# Revisiting the medial axis for planar shape decomposition

Nikos Papanelopoulos, Yannis Avrithis, Stefanos Kollias

Computer Vision and Image Understanding
journal homepage: www.elsevier.com

# Revisiting the Medial Axis for Planar Shape Decomposition

Nikos Papanelopoulos[a,**], Yannis Avrithis[b], Stefanos Kollias[c]

[a]National technical university of Athens, Zografou Campus 9 Iroon Polytechniou str, Zografou, 15780, Greece
[b]Inria Rennes–Bretagne Atlantique, Campus Universitaire de Beaulieu, Rennes, 35042, France
[c]School of Computer Science University of Lincoln, Brayford Pool, Lincoln LN6 7TS, UK

ABSTRACT

We present a 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 prioritized processing of candidate part-cuts according to a 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. With our single computational model, we outperform even an ensemble method on several other competing models.

## 1. Introduction

THE psychophysical and computational aspects of planar shape decomposition into parts have been studied for more than five decades Siddiqi and Kimia (1995). 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 Hoffman and Richards (1984).

In computer vision, object detection and recognition has deviated from such studies with the advent of deep learning: modern approaches learn to detect or segment objects from raw data without necessarily studying their silhouette or its part decomposition Ren et al. (2017); Liu et al. (2016); Noh et al. (2015), and even shape recognition in 2D Yu et al. (2017) or 3D Garcia-Garcia et al. (2016) does not necessarily consider object parts. To our knowledge, semantic part segmentation from 2D images Tsogkas et al. (2015) or 3D shapes (point clouds) Garcia-Garcia et al. (2016); Yi et al. (2017) based on deep learning is so far fully supervised by semantic part annotation and despite excellent performance, little is known on how to interpret the predictions of such models. Understanding visual perception towards learning better representations is always relevant, so the current study focuses on unsupervised 2D shape decomposition using simple interpretable rules.

### 1.1. Related work

According to psychophysical findings, the most recognized rules underpinning shape decomposition are the *minima rule* Hoffman and Richards (1984) and the *short-cut rule* Singh

**Corresponding author:
*e-mail:* papanelo@image.ntua.gr (Nikos Papanelopoulos)

et al. (1999), along with the definition of *part-cuts* Singh and Hoffman (2001). However, attempts to reflect these rules into simple computational models often resort to optimization and new rules Luo et al. (2015). Although the medial axis has been one of the first representations used even before the formulation of these rules Blum and Nagel (1978); August et al. (1999), it is not frequently used today.

Another popular rule is *convexity*, although the support from psychophysical studies is limited or absent Latecki and Lakamper (1999); Rosin (2000). In this work we observe that there is a direct connection between convexity and the minima rule: points of negative minima of curvature are detected early in our analysis, while a convexity measure is used at a later stage to prioritize part cuts. Models based on convexity are often based on iterative removal of the most non-convex features from the shape *boundary* Latecki and Lakamper (1999); Lien and Amato (2004). While this is intuitive and similar to our model, here we rather use the medial axis representation where it is easier to incorporate a rich set of additional saliency measures and rule.

Recent work on the subject has introduced complex computational models relying on *combinatorial optimization* Luo et al. (2015), and in many cases the objective or the constraints are still based on convexity Liu et al. (2010); Ren et al. (2011); Ma et al. (2013). While this may work better than greedy approaches Latecki and Lakamper (1999); Lien and Amato (2004), they are still based on a boundary representation, where pair-wise terms arise for all pairs of boundary points, unnecessarily increasing the cost. More importantly, global optimization over the entire shape is contradicting the robustness requirement Siddiqi and Kimia (1995), whereby decomposition at a point should only be affected by its local neighborhood.

Features related to the *medial axis* are present in a lot of models, though not always explicitly connected conceptually or computationally. Skeleton features can be combined with boundary features Zeng et al. (2008), where the boundary is most notably used in applying the minima rule. To our knowledge, we are the first to detect points of negative minima of curvature directly from the *exterior* medial axis, that is, the medial axis of the shape complement. *Smooth local symmetries* is an alternative representation that has been used for shape decomposition Mi and Decarlo (2007), which however is also not straightforward to incorporate features other than symmetry.

The medial axis is well known for its sensitivity to small changes in the boundary Marr (1982). This can be overcome *e.g*. by simplifying the boundary Bai et al. (2007) or simply thresholding the *chord residue* Ogniewicz and Ilg (1992). More importantly, such changes can be identified by ligatures Blum and Nagel (1978); August et al. (1999), which essentially give rise to the minima rule. Apart from its sensitivity, the medial axis is some times explicitly avoided due to its cost Luo et al. (2015). Here 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 Siddiqi and Kimia (1995) is not sought. For instance, introducing a small protrusion in a shape would result in an entire new branch of the medial axis but in terms of decomposition, this protrusion would be simply cut off.

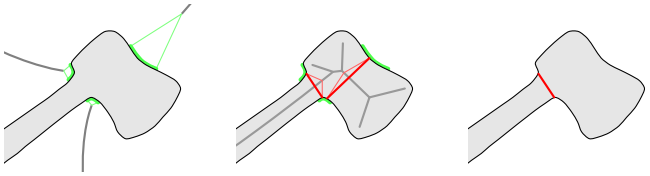
A shape is often discretized into a polygon, where the medial axis is replaced by a Voronoi diagram, and its dual Delaunay triangulation can indeed provide a limited set of candidate cuts Dey et al. (2003). This is similar to how we construct our own candidate cuts from the medial axis. While discretization is advantageous computationally, it is an unnecessary approximation if one can work efficiently, directly on the medial axis. Other works are morphological operations such as openings Kim et al. (2005). This is equivalent to using the medial axis but is clearly not the most efficient computational model. Conceptually, in terms of rules, the closest model to ours is Luo et al. (2015); computationally however it is very different, expressed as optimization over a boundary representation. We also use a richer set of rules and salience measures.

A major obstacle against progress in the field has been the lack of *annotated datasets* by human subjects, often limiting comparisons to a few qualitative examples. Psychophysical studies are typically based on experiments on multiple human subjects, but these experiments are not reproducible without sharing the data. Notable exceptions are the Kimia dataset Sid-

diqi and Kimia (1995), where human ground truth is given as a single decomposition per shape, and more recently, the annotated S&V dataset De Winter and Wagemans (2006), where cuts are defined by several human subjects per shape. The latter is crucial because in the absence of a concrete guidelines, there is typically little consensus between humans. This has been shown by Lewin et al. (2012b), where an ensemble method has been employed to find *majority* ground-truth cuts where most subjects agree.

A similar *clustering-based ensemble* Lewin et al. (2012a) has been used to aggregate the results of several existing computational models. Using a form of consensus, "majority" cuts are found where most algorithms agree. This yields the current state of the art on the S&V dataset. However, it is unsatisfying, not only computationally but also conceptually, because it is in principle a late fusion of algorithms that are used as black boxes. There is no intuition as to how rules of one algorithm might interact with rules of another, or how they may be correlated or even identical. Our main achievement in this work is to derive a single computational model based on a single shape representation that is compatible with a flexible set of interpretable rules and salience measures and outperforms individual models, the ensemble method, as well as individual humans against the majority ground truth.

In this work, we revisit the problem using the medial axis representation and introduce a new computational model, called *medial axis decomposition*. 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 constructing a 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*. This improves *robustness* Siddiqi and Kimia (1995) compared to global optimization models.



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

**Fig. 1. Main elements of our method. (a) *Minima rule*: exterior medial axis and concave corners (in green) as boundary arcs that are each determined by one medial axis end vertex. (b) *Symmetry*: 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) *Convexity rule*: cuts are prioritized and selected for each corner such that each shape part is locally convex at the corner, roughly forming an interior angle less than $\pi$ (up to tolerance).**

### *1.2. Overview*

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 Hoffman and Richards (1984), the part-cut endpoints are points of *negative minima of curvature* of the shape boundary curve. But it is known Choi et al. (1997) 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 according to our knowledge. 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* August et al. (1999) and single-minimum cuts Luo et al. (2015), a cut may also have only one corner point as endpoint Singh et al. (1999). In either case, endpoint pairs are readily available by a single traver-

sal of the medial axis. Comparing to the conventional definition, which requires part-cuts to cross an axis of local symmetry Singh and Hoffman (2001), this is a stronger definition in agreement with the definition of *necks* Siddiqi and Kimia (1995). We actually show that this can be in accordance to psychophysical evidence De Winter and Wagemans (2006) in some cases. In general, some ground-truth cuts may be lost but we introduce a way to recover them. 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 list of candidate cuts, the *short-cut* rule Singh et al. (1999) 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 Ren et al. (2011). 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 $\pi$ (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.

### *1.3. Prior work*

This work is an extension of our previous work (Papanelopoulos and Avrithis, 2015), referred to as MAD, which is also based on a medial representation and follows the ideas outlined above. In (Papanelopoulos and Avrithis, 2015), we select cuts by applying a local convexity rule independently to each corner. Additionally, cuts lying on a corner are prioritized according to the short-cut rule alone. But these choices often lead to cuts that are not consistent with human ground truth. In this work, we use four saliency measures to discard cuts before applying the local convexity rule. In particular, apart from *protrusion strength* that was used in (Papanelopoulos and Avrithis, 2015), we also use *flatness*, *expansion strength* and *extension strength* as discussed in section 6. In applying the local convexity rule, we first prioritize corners according to a measure of distance from the center of the shape as discussed in section 7.2. All cuts lying on a corner are then prioritized as discussed in section 7.1. In contrast to (Papanelopoulos and Avrithis, 2015), when we select cuts to achieve local convexity at a corner, we penalize the remaining cuts lying on this corner. The selection of cuts at a corner is thus no longer independent of the selections at other corners as in (Papanelopoulos and Avrithis, 2015).

