity by relaxing the latter to *local convexity*. Contrary to global optimization models, this guarantees robustness [25].

The main ideas of our work are illustrated in Fig. 1. As in most related work, a shape is decomposed into parts by defining a number of part-cuts which are line segments contained in the shape. According to the minima rule [8], the part-cut endpoints are points of *negative minima of curvature* of the shape boundary curve. But it is known [4] that such points are exactly projection points (boundary points of minimal distance) of end vertices of the exterior medial axis (the medial axis of the complement of the shape). Moreover, as shown in Fig. 1a, one may get from a medial axis vertex not just one boundary point but an entire arc. We call this arc a *concave corner* or simply *corner*. It is readily available and involves no differentiation, contrary to all previous work. We show there are advantages over the common single-point approach.

There is no constraint as to which pairs of minima (corner points) are candidate as part-cut endpoints, hence all prior work examines all possible pairs. On the contrary, as shown in Fig. 1b, we only consider pairs of points that are projection points of the same point of the interior medial axis (of the shape itself). Similarly to *semi-ligatures* [1] and single-minimum cuts [17], a cut may also have only one corner point as endpoint [27]. In either case, endpoint pairs are readily available by a single traversal of the medial axis. Comparing to the conventional definition, which requires part-cuts to cross an axis of local symmetry [26], this is a stronger definition in agreement with the definition of *necks* [25]. Contrary to common belief, we show that it can actually be in accordance to psychophysical evidence [5]. For each corner, we only select one cut per medial axis branch; this is a simple and intuitive rule that has not been observed before.

Now, given a candidate list of cuts, the *short-cut* rule [27] suggests that priority be given to the shortest over all cuts incident to each corner point; but it does not specify how many should be kept. On the other hand, convexity-based approaches attempt to find a minimal

number of cuts such that each shape part is convex [23]. Clearly, a concave smooth boundary curve segment would require an infinite partition, so convexity is only sought approximately. But negative minima of curvature are points where the shape is locally maximally concave. They are therefore the first points where one should establish convexity by cutting. Hence we introduce a *local convexity* rule whereby the minimal number of cuts is selected such that the interior angle of each part is less than π (up to tolerance) at each corner. Selection is linear in the number of candidate cuts and again, all information is merely read-off from the (exterior) medial axis. The final cuts are shown in Fig. 1c.

Structure. The remaining text is organized as follows. Our shape representation is given in section 2, followed by a more detailed account of our decomposition method in section 3. Experimental findings are presented in section 4 and conclusions are drawn in section 5.

2 Shape representation

A *planar shape* is a set $X \subset \mathbb{R}^2$ whose boundary ∂X is a finite union of mutually disjoint simple closed curves, such that for each curve there is a parametrization $\alpha : [0,1] \to \partial X$ by arc length that is piecewise real analytic. The (Euclidean) *distance map* $\mathcal{D}(X) : X \to \mathbb{R}$ is a function mapping each point $y \in X$ to

$$\mathcal{D}(X)(y) = \inf_{x \in \partial X} \|y - x\|, \tag{1}$$

where $\|\cdot\|$ denotes the ℓ^2 norm. For $y \in \mathbb{R}^2$, let

$$\pi(y) = \{x \in \partial X : \|y - x\| = \mathcal{D}(X)(x)\} \tag{2}$$

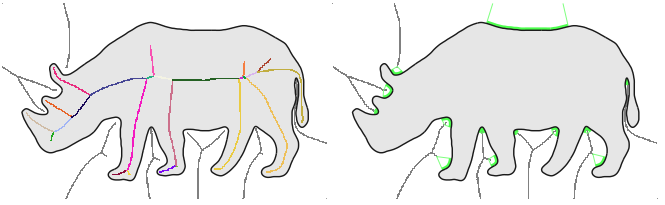
be the set of points on the boundary at minimal distance to y. This set is non-empty because ∂X is closed in $\mathbb{R}^2$ hence compact. It is called the *projection* [1] or *contact set* [4] of y on the boundary; each $x \in \pi(y)$ is called a projection or contact point of y.

