Parts of Visual Form: Computational Aspects

Kaleem Siddiqi, *Student Member, IEEE,* and Benjamin B. Kimia, *Member, IEEE*

Abstract— **Underlying recognition is an organization of objects and their parts into classes and hierarchies. A representation of parts for recognition requires that they be invariant to rigid transformations, robust in the presence of occlusions, stable with changes in viewing geometry, and be arranged in a hierarchy. These constraints are captured in a general framework using notions of a** PART-LINE **and a** PARTITIONING SCHEME**. A proposed general principle of "form from function" motivates a particular partitioning scheme involving two types of parts,** NECK-BASED **and** LIMB-BASED**, whose psychophysical relevance was demonstrated in [39]. Neck-based parts arise from narrowings in shape, or the local minima in distance between two points on the boundary, while limb-based parts arise from a *pair* of negative curvature minima which have "co-circular" tangents. In this paper, we present computational support for the limb-based and neck-based parts by showing that they are invariant, robust, stable and yield a hierarchy of parts. Examples illustrate that the resulting decompositions are robust in the presense of occlusion and clutter for a range of man-made and natural objects, and lead to natural and intuitive parts which can be used for recognition.**

I. Introduction

Underlying recognition is an organization of objects and their parts into classes and hierarchies. In computer vision, the notion of recognition based on "parts" has become increasingly popular. The task of forming high-level object-centered models from low-level image-based features requires intermediate representations and parts serve such a role for robust recognition. Perhaps the most compelling support for this idea is based on recognition in the presence of occlusion: while local features are sensitive to noise and other variations, global structures are susceptible to occlusion by other objects. In contrast, the stable computation of a few parts can lead to recognition that is robust in the presence of occlusions. In addition, objects are often composed of moving or growing parts: while the description of each part remains intact, the relationships between parts change.

How should parts of a shape be computed? Previous approaches have focused either on the decomposition of its interior region or on the segmentation of its boundary. Blum's idea of shape as a collection of overlapping disks leads to partitioning of the interior shape along branch points of the symmetric axis [6, 7]. Other region-based parts include maximally convex parts [38], a description based on the best combination of primitives such as generalized cylinders [5, 28, 29], and superquadrics [32, 3], or the *simplest* description in some language [22, 33]. In contrast, contour-based segmentations use boundary features such as high-curvature points [11] whose salience is demonstrated in [2, 4]. Koenderink [21] noted that three-dimensional shapes are perceived as the composition of elliptic regions with the hyperbolic patches as glue, as evidenced in the works of art, and advocated a decomposition along parabolic lines, or inflection points of the two-dimensional shape. Primitive-based boundary partitioning schemes approximate the boundary as the best combination of primitives such as polygons [34], primitive curvature changes, *e.g.*, cranks, bends, bumps, and ends [1], or constant curvature segments [41].

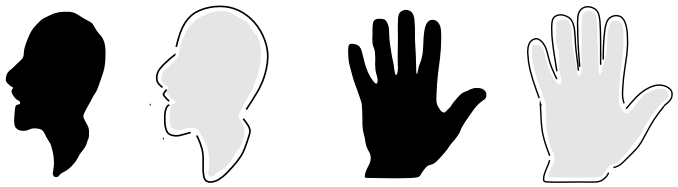

Fig. 1. Hoffman and Richards' theory of curve partitioning segments a contour at negative curvature minima.