Furthermore, ground truth cuts are commonly not found in (Papanelopoulos and Avrithis, 2015) due to a well-known limitation of the medial axis. In this work, we *recover* those missing cuts while still relying on the medial axis, as discussed in section 5.2. We also introduce a number of other extensions including the definition of *extended corners* (section 3.3), different *equivalence relations* among candidate part-cuts (section 4), as well as *protecting* certain part-cuts beyond the requirements of local convexity (section 5.1). We thereby improve the quantitative and qualitative results on a standard human-annotated shape dataset. Our extended model, referred to as MAD$^*$, is more complex than MAD; however it still relies on the medial axis representation alone and it outperforms all known 2d shape decomposition methods, including the ensemble method Lewin et al. (2012a).

### *1.4. 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. In section 4, we discuss two equivalence rules we use to reduce the number of candidate cuts. Recovered and protected cuts are discussed in section 5, while in section 6 we discuss the salience measures we apply to discard or prioritize cuts. The final selection of cuts is determined by the local convexity rule discussed in section 7. Experimental findings are presented in section 8 and conclusions are drawn in section 9.

## 2. Shape representation

### 2.1. Medial axis

A *planar shape* is a set $X \subset \mathbb{R}^2$ whose boundary $\partial X$ 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* $\mathcal{D}(X) : \mathbb{R}^2 \to \mathbb{R}$ is a function mapping each point $z \in \mathbb{R}^2$ to

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

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

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

be the set of points on the boundary at minimal distance to $z$, where we have omitted the dependence on $X$ for the sake of readability. This set is non-empty because $\partial X$ is closed in $\mathbb{R}^2$ hence compact. It is called the *projection set* August et al. (1999) or *contact set* Choi et al. (1997) of $z$ on the boundary; each $x \in \pi(z)$ is called a *projection* or *contact point* of $z$.

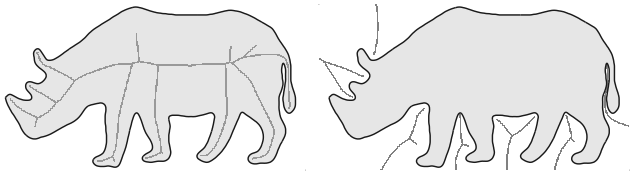
The (interior) *medial axis*

$$\mathcal{M}(X) = \{z \in X : |\pi(z)| > 1\} \tag{3}$$

is the set of points of $X$ with more than one projection points. This set is a finite linear graph embedded in $\mathbb{R}^2$ Choi et al. (1997). Each edge of $\mathcal{M}(X)$ is homeomorphic to the unit closed interval, and each point $z$ in an edge has exactly two projections; a vertex is called an *end vertex* (resp. *junction*) if it has degree 1 (resp. 3 or higher).

Given a point $z$ on an edge or a junction of $\mathcal{M}(X)$, $z$ is the center of a circle inscribed in $X$, which is tangent to $\partial X$ at the projections of $z$. Assuming $X$ is bounded, an end vertex is either a *convex vertex* of $X$ (point of discontinuity of the derivative $\alpha'$ on $\partial X$ with interior angle less than $\pi$) 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 $\alpha$ is positive and locally maximum at the projection Choi et al. (1997).

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 $\alpha'$ on $\partial X$ with interior angle greater than $\pi$) or the curvature is negative and locally minimum at the projection.



(a) interior medial axis (b) exterior medial axis

**Fig. 2. Medial axes of shape `#186` from S&V dataset De Winter and Wagemans (2006) which we use in experiments of section 8. (a) Interior medial axis. (b) Exterior medial axis.**

### 2.2. Computation

In practice, we compute the distance map with any algorithm that provides at least one representative of the projection $\pi(z)$ of each point Felzenszwalb and Huttenlocher (2004), and then compute the medial axis using the *chord residue* Ogniewicz and Ilg (1992); Avrithis and Rapantzikos (2011). Given two points $x, y \in \partial X$, the arc length $\ell(x, y)$ is the length of the minimal arc of $\partial X$ having $x, y$ as endpoints or $\infty$ if no such arc exists. Now, given a point $z \in \mathcal{M}(X)$, its chord residue

$$r(z) = \sup_{x,y \in \pi(z)} \ell(x, y) - \|x - y\| \tag{4}$$

is the maximal 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 $\sigma > 0$ Ogniewicz and Ilg (1992); Avrithis and Rapantzikos (2011).

This method 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 $\sigma$, without

simplifying the curve $\partial X$ in any way. Typically, $\sigma$ is only 1-2 pixels just to remove discretization noise. Unfortunately, it is constrained to two dimensions.

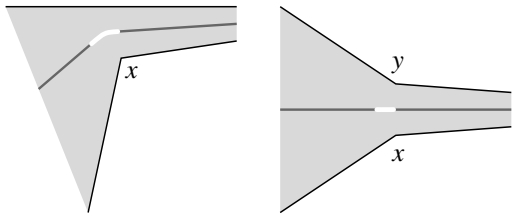
In the following, we assume that both the interior and exterior medial axes are available. Both are computed by a single traversal over a discrete representation of the plane on a regular grid. This operation is linear in the size of the representation, while the arc length is computed in constant time Ogniewicz and Ilg (1992). Fig. 2a,b, illustrates the two medial axes for a sample shape that will also serve as a running example in section 3 below.

For simplicity, we assume that for each point $z$ on the medial axis, the projection $\pi(z)$ contains exactly two points; in practice, only one projection point is stored for each $z$, while the second one is obtained from $z$'s neighbors. In fact, our prior implementation Avrithis and Rapantzikos (2011) yields a medial axis that is two pixels thick everywhere, so that two neighbors are always to be found. According to this assumption, given a point $z$ of the interior medial axis with projections $\pi(z) = \{x, y\}$, we define the *arc length* and *chord length* of $z$ as the corresponding arc length $\ell(x, y)$ and chord length $\|x - y\|$ between its two projections $x, y$.

## 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 $\partial X$. In some cases, cuts have been defined as curves, *e.g.* cubic splines, providing for *continuation* of boundary tangents at end points Siddiqi and Kimia (1995); but in all work discussed in section 1 or compared to in section 8, as well as in the current work, cuts are just line segments for simplicity Hoffman and Singh (1997); Singh and Hoffman (2001). In either case, the cut endpoints always lie on the boundary $\partial X$ and the cuts lie entirely on the closure of $X$ Singh and Hoffman (2001). Additional constraints 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 decomposition process up to selecting raw cuts is detailed below.




(a) semi-ligature (b) full ligature

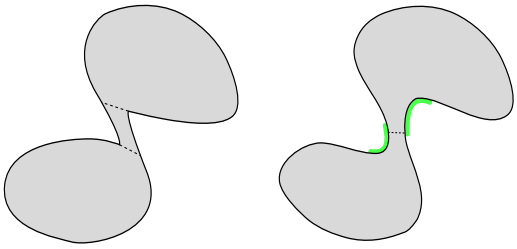
**Fig. 3. (a) Semi-ligature on $x$. (b) Full ligature on $x, y$ (in white) August et al. (1999).**

### *3.1. Minima, maxima and corners*

**Background.** According to the *minima rule* Hoffman and Richards (1984), the shape $X$ should be cut at points of negative minima of curvature of its boundary parametrization $\alpha$. In the theory of limbs and necks Siddiqi and Kimia (1995), 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* endpoint has negative curvature Singh et al. (1999). This condition is contained in the standard definition of part-cuts Singh and Hoffman (2001). This is in agreement with the earlier theory of ligatures August et al. (1999) and more recent studies Luo et al. (2015).

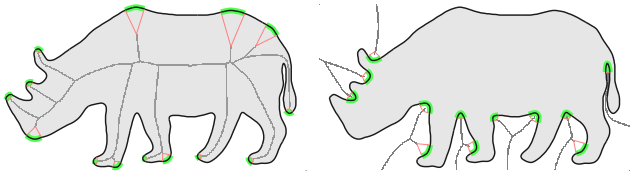
In particular, given a set of minima points $C$, a *full-ligature* (resp. *semi-ligature*) August et al. (1999) on two points $x, y \in C$ (resp. one point $x \in C$) is the set of interior medial axis points $z$ in whose projection $\pi(z)$ contains $x, y$ (resp. $x$ but no other point of $C$). Commonly referred to as ligatures, these sets are subsets of the interior medial axis and disconnect it such that subsequent shape reconstruction produces a rough decomposition into parts. They are illustrated in Fig. 3a,b. Accordingly, following Luo et al. (2015), we define *double cuts* (resp. *single cuts*) as the line segments 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? According to our knowledge, all relevant studies assume a discrete



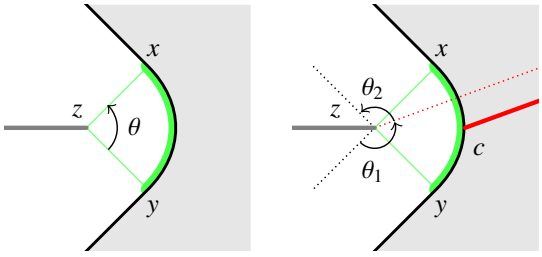
(a) sharp (b) weak

**Fig. 4. (a) Two nearby sharp concavities result in two different cuts Singh and Hoffman (2001). (b) Two nearby weak concavities should ideally result in one cut; this is possible if their *locale* Hoffman and Singh (1997) is known (in green).**