The (interior) *medial axis*

$$\mathcal{M}(X) = \{x \in \mathbb{R}^2 : |\pi(x)| > 1\} \tag{3}$$

is the set of points with more than one projection points. This set is a finite linear graph embedded in $\mathbb{R}^2$ [4]. Each edge of $\mathcal{M}(X)$ is homeomorphic to the unit closed interval, and each point x in an edge has exactly two projection points; a vertex is called an *end vertex* (*junction*) if it has degree 1 (3 or higher). Assuming X is bounded, an end vertex is either a *convex vertex* of X (point of discontinuity of α' on ∂X with interior angle less than π) or the center of an osculating circle inscribed in X with a connected projection that is either one point or a circular arc; hence the curvature of α is positive and locally maximum at the projection [4]. In this work, we also use the *exterior* medial axis of X, which is the medial axis of its complement $\mathbb{R}^2 \setminus X$. In this case an end vertex is either a *concave vertex* of X (point of discontinuity of α' on ∂X with interior angle greater than π) or the curvature is negative and locally minimum at the projection.

In practice, we compute the distance map with any algorithm that provides at least one representative of the projection $\pi(y)$ of each point [7], and then compute the medial axis using the *chord residue* [7, 20]. Given two points $x, y \in \partial X$, the arc length $\ell(x, y)$ is the length of the minimal arc of ∂X having x, y as endpoints or ∞ if no such arc exists. Now, given a point z, its chord residue $r(z) = \sup_{x,y \in \pi(z)} \ell(x, y) - \|x - y\|$ is the maximal



(a) medial axis (b) concave corners

Figure 2: (a) Exterior and interior medial axes of shape `#186` from S&V dataset [5]. Branches of the latter shown in random color. (b) *Minima*: exterior medial and concave corners (in green, along with lines connecting vertices to their projection points).

difference between arc length and chord length over all pairs of points in its projection. The residue is non-negative, attains a maximum at a single *center point* of each path component of $\mathcal{M}(X)$, and is a non-increasing function of distance to the center point on $\mathcal{M}(X)$.

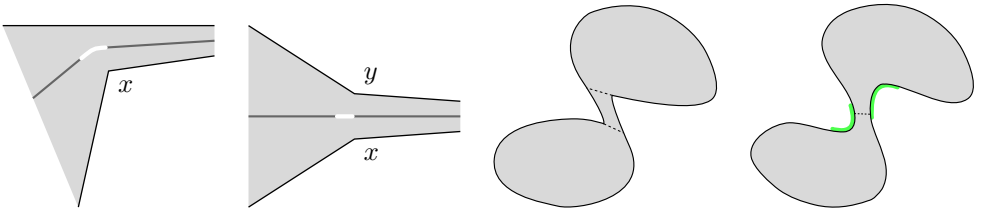
Construction of the medial axis begins at local maxima of the distance map and propagates as long as the residue, measured between single-point projections of neighboring points, is higher than a given threshold $\sigma > 0$. Contrary to common misinterpretation [22], this method is not constrained to polygons. It is very efficient, does not involve differentiation *e.g.* of the distance map, preserves shape topology under mild assumptions (in particular, yields one connected component of the medial axis for each component of X), and can simplify (in a sense, prune) the medial axis by merely adjusting σ, without simplifying the curve ∂X in any way. Typically, σ is only 1-2 pixels just to remove discretization noise. Unfortunately, it is constrained to two dimensions.

In the following, we assume that both the interior and exterior medial axes are available—in fact, both are computed on a single traversal over a discrete representation of the input shape X on a regular grid. For simplicity, we assume that for each point x of the medial axis, the projection $\pi(x)$ contains exactly two points. In practice, only one projection point is stored for each x; the second one is obtained from x's neighbors. The arc length is computed in constant time [20]. For the interior medial axis, we also parse its graph structure by a single traversal; we refer to the edges of the graph as medial axis *branches*. Fig. 2a illustrates the two medial axes and the branches found on the interior one for a sample shape that will also serve as a running example in section 3 below.

3 Shape decomposition

A shape X is decomposed into parts by defining a set of *part-cuts* or simply *cuts*, as common part boundaries. The cut endpoints, in turn, serve as boundaries between parts of ∂X. In some cases, cuts have been defined as curves, *e.g.* cubic splines, providing for *continuation* of boundary tangents at end points [25]; but in most relevant work, as well as in the current work, cuts are just line segments for simplicity [9, 26]. In either case, the cut endpoints always lie on the boundary ∂X and the cuts lie entirely on the closure of X [26]. Additional conditions apply as discussed below.