A significant departure from the traditional techniques is the boundary-based method proposed by Hoffman and Richards [14] who advocate that *part decomposition should precede part description.* They contrast primitive-based approaches with boundary-based methods and propose a theory of parts which relies not on the shape of parts, as captured by primitives, but rather on general principles underlying their formations, or "regularities of nature". The *transversality principle*, as an example of such a regularity, asserts that "when two arbitrarily shaped surfaces are made to interpenetrate they always meet in a contour of concave discontinuity of their tangent planes". Singularity regularity, or "lawful properties of the singularities of the retinal projection", is a second example. These two regularities lead to a partitioning rule for plane curves: "divide a plane curve into parts at negative minima of curvature." Figure 1 illustrates the successful application of this rule to two examples, a face and a hand, which produces intuitive results. This theory explains several figure-ground reversal illusions successfully; see [9] for further psychophysical evidence. This partitioning proposal leads to a representation of the shape boundary based on *codons*, pieces of the boundary bounded by negative curvature minima; the curvature maxima and zeros are then used to classify each piece as one of the six possible types [36, 35, 37].

Leyton [23, 25, 24, 26] contrasts Hoffman and Richards' notion of parts with one based on a *process* in which a part is not viewed as a rigid segment. Rather, Leyton's

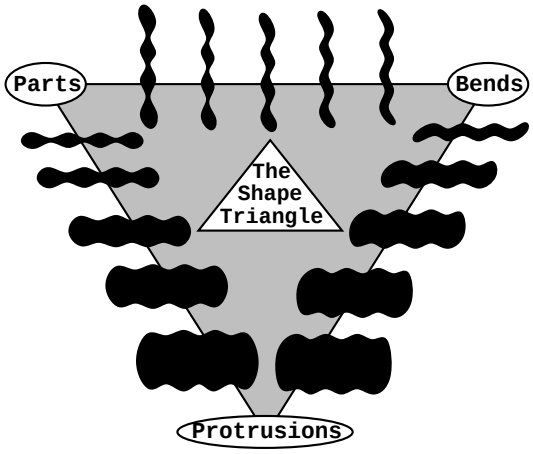

Fig. 2. The *shape triangle* represents three cooperative/competitive processes acting on shape, namely, parts, protrusions, and bends. Each process is biased to see the shape in its own terms. The sides of the triangle represent continuums of shapes whose extremes correspond to the node and whose perception varies from one description to another. For a more detailed description see [19].

process-analysis views a part as a causal explanation, or as a consequence of historical processes, which is based on a biologically-relevant Symmetry-Curvature duality principle [26]. To merge these extreme views, Kimia, Tannenbaum and Zucker [19] propose a continuum connecting two extremes, and capturing the distinction between, (1) the *parts* extreme where objects have clearly defined and distinct segments, *e.g.*, resulting from object composition, and (2) the *protrusions* extreme, where an object is best described as another whose boundary has been deformed, *e.g.*, due to growth. In [19], this continuum is viewed in the context of the *shape triangle* representing three cooperative/competitive processes acting on shape: *parts*, *protrusions*, and *bends*, Figure 2. The three sides of the triangle represent three separate continuums: the parts-protrusion continuum, the parts-bends continuum and the protrusions-bends continuum, capturing the tension between object composition, boundary deformation and region deformation. In this context, the proposal presented here may be viewed as a study focused on a single node of the shape triangle, namely the parts process.

Returning to the theory proposed by [14], the distinction presented between the primitive-based and boundary-based methods can be viewed, alternatively, as a distinction between those primitive-based methods acting on the interior region and primitiveless, general-purpose boundary-based techniques. In other words, the boundary/region and primitive/primitiveless distinctions are orthogonal and not necessarily in conflict [20]. We are in general agreement with Hoffman and Richards on the limited capability of primitive-based techniques, whether they are applied to the boundary or the interior. However, we argue for a role not only for the boundary, but also for the interior of the shape [20]. For example, by stretching a small portion of the contour of a shape, while retaining it's remaining contours, one can produce new shapes whose parts change drastically [20], Figure 3 (top). As another example, the