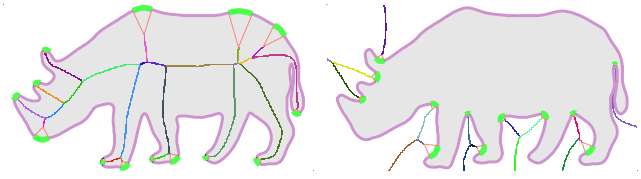
(a) convex corners (b) concave corners

**Fig. 5. (a) Convex corners obtained from the interior medial axis. (b) Concave corners (or simply corners) obtained from the exterior medial axis. Corners are shown in green. Pink lines connect the end-vertices of the medial axis to their projections.**




(a) interior angle (b) local convexity

**Fig. 6. (a) Interior angle $\pi + \theta$ of a concave corner (in green) determined by exterior medial vertex $z$ and its projection points $x$, $y$, where $\theta$ is the angle between the two line segments from $z$ to $x$, $y$. (b) By translating the two boundary segments starting at $x$, $y$ and the cut (in red) starting at $c$ to the same origin (vertex $z$ here), shown as dotted lines, we measure the interior angles $\theta_1$, $\theta_2$ of the two shape parts at this corner after cutting. Both are less than $\pi$, while $\theta_1 + \theta_2$ is not. Local convexity is achieved and there is no need for more cuts at this corner. See section 7.**



(a) interior branches (b) exterior branches

**Fig. 7. (a) Interior branches from convex corners and interior boundary components. (b) Exterior branches from concave corners and exterior boundary components. Boundary components, corners and branches are shown in purple, green and random color respectively. Pink lines connect end-vertices to their projections.**

parametrization of shape boundary $\partial X$ 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* Hoffman and Singh (1997) and illustrated in Fig. 4a,b.

**Our solution.** The background of section 2 specifies that end-vertex projections of the medial axis are either single points tangent to osculating circles, or circular arcs. In practice, the two projections determine a boundary arc that always approximates a circular arc. The radius of the circle is the inverse of the absolute curvature. In the case of the interior medial axis, the curvature is locally maximized on this arc (respectively minimized in the case of the exterior medial axis), which makes this arc particularly suitable for detecting a convexity (resp. concavity). We call this arc a *convex corner* (resp. a *concave corner* or simply *corner*). We give an illustration in Fig. 5. The three points involved—the end vertex and its two projections—directly determine the position, spatial extent, orientation and strength of the convexity (resp. concavity). All information comes for free from the medial axis.

The strength of a convexity (or concavity) can be measured in terms of both curvature and interior angle. Given an arc with endpoints $x$, $y$ that is a subset of a convex (resp. concave) corner specified by interior (exterior) medial axis end vertex $z$, we define its *interior angle* as $\theta$ (resp. $\pi + \theta$), where $\theta$ is the angle between the two line segments from $z$ to $x$, $y$. Its *curvature* is the inverse of the length of any of these line segments (all such lengths are equal). See Fig. 6a.

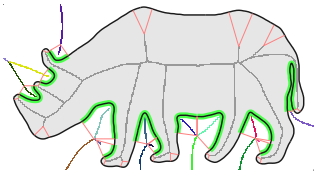


**Fig. 8. Extended corners, in green. Pink lines connect the end-vertices of the interior medial axis to their projections and the last visited points in the backward traversal of the exterior medial axis to their projections. Branches of the exterior medial axis are shown in random color.**

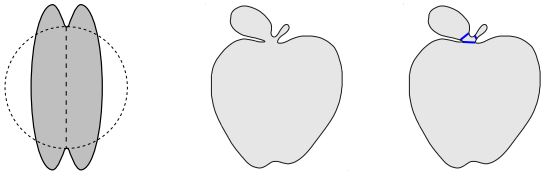
### *3.2. Branches*

For each medial axis, we also parse its graph structure by a single traversal. We refer to the edges of the graph as medial axis *branches*. First, we detect the end-vertices of interior (resp. exterior) medial axis. For each end-vertex, we compute the corresponding convex (resp. concave) corner. Subsequently, we subtract the set $C$ of convex (resp. concave) corners from the boundary $\partial X$ of the shape. We find the connected components of the difference $\partial X \setminus C$, which we call *interior* (resp. *exterior*) *boundary components*.

Now, interior (resp. exterior) medial axis points with projections belonging to the same pair of interior (resp. exterior) components, belong to the same *interior* (resp. *exterior*) *branch*. Additionally, interior (resp. exterior) medial axis points having at least one (8-connected) neighbor belonging to a different interior (resp. exterior) branch, are called *interior* (resp. *exterior*) *junctions*. The entire traversal operation is linear in the number of medial axis points. Fig. 7a,b illustrate the interior (resp. exterior) boundary components and branches.

### *3.3. Extended corners*

In this work we also introduce the concept of extended corners. For each end-vertex of exterior medial axis, we traverse the corresponding branch backwards until we meet a junction point or a point whose at least one projection lies on a convex corner. Similarly to the corners discussed above, the projections of the point visited last in this traversal determine an *extended concave corner* or simply *extended corner*. Because traversal is limited to a single branch there is a unique extended corner for each corner and for each end-vertex. Fig. 8 illustrates the extended corners.



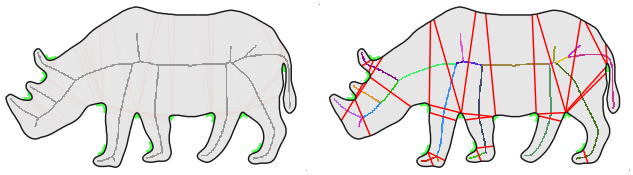
(a) local symmetry (b) all subjects (c) majority cuts

**Fig. 9. (a) Example from Singh and Hoffman Singh and Hoffman (2001) illustrating that a cut across a local symmetry axis fails to be captured by the medial axis or equivalently by the definition of neck Siddiqi and Kimia (1995) because a circle cannot be inscribed. (b) A counter-example of shape `#006` from ground-truth data of DeWinter and Wagemans De Winter and Wagemans (2006) where most subjects do not cut in a similar case. Cuts of all subjects are overlaid in blue, 85% transparent. (c) Majority cuts of (b), in blue (see text).**

### *3.4. Symmetry*

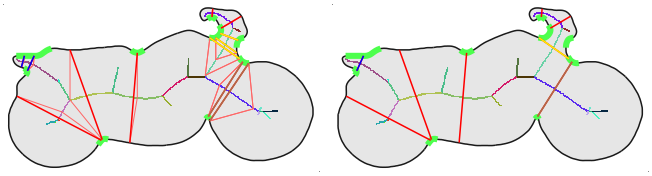
**Background.** According to the minima rule, all pairs of points on (distinct) corners are potential cuts. Several methods actually examine all pairs Siddiqi and Kimia (1995); Liu et al. (2010); Ren et al. (2011); Ma et al. (2013); Luo et al. (2015), hence are at least quadratic in the number of samples of the boundary. More importantly, they may involve solving an optimization problem or introduce new rules to resolve conflicts (*e.g.* that cuts do not intersect). But the standard definition of part-cuts Singh and Hoffman (2001) includes the additional condition that they cross an axis of local symmetry. We modify the condition such that the cut endpoints are projections of the same point of the interior medial axis (recall that a cut lies in the shape). This is in line with the definition of ligatures August et al. (1999).

In most cases, this is a stronger condition. Combined with the minima rule, it implies that endpoints are exactly projections of the same point of a ligature. We observe that this condition most often agrees with ground truth data from psychophysical experiments De Winter and Wagemans (2006), as shown in Fig. 9a-c. This figure illustrates human annotation, in particular cuts specified by all subjects and *majority cuts* for which most subjects



(a) raw cuts (b) representative cuts

**Fig. 10. (a) *Symmetry*: all cuts (in red, 95% transparent) whose endpoints are projections of the same interior medial axis point, with at least one endpoint on a concave corner. (b) *Strong equivalence*: representative cuts (in red) are selected such that for each corner there are at most two cuts (one double and one single) per interior medial axis branch. Corners and branches are shown in green and random color respectively.**



(a) strong equivalence (b) weak equivalence

**Fig. 11. *Weak equivalence*: representatives of (a) strong and (b) weak equivalence classes. In (a), cuts in blue, yellow and brown form three weakly equivalent classes, while the remaining representatives are in red. In the blue and brown class there are only two cuts (one double and single), while in the yellow class there are three cuts (two double and one single). In (b), we select the double cut as representative of the blue and brown class. For the yellow class, we select the double cut having the maximum number of votes. Corners and branches are shown in green and random color respectively. Pink lines connect cut points to their projections.**

agree. The latter are found according to a clustering-based ensemble Lewin et al. (2012a); see section 8 for more details.

**Our solution.** So what we do in practice is, traverse the interior medial axis once, and collect all pairs of projections such that at least one lies on a corner. The line segments between these pairs of points are called *raw cuts* and are illustrated in Fig. 10a. It is easily shown that they do not intersect by construction. Depending on the number of corners, we call the cuts *double* or *single*.

To each cut we assign the corresponding interior medial axis point. We call this point the *cut point* and its chord residue the *cut residue*. Recall that a cut is a line segment in this work so by *cut length* we refer to the length of this line segment. We say that a cut *lies on a corner* (or arc in general) if one of its endpoints lies on this corner (or arc). We also say a cut *lies on a branch* if its cut point lies on this branch. Given a cut $c$ with endpoints $x, y$, we define its *minimal arc*, denoted by $\text{arc}(c)$, as the minimal arc of $\partial X$ having $x, y$ as endpoints.