In this work, a large number of *raw cuts* is initially extracted by traversing the interior medial axis; a short list of *candidate cuts* is selected by means of an equivalence relation, and a final cut selection follows by seeking local convexity at each endpoint along with a few simple salience measures. The entire decomposition process is detailed below.



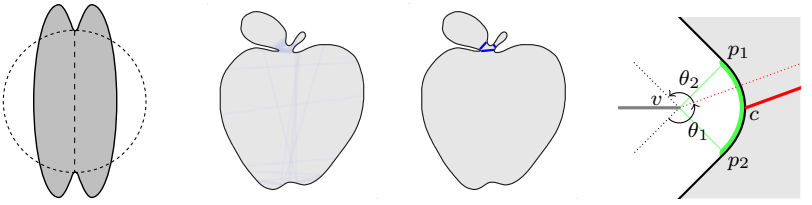
(a) full ligature (b) semi-ligature (c) sharp (d) weak

Figure 3: (a) Full ligature on x. (b) Semi-ligature on x, y (in white) [1]. (c) Two nearby sharp concavities result in two different cuts [26]. (d) Two nearby weak concavities should ideally result in one cut; this is possible if their *locale* [9] is known (in green).

Minima. According to the *minima rule* [8], the shape X should be cut at points of negative minima of curvature of its boundary parametrization α. In the theory of limbs and necks [25], this rule is taken to mean that *both* cut endpoints are such minima points. However, the rule has been subsequently relaxed by requiring that *at least one* of each cut endpoints have negative curvature [27]. This condition is contained in the standard definition of part-cuts [26]. This is in agreement with the earlier theory of ligatures [1] and more recent studies [17]. In particular, given a set of minima points C, a *full(semi)-ligature* [1] on two points $x, y \in C$ (resp. one point $x \in C$) is the set of points z whose projection $\pi(z)$ contains x, y (resp. x but no other point of C). Commonly referred to as ligatures, these sets are subsets of the medial axis and disconnect it such that subsequent shape reconstruction produces a rough decomposition into parts. They are illustrated in Fig. 3a,b. Accordingly, *double(single)-minima cuts* [17] are defined as having both endpoints (resp. exactly one endpoint) in the minima set C. We follow the same idea.

But how is the minima set C exactly determined? All relevant studies assume a discrete parametrization of shape boundary ∂X and compute negative minima of a discrete approximation of curvature. Apart from numerical sensitivity and the further assumption of a scale parameter in every discrete derivative approximation, the limitation is that detected minima are isolated points that provide no information on the spatial extent of concavities—referred to as *locale* [9]—as illustrated in Fig. 3c,d. The background of section 2 specifies that end-vertex projections of the *exterior* medial axis are either single points tangent to osculating circles, or circular arcs. In practice, the two projection points determine a boundary arc that always approximates a circular arc. We call this arc a *concave corner* or simply *corner*. The radius of the circle is the inverse of the absolute curvature. The three points involved—the end vertex and its two projection points—directly determine the position, spatial extent, orientation and strength of the concavity, including both curvature and turning angle. All information comes for free from the medial axis. Fig. 2b illustrates this idea.

Symmetry. Now potential cuts are determined by all pairs of points in two different corners. Most relevant work actually examines all pairs [17, 25]. This is not only inefficient, but may involve all sorts of new ad-hoc rules to resolve conflicts (*e.g.* that cuts do not intersect) as well as solving an optimization problem. But the standard definition of part-cuts [26] includes the additional condition that they cross an axis of local symmetry. We modify the condition such that the cut endpoints are projection points of the same point of the interior medial axis (recall that a cut lies in the shape). In most cases this is a stronger condition, but we observe that it most often agrees with ground truth data from psychophysical experiments [5], as shown in Fig. 4a-c. Combined with the minima rule, it implies that endpoints are exactly projection



(a) local symmetry (b) all subjects (c) majority (d) local convexity

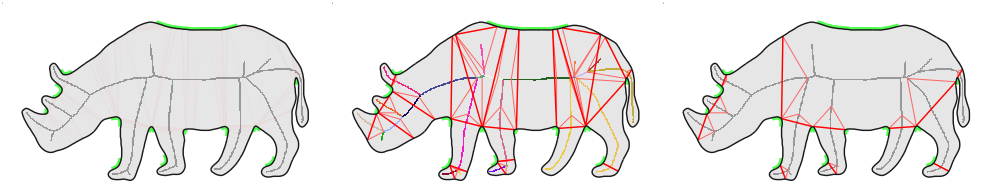
Figure 4: (a) Example from Singh and Hoffman [26] illustrating that a cut across a local symmetry axis fails to be captured by the medial axis or equivalently by the definition of neck [25] because a circle cannot be inscribed. (b) A counter-example from ground-truth data of DeWinter and Wagemans [5] where most subjects do not cut in a similar case (cuts of all subjects overlaid in blue, 85% transparent). (c) *Majority* cuts of (b) (in blue) according to clustering-based ensemble [12]; see section 4 for more details. (d) Measurement of interior angles at a corner. Exterior medial vertex v and its projection points p_1, p_2 are known, hence also the inward/outward orientations of the boundary—the lines joining v and p_1, p_2 are normal to the boundary. By translating the two boundary segments starting at p_1, p_2 and the cut starting at c to the same origin (vertex v here), we measure the interior angles θ_1, θ_2 of the two shape parts at this corner after cutting. Both are less than π, while $\theta_1 + \theta_2$ is not. *Local convexity* is achieved and there is no need for more cuts at this corner.

points of the same point of a ligature. So what we do in practice is, traverse the interior medial axis once, and collect all pairs of projection points such that at least one lies in a corner. Depending on the number of corners, we call the cuts *double* or *single*. The cuts obtained this way are called *raw cuts* and illustrated in Fig. 5a. It is easily shown that they do not intersect by construction.

Equivalence. Observing Fig. 5a, raw cuts are clearly too many, but they tend to appear in groups. As shown Fig. 5b, we select a small number of *candidate cuts* before applying other rules by defining two equivalence relations on cuts and selecting one representative from each equivalence class. According to the first relation, *corner equivalence*, two (double) cuts are equivalent if their endpoints lie on the same pair of corners. In this case, the representative is chosen that maximizes the protrusion strength measure discussed below. The second, *branch equivalence*, specifies that two cuts (double or single) are equivalent if they are on the same branch and their endpoints share at least one corner; we say a cut is on a branch if the medial axis point whose projection points are the cut endpoints lies on this branch.

This rule is intuitive and always maintains all correct cuts in our experiments. Observe in Fig. 5b that whenever two groups of cuts are on the same corner but on two different branches, there is also a junction and a third branch in the outward direction from the corner, such that the shape is expanding between the two groups. Hence there should be a representative from both cut groups. If there are both double-cuts and single-cuts in the same equivalence class, the representative is always a double-cut. In either case, the representative cut is chosen such that its endpoints are closest to the midpoint of the corner arc(s). This rule is almost always superior to the short-cut rule. For instance, observe the cuts at the tail in Fig. 5a,b and compare to Fig. 7b of De Winter and Wagemans [5].

Local convexity. Although the psychophysical evidence concerning convexity as a rule for shape decomposition is limited, most recent studies are based on optimization targeting approximate convexity. We rather avoid global optimization, not only for its complexity



(a) raw cuts (b) candidate cuts (c) selected cuts

Figure 5: (a) *Symmetry*: all cuts (in red, 95% transparent) for the shape of Fig. 2, whose endpoints are projection points of the same interior medial axis point, with at least one endpoint on a concave corner. (b) *Equivalence*: candidate cuts selected such that for each corner, there is a single cut per medial axis branch (each shown in random color); priority is given to cuts with both endpoints on corners, while representatives are selected such that endpoints are as close as possible to the midpoint of the associated corner arc. (c) Cuts selected independently at each corner by descending priority until *local convexity* is achieved; priority specified by short-cut rule and other salience measures.

but also because according to robustness requirement [25], decomposition at a point should only be affected by its local neighborhood, such that partial occlusion and part movement do not affect the remaining parts. We observe that the minima rule is inherently related to convexity, since boundary points of negative minima of curvature are in fact points where the shape is locally maximally concave. We therefore select cuts independently at each corner in order to achieve *local convexity* at the corner. In particular, for every corner, we prioritize all cuts incident to the corner according to criteria discussed below, and we select cuts by descending priority until the interior angle of all parts after cutting is less than $\pi + \theta$, where θ is a tolerance. Measurement of interior angles is illustrated in Fig. 4d; once more, all information is readily available from the medial axis. Selected cuts according to local convexity and short-cut rule (see below) are illustrated in Fig. 5c.