## 4. Equivalence

The selection of candidate cuts for shape decomposition out of all raw cuts is based on two equivalence relations and the choice of one representative per equivalence class.

### *4.1. Strong equivalence*

Observing Fig. 10a, raw cuts are clearly too many, but they tend to appear in groups. As shown in Fig. 10b, we select a small number of *representative cuts* by defining a strong equivalence relation on raw cuts and selecting one representative from each equivalence class. We say that two cuts are *strongly equivalent* or simply *equivalent* if they (a) are both double or both single, (b) lie on the same branch and (c) lie on the same set of corners. This rule is intuitive and always maintains all correct cuts in our experiments.

Fig. 10b shows that whenever two groups of cuts lie on a common 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. The representative cut is chosen such that its endpoints are closest to the midpoint of the corner(s).

### *4.2. Weak equivalence*

We now consider representative cuts of strong equivalence classes that lie on (a) the same branch and (b) a common corner. We say that such cuts are *weakly equivalent*. Clearly, two strongly equivalent cuts are also weakly equivalent. As shown in Fig. 11a, these cuts have approximately the same minimal arc on the boundary. Fig. 11b illustrates representative cuts of weak equivalence classes. We call these representatives *candidate cuts* or simply *cuts* in the following.

Similarly to strong equivalence discussed above, we choose

only one representative cut from each weak equivalence class. The choice only refers to the case where there is at least one double cut in a weak equivalence class; otherwise, the cuts are strongly equivalent and have been processed in advance. In this case, the representative cut is always double and is chosen as the one having the maximum number of votes. The *votes* of a point on the boundary $\partial X$ is the number of cuts (single or double) having this point as an endpoint. The *votes* of a cut is the sum of votes of its endpoints. This voting process is inspired by *Hough transform*, where the voting space is the boundary $\partial X$. It is a local measure for concavity that is more detailed than the definition of a corner because it lets us specify a precise point rather than an entire arc.

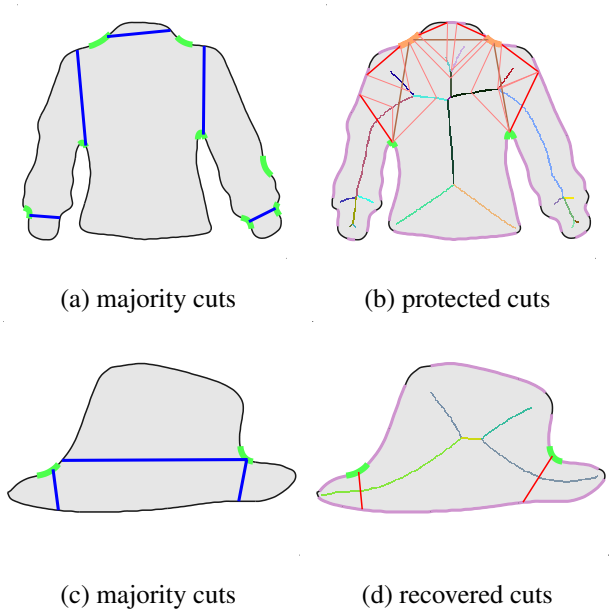
## 5. Protected and recovered cuts

Given a set of candidate cuts, the final selection is based on local convexity discussed in section 7, using salience measures discussed in section 6. Here we introduce two mechanisms to protect cuts from being discarded during the local convexity process and to recover cuts that cannot be found as raw cuts in the first place.

### 5.1. Protected cuts

As discussed in section 7, we apply our local convexity rule to select cuts at each corner, until local convexity is achieved. But there are cases where local convexity has been achieved at a corner before certain essential cuts according to majority cuts discussed in section 8 are found. To maintain these cuts regardless of the local convexity rule, we mark them as *protected*.

Let $a$ and $b$ be two cuts. We say $a$, $b$ are *disjoint* if $\text{arc}(a)$ and $\text{arc}(b)$ are disjoint i.e. $\text{arc}(a) \cap \text{arc}(b) = \varnothing$. Otherwise, we say that $b$ *contains* $a$ if $\text{arc}(a)$ contains $\text{arc}(b)$ i.e. $\text{arc}(a) \subset \text{arc}(b)$. Now, suppose we have a pair of disjoint single cuts satisfying the following properties: (a) they do not share the same corner, (b) their endpoints share only one interior component, and (c) their cut points are on branches which meet at a junction. Suppose also that there is a double cut containing these two cuts and that its endpoints lie on the same two corners of these cuts. We call this double cut *protected* and the corresponding single cuts its *generators*. In practice, we detect protected cuts in a single traversal of interior medial axis. Fig. 12a,b illustrates majority and protected cuts respectively.



(a) majority cuts (b) protected cuts

(c) majority cuts (d) recovered cuts

**Fig. 12. (a) Majority cuts of shape `#029` from S&V dataset De Winter and Wagemans (2006), in blue. (b) Local convexity at orange corners can be achieved by selecting only one of the brown cuts. However, in agreement with (a), all three brown cuts are marked as protected, as found via their generator cuts, in red. Remaining corners are shown in green. Interior boundary segments and branches are shown in purple and random color respectively. Pink lines connect the cut points with their projections. (c) Majority cuts of shape `#118`. (d) The cuts in red can be found directly as raw cuts, but the brown one cannot. However, in agreement with (c), it is marked as recovered, as found via its generator cuts, in red. Remaining colors as in (b).**

### 5.2. Recovered cuts

Although the detection of cuts using interior medial axis often agrees with the ground truth De Winter and Wagemans (2006), it has a weakness. In particular, considering the three majority cuts in Fig. 12c, only two can be found directly as shown in Fig. 12d. This is due to the construction of medial axis whereby there is no interior medial axis point having projections on both corners—a circle cannot be inscribed in the shape. Such cuts cannot be found in the set of raw cuts and accordingly among the equivalence class representatives, but we recover them as follows.

Suppose we have a pair of disjoint single cuts satisfying the following properties: (a) they do not share the same corner, (b) their endpoints share only one interior component. For this pair of single cuts, we create a new double cut between the two single cut endpoints lying on corners. We call such a double cut *recovered*. Similarly to protected cuts, we call the corresponding single cuts its *generators* and we detect recovered cuts in a single traversal of the interior medial axis. Fig. 12c,d illustrates majority and recovered cuts respectively.

## 6. Salience measures

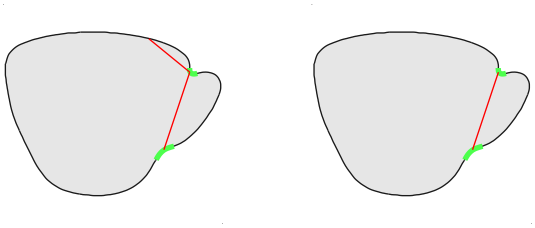
In our previous work Papanelopoulos and Avrithis (2015), the local convexity rule selects the appropriate number of part-cuts independently per corner, so it is also completely blind to their prioritization. In this work, before we apply our local convexity rule discussed in section 7, we first discard a number of representative cuts and prioritize them using a number of salience measures.

**Discussion.** Although there is no complete theory, several such measures have been suggested as plausible in the literature, going back to at least Gestalt psychologists Hoffman and Singh (1997); Singh and Hoffman (2001). These refer to boundary strength at cut endpoints, including turning angle for cusps and normalized curvature for smooth boundary Hoffman and Singh (1997), continuation of boundary at endpoints Singh and Hoffman (2001), as well as of salience of cuts or parts themselves, including relative area, protrusion Hoffman and Singh (1997), and cut length Singh et al. (1999).

In the following, we describe the salience measures we use in this work: *protrusion strength*, *flatness*, *expansion strength* and *extension strength*. Protrusion strength is known, while the remaining are new. As discussed in section 7.1, only protrusion and extension strength are used for the prioritization of cuts.

### 6.1. Protrusion strength

Observing Fig. 13a, the length of the minimal arc between the endpoints of the single cut is not much greater than the length of the cut. The minimal arc corresponds to an insignificant protrusion of the boundary $\partial X$. This cut does not improve the decomposition of the shape and is insignificant.



(a) before protrusion strength (b) after protrusion strength

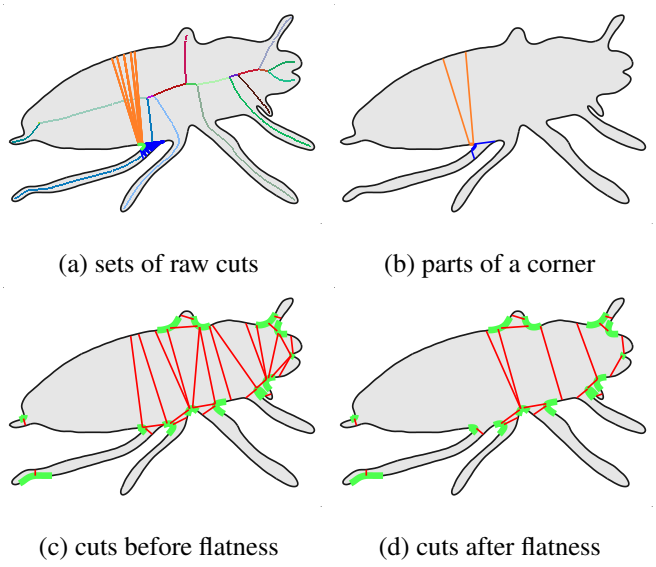
**Fig. 13. Cuts (a) before and (b) after discarding cuts according to protrusion strength. Corners are shown in green and cuts in red.**

For this reason, following Hoffman and Singh (1997) and in particular the simpler definition of Zeng et al. (2008), we define the *protrusion strength* of a cut as the ratio of its length to the length of its minimal arc. This ratio takes values in the interval $[0, 1]$. The lower this ratio, the more salient a cut is. A cut having protrusion strength above a threshold $\tau_a$ is discarded if it is single and prioritized if it is double, as discussed in section 7.1. Fig. 13b illustrates the remaining cuts after this process.