Salience measures. Our local convexity rule selects the appropriate number of part-cuts independently per corner but is otherwise completely agnostic to their prioritization. This enables the use of arbitrary salience measures for cuts. Although there is no complete theory, several such measures have been suggested as plausible, going back to at least Gestalt psychologists [9, 26]. These refer to boundary strength at cut endpoints [9], including turning angle for cusps and normalized curvature for smooth boundary, continuation of boundary at endpoints [26], as well as of salience of cuts or parts themselves, including relative area, protrusion [9], and cut length [27]. Studying their role is beyond the scope of this work. We rather focus on a general framework that can easily incorporate any of these measures; all are readily available in our representation with the exception of relative area. In practice, we follow a minimal approach by only using cut length (shorter cuts are preferred) after first discarding cuts with protrusion less than a given (inverse) threshold p. The latter is defined as the ratio of cut length to the perimeter (arc length) of the part excluding the cut.

4 Experiments

4.1 Experimental setup

Datasets. In most related work [25], even in recent methods [16, 19, 30], evaluation is

only qualitative, while quantitative evaluation is often limited to datasets that are not public like arbitrary subsets of MPEG-7 shape dataset [18, 23]. To our knowledge, there are two public datasets with ground-truth from human subjects [5, 15]. The former by Liu *et al*. is focusing on the classification of holes as structurally important or topological noise, which is a different problem. We use the latter by de Winter and Wagemans, which evaluates exactly segmentation of object outlines. It is a subset of the Snodgrass and Vanderwart (S&V) everyday object dataset [28], consisting of 260 line drawings. The subset refers to 88 of the drawings, which have been converted to smooth outlines and each segmented by 39,5 subjects (psychology students) on average. For each shape there are 122,4 part-cuts, that is 3,1 cuts per subject on average. The same dataset, referred to as S&V, has been subsequently used for quantitative comparison of different computational models [12, 13, 17]. An example illustrating the cuts of all subjects on a single outline is shown in Fig. 4b.

Majority voting. Because part-cuts of human subjects are typically inconsistent, it is common practice to perform some form of majority voting before using the ground-truth to evaluate a computational model [25]. There are different alternatives, which take the form of either a majority decomposition by clustering [12, 15], or spatial density used directly for evaluation [17]. We follow the framework of Lewin *et al*. [12]. In particular, given two cuts c_1, c_2 with endpoints $\{x_1, y_1\}$, $\{x_2, y_2\}$ respectively, their *arc distance* is defined as

$$d(c_1, c_2) = \min\{\ell(x_1, x_2) + \ell(y_1, y_2), \ell(x_1, y_2) + \ell(y_1, x_2)\}, \tag{4}$$

where ℓ is the arc length function defined in section 2. Using this distance, cuts are subject to average-linkage agglomerative clustering and a cluster is only kept if contains cuts from a given proportion of the subjects. A representative cut is chosen from each cluster whose endpoints are averaged over the endpoints of individual cuts in the cluster, where averaging takes place on the parametrization of the boundary curve. The result is a *majority decomposition* per shape. An example is given in Fig. 4b,c.

Evaluation measures. Unfortunately, since quantitative evaluation is relatively new, there is nearly one different protocol for every relevant publication. We use two different measures, both of which assume a decomposition of shape X is represented by a partition $A = \{A_i\}$ of X, where both X and each part A_i are represented by sets of pixels in practice. The *Hamming distance* [12] of partitions A, B is then

$$H(A, B) = \frac{1}{2|X|}[h(A|B) + h(B|A)], \tag{5}$$

where $|X|$ is the area of X in pixels, $h(A|B) = \sum_i |A_i \setminus B_{\pi_i}|$ is the sum over all parts of A of the area of part A_i not covered by its best match B_{π_i} in B, and the best match is defined by $\pi_i = \arg\max_j |A_i \cap B_j|$. On the other hand, the *Rand Index* (RI) [15] of A, B, is