### 6.2. Flatness

A corner may have non-negligible spatial extent and several cuts lying on it, not all of which are equally important. For this reason, we split corners into parts and then discard cuts lying on certain parts. In particular, for a given corner, we partition all raw cuts lying on the corner according to their branch. From each set in the partition, we select the two cut endpoints that lie on this corner and are closest to the endpoints of the corner. The minimal arc on the boundary $\partial X$ between the selected endpoints is defined as a *part* of the corner.

For each part, we compute its interior angle and we normalize all angles such that the maximum is one. We call this measure *flatness*; it also takes values in $[0, 1]$. The higher this measure, the more salient a part is. A part is called *flat* if this ratio is below a threshold $\tau_b$. We discard cuts having at least one endpoint lying on a flat part. Fig. 14 illustrates this selection process.



(a) sets of raw cuts (b) parts of a corner

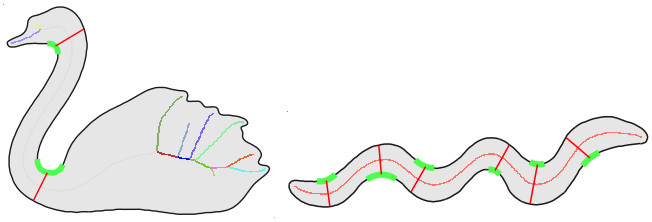
(c) cuts before flatness (d) cuts after flatness

**Fig. 14. (a) Two sets of raw cuts lying on a corner, shown in orange and blue. The corresponding cut points lie on two different branches. Branches are shown in random colors. (b) The two selected raw cuts from each set and the corresponding parts of the corner, shown again in orange and blue. Cuts (in red) (c) before and (d) after discarding according to flatness. Corners are shown in green.**

### 6.3. Expansion strength

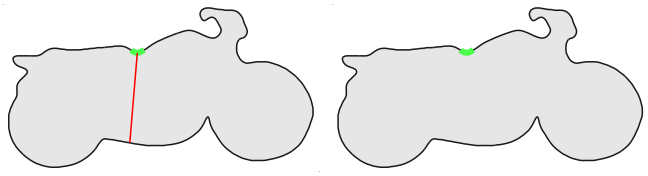
**Background.** According to Luo et al. (2015), a single cut should be *expanding* on at least on side of the cut. This means that given a single cut lying on a corner that contains only one part, the cut length of all raw single cuts lying on the same corner varies significantly and increases at least on one side. Otherwise, if the length is roughly constant, the cut is *non-expanding*. The cuts in Fig. 15a are expanding, while in Fig. 15b are non-expanding and must be discarded.

**Our solution.** Instead of using neighborhood histograms to discard these single cuts as in Luo et al. (2015), we simply use their distance map value. In particular, given a single cut lying on a corner that contains only one part, we compute the ratio of the minimum to the maximum distance map value of the cut point over all raw cuts lying on the part. Inspired from Luo et al. (2015) we call this ratio *expansion strength*; again, it takes values in $[0, 1]$. The lower this ratio, the more salient a cut is. We discard cuts having expansion strength above $1 - \tau_b$.



(a) expanding cuts (b) non-expanding cuts

**Fig. 15. (a) Expanding single cuts are not discarded. (b) Non-expanding single cuts are discarded. All cuts are shown in red. Corners and branches are shown in green and random color respectively.**



(a) before extension strength (b) after extension strength

**Fig. 16. Cuts lying on an extended corner (a) before and (b) after discarding according to extension strength. Cuts are shown in red and extended corners in green.**

### 6.4. Extension strength

Given a concave corner, we define its *extended arc length* as the arc length of the corresponding extended corner. Then, given a cut lying on a corner, we define the ratio of its extended arc length to the cut length as its *extension strength*. This measure takes values in $(0, \infty)$. The higher this ratio, the more salient is this cut with respect to the particular corner.

Single and double cuts have one and two extension strength(s) respectively. Single cuts having extension strength below a threshold $\tau_c$ are discarded. A double cut having only one extension strength ratio below $\tau_c$ is marked as single and we consider that only one endpoint lies on a corner. Otherwise, if both extension strength ratios are below $\tau_c$, a double cut is discarded.

Fig. 16a,b illustrates this process for a single cut. In this example, and in most shapes in S&V dataset, corners and extended corners are nearly the same for this rule to apply. Still, in general, extended corners are a more robust option in case *e.g.* corners are very sharp because they express how much a concavity penetrates the shape.

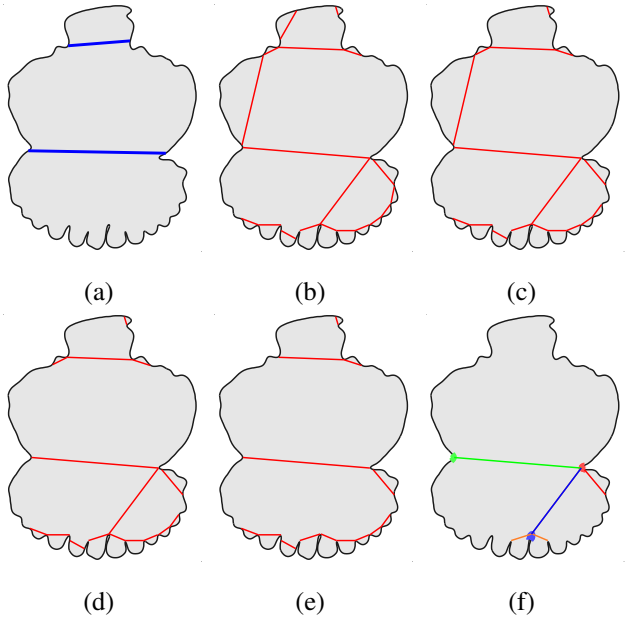



**Fig. 17.** ***Local convexity*****. (a) Majority cuts of shape `#063` from S&V dataset. (b) Cuts (in red) before applying the local convexity rule. (c) Final selection of cuts without penalization and prioritization only according to short-cut rule, as in Papanelopoulos and Avrithis (2015). It is clear that more cuts are selected than what would be needed to achieve local convexity at each corner. (d) Final cuts with prioritization as defined in section 7.1. (e) Final cuts with priorities initialized as in section 7.1 and dynamically updated by penalization as discussed in section 7.2. Less cuts are selected comparing to (c). (f) The *penalization* process. All illustrated cuts have the same priority. The blue corner precedes the red one and as a result cuts lying on it are examined first. When we select the two orange cuts for the blue corner, local convexity is already achieved. If we did not penalize the blue cut, we would eventually select the red and blue cut to achieve local convexity at the red corner, because these cuts are shorter. With penalization, we rather select the green and red cut for the red corner. This works because corners far away from the shape center are examined first.**

## 7. Local convexity

**Background.** 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 but also because according to the *robustness* requirement Siddiqi and Kimia (1995), 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.

**Our solution.** 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 at each corner in order to achieve *local convexity* at the corner. In particular, for every corner, we prioritize all cuts lying on the corner as discussed in section 7.1, and we select cuts by descending priority until the interior angle of all parts after cutting is less than $\pi + \phi$, where $\phi$ is a tolerance. The process is illustrated in Fig. 6b. Once more, all information is readily available from the medial axis. Fig. 17a,b illustrate respectively majority cuts and cuts before applying our local convexity rule.

In our previous work Papanelopoulos and Avrithis (2015), the local convexity rule is based on the following: (a) for each corner, we select cuts independently of other corners; (b) according to the short-cut rule Singh et al. (1999), cuts lying on a corner are prioritized by ascending order according to their cut length. The selection is such that corners can be examined in any order. Unfortunately, these choices often lead to a final selection of cuts that is not consistent with majority cuts as illustrated by comparing Fig. 17a,c. We therefore reconsider the selection process in this work, as discussed below.

### *7.1. Initial prioritization*

We initialize the priority of single and double cuts to zero and one respectively. We then increase by one the priority of double cuts that are disjoint with all other cuts on either corner. On the other hand, we decrease by one the priority of cuts having protrusion strength above $\tau_a$ or extension strength below $\tau_c$. Finally, we set the priority of generator cuts to $-\infty$. In case two cuts have the same priority, cuts with shorter cut length are examined first.

### *7.2. Selection process*

By using multiple prioritization criteria, our cut selection process is no longer independent of the order in which we examine corners as in our previous work Papanelopoulos and Avrithis (2015). For this reason, we also specify a particular order in this work.

We remind that the residue attains a global maximum on a single point of the medial axis, which we call the *center* of the shape, and is decreasing as we move away from this point on the medial axis Ogniewicz and Ilg (1992). We use this property such that corners far away from the center are examined first. In particular, for each corner, we define its *residue* as the maximum residue over cut points of all raw cuts lying on this corner—recall that cut points lie on the medial axis. We examine corners by ascending order of residue.

Next, at each corner we first select all protected cuts unconditionally and then examine the rest by descending priority until local convexity is achieved. If local convexity has been achieved and there are remaining cuts to be examined, we *penalize* them by decreasing their priority by one. This means that a single cut is effectively discarded, while a double cut may be examined at lower priority on the second corner it lies on. When local convexity has been achieved at all corners, all remaining cuts are discarded. Fig. 17e,f illustrates the final selection of cuts with penalization and the penalization process respectively. This example justifies why it makes sense to examine first corners that are far away from the center.