$$R(A, B) = \binom{n}{2}^{-1} (|P_{A\bar{B}}| + |P_{B\bar{A}}|), \tag{6}$$

where $P_{A\bar{B}} = \{(x_i, x_j) \in X^2 : j > i \wedge A(x_i) = A(x_j) \wedge B(x_i) \neq B(x_j)\}$ are the ordered pairs of pixels in X that are in the same part of A and in different parts of B, and $A(x_i)$ is the part of A where pixel x_i belongs. It is also referred to as Jaccard measure [12]. We evaluate against both individual human subjects by averaging per shape and to majority decompositions.

Compared methods. We perform quantitative comparison to approximate convex decomposition (ACD) [14], discrete contour evolution (DCE) [11], combined skeleton-boundary features (SB) [30], flow discretization (FD) [8], constrained morphological decomposition (MD) [10] and clustering-based ensemble (CBE) [13]. Quantitative results on all methods are used as provided by Lewin *et al*. [13], which propose their own ensemble method CBE. The latter is in fact applying to all previous five methods the same clustering approach that is also applied to human subject decompositions as part of majority voting; therefore it may be considered as a meta-decomposition method. Our own method is referred to as *medial axis decomposition* (MAD). We also compare to human subjects, evaluated individually against their own average or majority [15], as well as to the baseline case of not cutting anywhere. Qualitative results, apart from ground truth (GT), are additionally compared to relatability (REL) [19], convex shape secomposition (CSD) [16], minimum near-convex decomposition (MNCD) [23] and computational model of short-cut rule (CSR) [17].

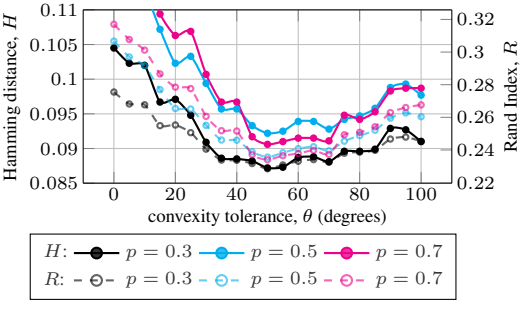


Figure 6: Hamming (H) and RI (R) measures *vs*. convexity tolerance θ for varying protrusion threshold p on S&V dataset.

	average		majority	
	H	R	H	R
DCE	0.208	0.497	0.188	0.466
SB	0.163	0.402	0.131	0.335
MD	0.151	0.371	0.126	0.328
FD	0.145	0.350	0.112	0.267
ACD	0.128	0.323	**0.092**	0.251
MAD	**0.126**	**0.317**	0.096	**0.247**
MAD-opt	0.118	0.303	0.085	0.225
CBE	**0.111**	**0.288**	**0.069**	**0.186**
Human	0.128	0.312	0.093	0.245
Baseline	0.160	0.424	0.140	0.376

Table 1: Hamming (H) and RI (R) measures for average and majority voting on S&V dataset.

4.2 Results

Timing. Implemented in C++ and Matlab, MAD takes 78ms on average per S&V shape on a single core, excluding medial axis preprocessing, which is 202ms on average.

Tuning. There are three parameters in MAD: medial scale threshold σ, convexity tolerance θ, and (the inverse of) protrusion strength threshold p. After both quantitative tuning and qualitative inspection, we choose $\sigma = 2$. Fig. 6 shows quantitative results for different configurations of θ, p. Performance is best for $p = 0.3$ (lower is better for both measures H and R), while the optimal range for θ is $[45, 70]$ degrees. However, due to qualitative inspection, we rather choose $p = 0.45, \theta = 40°$. Discrepancies are attributed to limitations of the measures used [15].