## 8. Experiments

### *8.1. Experimental setup*

#### *8.1.1. Datasets*

In most related work Siddiqi and Kimia (1995), even in recent methods Mi and Decarlo (2007); Zeng et al. (2008); Liu et al. (2010), evaluation is only qualitative, while quantitative evaluation is often limited to datasets that are not public like arbitrary subsets of MPEG-7 shape dataset Ren et al. (2011); Ma et al. (2013). To our knowledge, there are two public datasets with ground-truth from human subjects Liu et al. (2014); De Winter and Wagemans (2006). 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 decomposition of object outlines.

This dataset is a subset of the Snodgrass and Vanderwart (S&V) everyday object dataset Snodgrass and Vanderwart (1980), consisting of 260 line drawings. The subset refers to 88 of the drawings, which have been converted to smooth outlines and each decomposed by 39.5 subjects (psychology students) on average. For each shape there are 122.4 part-cuts on average, that is 3.1 cuts per subject. The same dataset, referred to as S&V, has been subsequently used for quantitative comparison of different computational models Lewin et al. (2012a,b); Luo et al. (2015). An example illustrating the cuts of all subjects on a single outline is shown in Fig. 9b.

For qualitative comparisons, we also use the Kimia dataset Siddiqi and Kimia (1995). This contains shapes decomposed by 14 subjects each as well as a ground-truth majority decomposition per shape. In the absence of published quantitative results on some consistent evaluation protocol, here we only focus on a small number of shapes that allows visual comparison to examples found in the bibliography.

#### *8.1.2. 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 Siddiqi and Kimia (1995). There are different alternatives, which take the form of either a majority decomposition by clustering Lewin et al. (2012b); Liu et al. (2014), or spatial density used directly for evaluation Luo et al. (2015). We follow the framework of Lewin *et al*. Lewin et al. (2012b). 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

$$d(c_1, c_2) = \min\{\ell(x_1, x_2) + \ell(y_1, y_2), \ell(x_1, y_2) + \ell(y_1, x_2)\}, \quad (5)$$

where $\ell$ 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. 9b,c.

### 8.1.3. Evaluation measures

Unfortunately, since quantitative evaluation is relatively new, there is nearly one different protocol for every relevant publication. Here we follow the framework by Lewin *et al*. Lewin et al. (2012b), which facilitates comparisons to a number of existing methods. We use two different measures, both of which assume a decomposition of shape $X$ is represented by a partition $A = \{A_i\}$ of $X$. The *Hamming distance* Lewin et al. (2012b) of partitions $A, B$ is then

$$d_H(A, B) = \frac{1}{2|X|}[h(A|B) + h(B|A)], \tag{6}$$

where $|X|$ is the area of $X$, $h(A|B) = \sum_i |A_i \setminus B_{\pi_i}|$ is the sum over all parts of $A$ of the area of part $A_i$ not covered by its best match $B_{\pi_i}$ in $B$, and the best match is defined by $\pi_i = \arg\max_j |A_i \cap B_j|$. In practice, both $X$ and each part $A_i$ are represented by binary masks on a discrete 2d grid and area is measured in pixels.

Another measure is the Jaccard distance, referred to as *Jaccard measure* in Lewin et al. (2012b) and defined as follows. Assuming $X$ is a finite set represented as $\{x_i\}$, let $P = \{(x_i, x_j) \in X^2 : j > i\}$ be the set of ordered pairs of points in $X$. Let also $P_A = \{(x, y) \in P : A(x) = A(y)\}$ be the pairs of points in $P$ that are in the same part of $A$, where $A(x)$ is the part of $A$ where point $x \in X$ belongs. Then, the *Jaccard index* or *intersection over union* of partitions $A, B$ is given by

$$J(A, B) = \frac{|P_A \cap P_B|}{|P_A \cup P_B|}, \tag{7}$$

and their *Jaccard distance* by

$$d_J(A, B) = 1 - J(A, B) = \frac{|P_A \triangle P_B|}{|P_A \cup P_B|} \tag{8}$$

where $\triangle$ denotes symmetric set difference.

Given a number of ground truth decompositions per shape, each by a different human subject, we follow Lewin et al. (2012b) in defining two different evaluation measures over a dataset of shapes:

- *Majority*: this is the Hamming or Jaccard distance between the decomposition computed by a method and the majority decomposition, averaged over all shapes.
- *Average*: this is the Hamming or Jaccard distance between the decomposition computed by a method and an individual subject's decomposition, averaged over all subjects and all shapes.

### 8.1.4. Compared methods

Our own method is referred to as *medial axis decomposition* (MAD*). We perform quantitative comparison to our previous work Papanelopoulos and Avrithis (2015), referred to as MAD, to the clustering-based ensemble (CBE) method Lewin et al. (2012a), and to five individual methods, namely approximate convex decomposition (ACD) Lien and Amato (2004), discrete contour evolution (DCE) Latecki and Lakamper (1999), combined skeleton-boundary features (SB) Zeng et al. (2008), flow discretization (FD) Dey et al. (2003) and constrained morphological decomposition (MD) Kim et al. (2005).

CBE is applying to the five latter individual methods the same clustering approach that is also applied to human subject decompositions as part of majority voting; it is therefore an ensemble decomposition method. Quantitative results on CBE and the five individual methods are reported as provided by Lewin et al. (2012b), where all methods have had their parameters optimized quantitatively on the S&V dataset. We also compare to human subjects, each evaluated individually using either majority or average evaluation, exactly like automated methods Liu et al. (2014). Finally, we compare to the baseline case of not cutting anywhere.

Qualitative results, apart from MAD and CBE, are additionally compared to relatability (REL) Mi and Decarlo (2007), convex shape secomposition (CSD) Liu et al. (2010), minimum near-convex decomposition (MNCD) Ren et al. (2011) and computational model of short-cut rule (CSR) Luo et al. (2015). Human ground truth (GT) per shape is given as a single decomposition for the Kimia dataset Siddiqi and Kimia (1995) and as an overlay of all subjects' cuts for S&V De Winter and Wagemans (2006).

### 8.2. Results

#### 8.2.1. Timing

Medial axis computation, implemented in C++, takes on average 84ms per S&V shape at a resolution of $500 \times 500$ on an AMD A8-4500M processor at 1.9GHz. On the other hand, MAD*, implemented in Matlab, takes on average 693ms. There has been no effort to optimize the code. In fact, medial axis computation is linear in the number of pixels while MAD* is linear in the number of points on the (external and internal) medial axis, which is a much smaller subset.

#### 8.2.2. Parameter tuning

There are five parameters in MAD*: medial scale threshold $\sigma$, convexity tolerance $\phi$ and the thesholds $\tau_a$, $\tau_b$, $\tau_c$ for the (inverse of) protrusion strength, flatness and extension strength respectively. Medial scale $\sigma$ is measured in pixels and although it is not scale invariant it is only meant to deal with discretization noise and $\sigma = 2$ or 3 pixels is known to work smoothly for medial axis computation Avrithis and Rapantzikos (2011). The remaining parameters are dimensionless, scale-invariant quantities: $\phi$ is an angle and $\tau_a$, $\tau_b$, $\tau_c$ are ratios in $[0, 1]$.

We perform grid search to find the optimal values of our parameters. *This is possible because the parameters are only a few and the dataset is quite small.* In particular, we tune the values of $\sigma$, $\phi$, $\tau_a$, $\tau_b$, $\tau_c$ in the interval [1,4], [0,120], [0,1], [0,0.2], [0,1] respectively. We uniformly sample each interval with a step of 0.6, 15, 0.2, 0.15, 0.04 respectively between samples. We find the performance to be globally optimal according to majority evaluation with respect to both Hamming and Jaccard distance for $\sigma = 2.8$, $\phi = 90°$, $\tau_a = 0.6$, $\tau_b = 0.75$ and $\tau_c = 0.16$. We refer to our method with this set of parameters as MAD*-opt. Fig. 18 illustrates quantitative results for different configurations of $\sigma$, $\phi, \tau_a, \tau_b$, $\tau_c$, while keeping the remaining parameters fixed to their optimal value.

Although not shown here, we have found the dependence of performance to different parameters to be largely uncorrelated, that is, slightly changing the value of one parameter does not affect much the local minimum of the others. This indicates that different rules are largely independent and complementary. To further investigate this, we initialize each parameter as arising from the development of the associated rule and we manually fine-tune them one by one by qualitative inspection on a few random examples, working only once with each parameter in random order. The interpretability of the rules helps in knowing exactly what to observe for each one.

This yields the sub-optimal set of values $\sigma = 2.8$, $\phi = 60°$, $\tau_a = 0.68$, $\tau_b = 0.75$, $\tau_c = 0.174$. We refer to our method with this set of parameters as MAD*. We use MAD* (resp. MAD for our prior work Papanelopoulos and Avrithis (2015)) for qualitative evaluation, and report both measurements MAD* and MAD*-opt (resp. MAD and MAD-opt for our prior work Papanelopoulos and Avrithis (2015)) in quantitative evaluation. The discrepancy between quantitative and qualitative fine-tuning can be attributed to limitations of the evaluation measures used Liu et al. (2014).

#### 8.2.3. Quantitative evaluation

Table 1 compares our method MAD* to a number of relevant methods and our previous work MAD (Papanelopoulos and Avrithis, 2015). On all measurements, MAD* outperforms all individual methods and human subjects, and is on par with CBE in Hamming distance on average evaluation; while MAD*-opt is always better than all methods and human subjects. The relative gain (decrease) of MAD* (resp. MAD*-opt) over CBE on majority evaluation is 5.8% Hamming and 8.0% Jaccard (resp. 8.7% Hamming and 10.8% Jaccard). CBE is an expensive ensemble method that involves all five methods that precede it in the Table 1. 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.

Against our previous method MAD (Papanelopoulos and Avrithis, 2015), we achieve a relative gain of 31% Hamming and 32% Jaccard on majority evaluation. The gain on average evaluation is 12% Hamming and 11% Jaccard. More importantly, MAD* is superior to CBE, while MAD is not. Human results are not very competitive, which is expected as we average the performance over subjects, and subjects are not always consistent with each other. It is interesting that SB and DCE

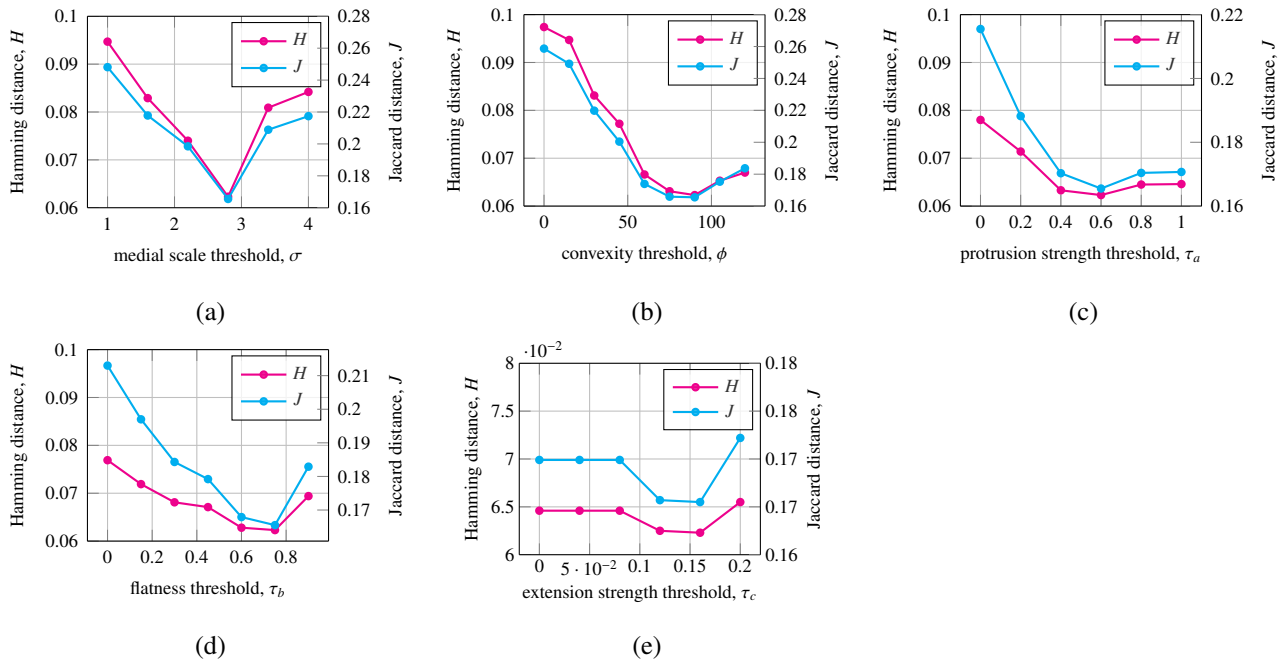



**Fig. 18.** ***Parameter tuning*****. Hamming (*H*) and Jaccard (*J*) distance *vs*. thresholds of (a) medial scale $\sigma$, (b) convexity $\phi$, (c) protrusion strength $\tau_a$, (d) flatness $\tau_b$, (e) extension strength $\tau_c$. Majority evaluation on S&V dataset. Lower is better for both evaluation measures. In each plot we fix the remaining parameters to their optimal value.**

are close to or even worse than the baseline of not cutting anywhere.

### *8.2.4. Ablation study*

In Table 2 we study the effect of individual rules or combination of rules introduced in this work. In particular, we evaluate versions or our optimal model MAD*-opt with one or more rules removed and we compare quantitatively to MAD*-opt and our baseline previous work MAD-opt Papanelopoulos and Avrithis (2015).

The rules are studied separately in three groups. In the first group, removing *recovery*, *protection* and their combination causes a significant drop in performance, while in the case of *weak* there is only a slight drop. In the second group, removing *penalization* and *prioritization* has comparable negative effects, which are however not as strong as those of the first group. Finally, in the third group, removing *flatness*, *extension*, *expansion*, or their combinations has the most severe effects in performance compared to the first two groups.

In conclusion, the most important individual rules or salience measures appear to be flatness, expansion strength and recovery, while combinations bring additive effects. The latter means that all rules are complementary.

### *8.2.5. Qualitative evaluation*

Fig. 19 illustrates qualitative results on a number of representative shapes on two datasets. Our method MAD* gives natural results on Kimia dataset and is the only one to capture the ground truth for the bottom part of the rabbit correctly. S&V is harder, but still MAD* yields the highest quality results compared to other methods. Our previous method MAD often tends to prefer cuts near the mouth than on the neck. This is attributed to the shortcut rule which is not always enough.

The selection process of the local convexity rule, introduced in (Papanelopoulos and Avrithis, 2015), is very open to using additional measures. Indeed, we add several other measures in the current work, yielding even better results. For instance, observe in Fig. 19 the blouse and motorcycle from S&V dataset, and the kangaroo and elephant from Kimia dataset. In general, MAD is inferior to the ensemble method CBE, which seeks consensus among all others, while MAD* is superior.

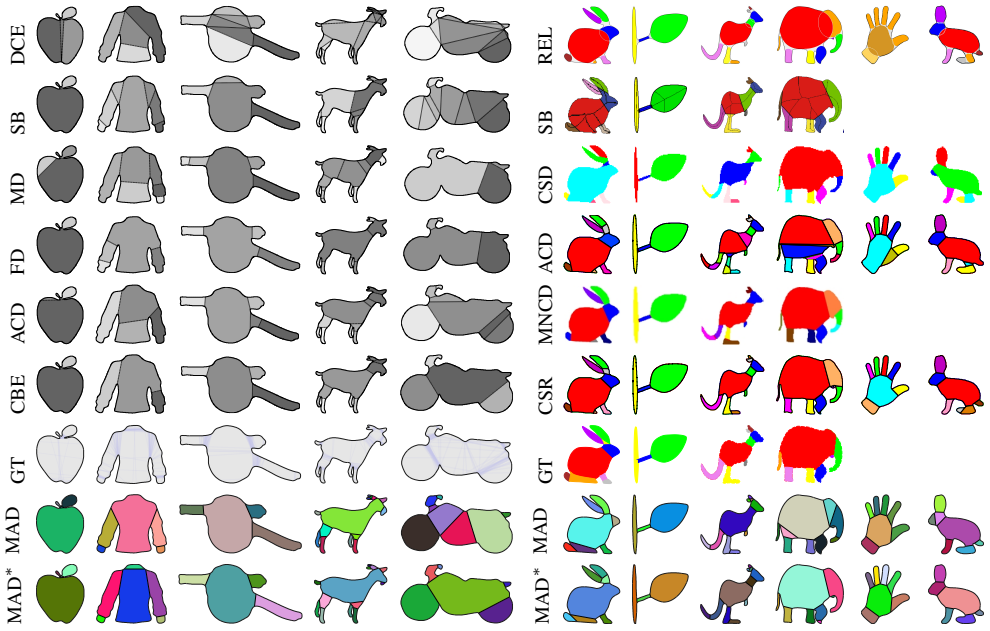



**Fig. 19.** ***Qualitative results*** **on representative shapes of S&V De Winter and Wagemans (2006) (left) and Kimia Siddiqi and Kimia (1995) (right) datasets for a number of methods. Examples of competing methods on S&V are taken from Lewin et al. (2012a). Examples of competing methods on Kimia dataset are taken from the respective publications. Ground truth (GT) is depicted with cuts of all subjects overlaid in blue, 85% transparent, as in Fig. 9c.**

## 9. Discussion

Both qualitative and quantitative evaluation suggests that a simple computational model based on an appropriate representation can outperform all existing models, including ensemble methods. More than that, our model is inherently connected to most rules suggested by human vision studies and highlights their connection. We have first introduced this model in our previous work (Papanelopoulos and Avrithis, 2015), where we have shown that planar shape decomposition based on the medial axis representation can be very simple and effective. Here we show that this model is very flexible in incorporating additional rules always based on the same representation.

In particular, except protrusion strength that we used in (Papanelopoulos and Avrithis, 2015), we also incorporate more salience measures like flatness, expansion strength and extension strength before we apply our local convexity rule. Additionally, we recover cuts that cannot be captured directly from the medial axis, such that all detected cuts are consistent with humans. Contrary to (Papanelopoulos and Avrithis, 2015), our local convexity rule examines corners in a particular order, while the selection of cuts at a corner is not independent with other corners. We show that each additional rule contributes positively to the quality of the decomposition, and their combination even more so—hence they are complimentary.

Other aspects that could be naturally incorporated are detection of bends and continuation of boundaries across parts. The fact that part-cut selection is based on simple local decisions can 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) Perdoch et al. (2007) can be seen as cuts on either figure or ground shape, while distance map saddle points Avrithis and Rapantzikos (2011) correspond to necks Siddiqi and Kimia (1995); our work can provide for a richer set of cuts hence candidate local features.