Quantitative evaluation. Table 1 compares our method to a number of relevant methods. On most measurements, our method is outperforming all individual methods and very close to or even better than human subjects. As a meta-method, CBE is the best, but is also significantly more complex and involves all the other five methods shown in the Table. It is expected to perform well since it applies to algorithms the same idea of majority voting that is applied to human subjects at ground truth construction. Because parameters were chosen mostly based on qualitative criteria, we also include for reference the optimal performance obtained with $p = 0.35, \theta = 50°$ (MAD-opt), which is always better than human subjects

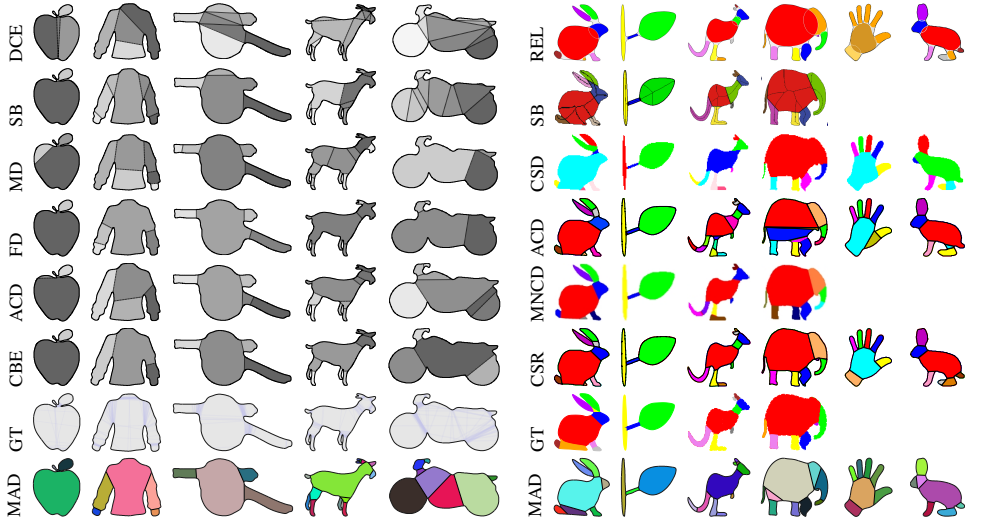


Figure 7: Qualitative results on representative shapes of S&V [5] (left) and Kimia [25] (right) datasets for a number of methods, including ground truth (GT), depicted as in Fig. 4c.

and not too far from CBE. All results are inferior on average evaluation against individual subjects, which is expected as subjects are not always consistent. It is interesting that SB, DCE are close to or even worse than the baseline of not cutting anywhere.

Qualitative evaluation. Fig. 7 illustrates qualitative results on a number of representative shapes. Our method gives natural results on Kimia dataset and is the only one to capture the ground truth for the bottom part of the rabbit correctly. It often tends to prefer cuts near the mouth than on the neck. This is attributed to the shortcut rule which we observe is not always enough, but our method is very open to using other measures. S&V is harder, but still MAD yields the highest quality results comparing to the other individual methods. That is, apart from the ensemble method CBE, which seeks consensus among all others.

5 Discussion

Both qualitative and quantitative evaluation suggests that an extremely simple computational model based on an appropriate representation can be competitive comparing to more complex models or ensemble methods. More than that, our model is inherently connected to most rules suggested by human vision studies and highlights their connection. There are more aspects that we have explored in the same model, which we have not been able to expose here due to limited space, including an extended definition of concave corners that captures semi-local boundary arcs and a proximity measure on part-cuts. Other aspects that could be naturally incorporated are detection of bends, continuation of boundaries across parts and local symmetry beyond what is captured by the medial axis. The fact that part-cut selection is based on simple local decisions will enable the investigation of a more general model beyond closed curves towards local feature detection on arbitrary natural images. For instance, bitangents on isophotes (level sets of intensity) [21] can be seen as cuts on either figure or ground shape, while distance map saddle points [2] correspond to necks [25]; our work can provide for a richer set of cuts hence candidate local features.

References

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

[2] Yannis Avrithis and Konstantinos Rapantzikos. The medial feature detector: Stable regions from image boundaries. In *ICCV*, 2011.

[3] Harry Blum and Roger N Nagel. Shape description using weighted symmetric axis features. *Pattern Recognition*, 10(3):167–180, 1978.

[4] H.I. Choi, S.W. Choi, and H.P Moon. Mathematical theory of medial axis transform. *Pacific Journal of Mathematics*, 181(1):57–88, 1997.

[5] Joeri De Winter and Johan Wagemans. Segmentation of object outlines into parts: A large-scale integrative study. *Cognition*, 99(3):275–325, 2006.

[6] Tamal K Dey, Joachim Giesen, and Samrat Goswami. Shape segmentation and matching with flow discretization. In *Algorithms and Data Structures*, pages 25–36. Springer, 2003.

[7] P.F. Felzenszwalb and D.P Huttenlocher. Distance transforms of sampled functions. Technical report, 2004.

[8] Donald D Hoffman and Whitman A Richards. Parts of recognition. *Cognition*, 18(1): 65–96, 1984.

[9] Donald D Hoffman and Manish Singh. Salience of visual parts. *Cognition*, 63(1): 29–78, 1997.

[10] Duck Hoon Kim, Il Dong Yun, and Sang Uk Lee. A new shape decomposition scheme for graph-based representation. *Pattern Recognition*, 38(5):673–689, 2005.

[11] Longin Jan Latecki and Rolf Lakamper. Convexity rule for shape decomposition based on discrete contour evolution. *Computer Vision and Image Understanding*, 73(3):441–454, 1999.

[12] Sergej Lewin, Xiaoyi Jiang, and Achim Clausing. Framework for quantitative performance evaluation of shape decomposition algorithms. In *ICPR*, 2012.

[13] Sergej Lewin, Xiaoyi Jiang, and Achim Clausing. A clustering-based ensemble technique for shape decomposition. In *Structural, Syntactic, and Statistical Pattern Recognition*, pages 153–161. Springer, 2012.

[14] Jyh-Ming Lien and Nancy M Amato. Approximate convex decomposition of polygons. In *Proceedings of Symposium on Computational Geometry*, pages 17–26. ACM, 2004.

[15] Guilin Liu, Zhonghua Xi, and Jyh-Ming Lien. Dual-space decomposition of 2d complex shapes. In *CVPR*, 2014.

[16] Hairong Liu, Wenyu Liu, and Longin Jan Latecki. Convex shape decomposition. In *CVPR*, 2010.

[17] Lei Luo, Chunhua Shen, Xinwang Liu, and Chunyuan Zhang. A computational model of the short-cut rule for 2d shape decomposition. *IEEE Transactions on Image Processing*, 24(1):273–283, 2015.

[18] Chang Ma, Zhongqian Dong, Tingting Jiang, Yizhou Wang, and Wen Gao. A method of perceptual-based shape decomposition. In *ICCV*, 2013.

[19] Xiaofeng Mi and Douglas Decarlo. Separating parts from 2d shapes using relatability. In *ICCV*, 2007.

[20] R. Ogniewicz and M. Ilg. Voronoi skeletons: Theory and applications. In *Computer Vision and Pattern Recognition*, 1992.

[21] M. Perdoch, J. Matas, and S. Obdrzalek. Stable affine frames on isophotes. In *ICCV*, volume 104, 2007.

[22] S.M. Pizer, K. Siddiqi, G. Szekely, J.N. Damon, and S.W Zucker. Multiscale medial loci and their properties. *International Journal of Computer Vision*, 55(2):155–179, 2003.

[23] Zhou Ren, Junsong Yuan, Li Chunyuan, and Liu Wenyu. Minimum near-convex decomposition for robust shape representation. In *ICCV*, 2011.

[24] Paul L Rosin. Shape partitioning by convexity. *IEEE Transactions on Systems, Man, and Cybernetics, Part A*, 30(2):202–210, 2000.

[25] Kaleem Siddiqi and Benjamin B Kimia. Parts of visual form: Computational aspects. *PAMI*, 17(3):239–251, 1995.

[26] Manish Singh and Donald D Hoffman. Part-based representations of visual shape and implications for visual cognition. *Advances in Psychology*, 130:401–459, 2001.

[27] Manish Singh, Gregory D Seyranian, and Donald D Hoffman. Parsing silhouettes: the short-cut rule. *Perception and Psychophysics*, 1999.

[28] Joan G Snodgrass and Mary Vanderwart. A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6(2):174, 1980.

[29] Christian Szegedy, Scott Reed, Dumitru Erhan, and Dragomir Anguelov. Scalable, high-quality object detection. Technical report, 2014.

[30] Jingting Zeng, Rolf Lakaemper, Xingwei Yang, and Xin Li. 2d shape decomposition based on combined skeleton-boundary features. In *Advances in Visual Computing*, pages 682–691. Springer, 2008.