Like all related work we have studied and compared to, the problem is to decompose “clean” shapes that have not been degraded in any way as would happen with shapes captured from images, *e.g.* by edge detection. Partial occlusion and deformation should not be a problem if the method is robust as defined by Siddiqi and Kimia (1995), that is, decomposition at a point is only affected by its local neighborhood, which largely holds for

**Table 1.** ***Quantitative results*****. Hamming (*H*) and Jaccard (*J*) distance for average and majority evaluation on S&V dataset. Lower is better for both evaluation measures. CBE is an ensemble method on all five methods DCE, SB, MD, FD, ACD. The parameters of our method are tuned based on quantitative and qualitative criteria for MAD*-opt and MAD* respectively; similarly for MAD Papanelopoulos and Avrithis (2015). Human and baseline are computed by us on the same framework by Lewin et al. (2012b). Results for all other methods are reported as provided by Lewin et al. (2012b), where all methods have had their parameters optimized quantitatively on the S&V dataset.**

| | average | | majority | |
|---|---|---|---|---|
| | *H* | *J* | *H* | *J* |
| 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 |
| CBE | **0.111** | 0.288 | 0.069 | 0.186 |
| MAD | 0.126 | 0.317 | 0.096 | 0.247 |
| MAD-opt | 0.118 | 0.303 | 0.085 | 0.225 |
| MAD* | **0.111** | **0.282** | **0.065** | **0.171** |
| MAD*-opt | **0.109** | **0.280** | **0.063** | **0.166** |
| Human | 0.128 | 0.312 | 0.093 | 0.245 |
| Baseline | 0.160 | 0.424 | 0.140 | 0.376 |

our method. Gaps along the boundary are relatively easy to fill according to the Gestalt principle of closure Avrithis and Rapantzikos (2011). However, in the presence of additional structures in the interior of the shape that change its topology, all such methods would fail. This problem is studied by Liu et al. (2014) for instance.

**Table 2.** ***Ablation study*****. Comparison of our baseline previous work MAD-opt Papanelopoulos and Avrithis (2015) and our best result MAD*-opt to versions of the latter with individual rules or combinations of rules removed. Rules are studied in three groups. We remove one or more rules from each group separately, as shown in the leftmost column. Next to each individual rule we refer to the section where the rule is described. We measure Hamming (*H*) and Jaccard (*J*) distance for average and majority evaluation on S&V dataset. Lower is better for both evaluation measures. All parameters are tuned based on quantitative criteria.**

| | average | | majority | |
|---|---|---|---|---|
| | *H* | *J* | *H* | *J* |
| MAD-opt | 0.118 | 0.303 | 0.085 | 0.225 |
| MAD*-opt | 0.109 | 0.280 | 0.063 | 0.166 |
| *MAD*-opt without the following*: | | | | |
| protection (5.1) | 0.112 | 0.287 | 0.069 | 0.180 |
| recovery (5.2) | 0.114 | 0.292 | 0.069 | 0.184 |
| weak (4.2) | 0.110 | 0.283 | 0.063 | 0.168 |
| protection, recovery | 0.115 | 0.295 | 0.077 | 0.199 |
| protection, recovery, weak | 0.116 | 0.296 | 0.077 | 0.200 |
| penalization (7.2) | 0.111 | 0.288 | 0.065 | 0.176 |
| prioritization (7.1) | 0.112 | 0.289 | 0.065 | 0.174 |
| penalization, prioritization | 0.114 | 0.293 | 0.065 | 0.174 |
| flatness (6.2) | 0.118 | 0.303 | 0.077 | 0.203 |
| extension (6.4) | 0.112 | 0.287 | 0.066 | 0.176 |
| expansion (6.3) | 0.116 | 0.296 | 0.071 | 0.188 |
| flatness, extension | 0.119 | 0.304 | 0.078 | 0.204 |
| flatness, extension, expansion | 0.125 | 0.315 | 0.084 | 0.218 |

## References


August, J., Siddiqi, K., Zucker, S.W., 1999. Ligature instabilities in the perceptual organization of shape. CVIU 76, 231–243.

Avrithis, Y., Rapantzikos, K., 2011. The medial feature detector: Stable regions from image boundaries, in: ICCV.

Bai, X., Latecki, L., Liu, W., 2007. Skeleton pruning by contour partitioning with discrete curve evolution. IEEE Transactions on Pattern Analysis and Machine Intelligence 29, 449–462.

Blum, H., Nagel, R.N., 1978. Shape description using weighted symmetric axis features. Pattern Recognition 10, 167–180.

Choi, H., Choi, S., Moon, H., 1997. Mathematical theory of medial axis transform. Pacific Journal of Mathematics 181, 57–88.

De Winter, J., Wagemans, J., 2006. Segmentation of object outlines into parts: A large-scale integrative study. Cognition 99, 275–325.

Dey, T.K., Giesen, J., Goswami, S., 2003. Shape segmentation and matching with flow discretization, in: Algorithms and Data Structures. Springer, pp. 25–36.

Felzenszwalb, P., Huttenlocher, D., 2004. Distance Transforms of Sampled Functions. Technical Report.

Garcia-Garcia, A., Gomez-Donoso, F., Rodríguez, J.G., Orts-Escolano, S., Cazorla, M., López, J.A., 2016. Pointnet: A 3d convolutional neural network for real-time object class recognition, in: International Joint Conference on Neural Networks.

Hoffman, D.D., Richards, W.A., 1984. Parts of recognition. Cognition 18,

65–96.

Hoffman, D.D., Singh, M., 1997. Salience of visual parts. Cognition 63, 29–78.

Kim, D.H., Yun, I.D., Lee, S.U., 2005. A new shape decomposition scheme for graph-based representation. Pattern Recognition 38, 673–689.

Latecki, L.J., Lakamper, R., 1999. Convexity rule for shape decomposition based on discrete contour evolution. Computer Vision and Image Understanding 73, 441–454.

Lewin, S., Jiang, X., Clausing, A., 2012a. A clustering-based ensemble technique for shape decomposition, in: Structural, Syntactic, and Statistical Pattern Recognition. Springer, pp. 153–161.

Lewin, S., Jiang, X., Clausing, A., 2012b. Framework for quantitative performance evaluation of shape decomposition algorithms, in: ICPR.

Lien, J.M., Amato, N.M., 2004. Approximate convex decomposition of polygons, in: Proceedings of Symposium on Computational Geometry, ACM. pp. 17–26.

Liu, G., Xi, Z., Lien, J.M., 2014. Dual-space decomposition of 2d complex shapes, in: CVPR.

Liu, H., Liu, W., Latecki, L.J., 2010. Convex shape decomposition, in: CVPR.

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.Y., Berg, A.C., 2016. SSD: Single shot multibox detector, in: European Conference on Computer Vision.

Luo, L., Shen, C., Liu, X., Zhang, C., 2015. A computational model of the short-cut rule for 2d shape decomposition. IEEE Transactions on Image Processing 24, 273–283.

Ma, C., Dong, Z., Jiang, T., Wang, Y., Gao, W., 2013. A method of perceptual-based shape decomposition, in: ICCV.

Marr, D., 1982. Vision. Freeman, New York.

Mi, X., Decarlo, D., 2007. Separating parts from 2d shapes using relatability, in: ICCV.

Noh, H., Hong, S., Han, B., 2015. Learning deconvolution network for semantic segmentation, in: International Conference on Computer Vision.

Ogniewicz, R., Ilg, M., 1992. Voronoi skeletons: Theory and applications, in: Computer Vision and Pattern Recognition.

Papanelopoulos, N., Avrithis, Y., 2015. Planar shape decomposition made simple, in: British Machine Vision Conference.

Perdoch, M., Matas, J., Obdrzalek, S., 2007. Stable affine frames on isophotes, in: ICCV.

Ren, S., He, K., Girshick, R., Sun, J., 2017. Faster r-cnn: Towards real-time object detection with region proposal networks. IEEE Transactions on Pattern Analysis and Machine Intelligence 39, 1137–1149.

Ren, Z., Yuan, J., Chunyuan, L., Wenyu, L., 2011. Minimum near-convex decomposition for robust shape representation, in: ICCV.

Rosin, P.L., 2000. Shape partitioning by convexity. IEEE Transactions on Systems, Man, and Cybernetics, Part A 30, 202–210.

Siddiqi, K., Kimia, B.B., 1995. Parts of visual form: Computational aspects. PAMI 17, 239–251.

Singh, M., Hoffman, D.D., 2001. Part-based representations of visual shape and implications for visual cognition. Advances in Psychology 130, 401–459.

Singh, M., Seyranian, G.D., Hoffman, D.D., 1999. Parsing silhouettes: the short-cut rule. Perception and Psychophysics .

Snodgrass, J.G., Vanderwart, M., 1980. 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, 174.

Tsogkas, S., Kokkinos, I., Papandreou, G., Vedaldi, A., 2015. Semantic part segmentation with deep learning. CoRR abs/1505.02438.

Yi, L., Su, H., Guo, X., Guibas, L.J., 2017. Syncspeccnn: Synchronized spectral CNN for 3d shape segmentation, in: IEEE Conference on Computer Vision and Pattern Recognition, pp. 6584–6592.

Yu, Q., Yang, Y., Liu, F., Song, Y.Z., Xiang, T., Hospedales, T.M., 2017. Sketch-a-net: A deep neural network that beats humans. International Journal of Computer Vision 122, 411–425.

Zeng, J., Lakaemper, R., Yang, X., Li, X., 2008. 2d shape decomposition based on combined skeleton-boundary features, in: Advances in Visual Computing. Springer, pp. 682–691.