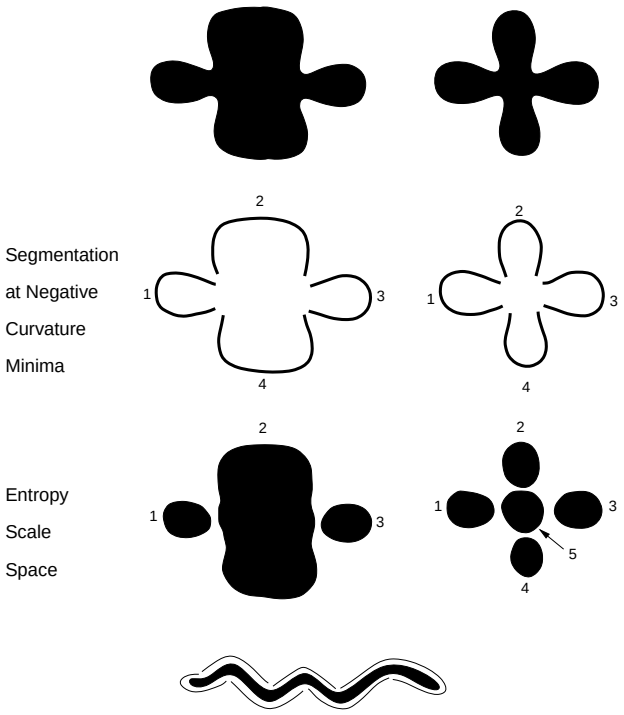

Fig. 3. TOP: Partitioning of a two-dimensional shape requires not only boundary, but also region information [20]. BOTTOM: The interaction of negative curvature minima through the region may produce "bending" rather than parts [20].

interaction of negative curvature minima through the interior region can produce "bending", rather than parts, Figure 3 (bottom). In addition, Kimia, Tannenbaum, and Zucker [20] discover parts which are not derived from curvature extrema of the boundary, but are based on "necks" of the objects, Figure 4.

We now proceed to develop a formal framework to investigate requirements of a partitioning scheme for recognition.

II. A Framework for Partitioning Shape

In this section we develop a framework for studying partitioning schemes for decomposing visual form. Natural constraints from recognition demand that (1) a partitioning scheme be *invariant* to rigid transformations, be *robust* to local deformations and be *stable* with global transformations; (2) a partitioning scheme arrange parts in a hierarchy so that it is suitable for the recognition of objects organized in a taxonomy, for recognition in the presence of resolution effects, and for efficient recognition. We define notions of a *part-line* and a *partitioning scheme* and then formally develop the above constraints in this framework towards a general-purpose shape decomposition scheme.

A significant aspect of a partitioning scheme is the way decomposed parts behave under the various changes that may arise in the visual projection. Earlier approaches to partitioning have demanded that part computations be *reliable*, or invariant with time and viewing geometry [28, 14]. However, observe that three distinct types of changes can occur in the visual world and parts should behave appro-

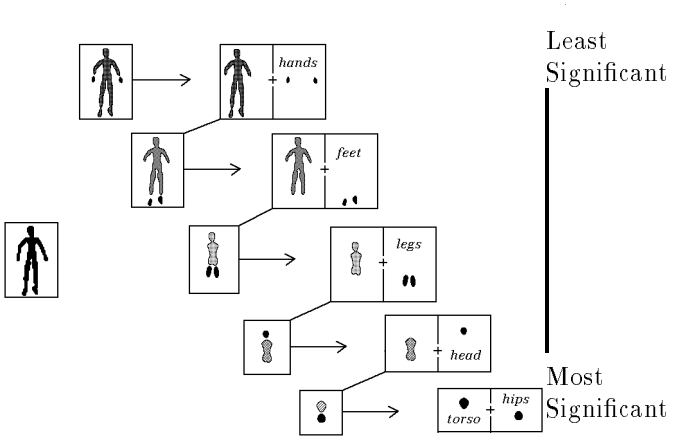

Fig. 4. The hierarchy of parts for the doll obtained from the evolution of shocks in the entropy scale space [20].

priately with these changes. First, we consider changes in the visual image formation that map to a combination of translations, rotations, and scalings of the two dimensional shape. It is clear that parts of the transformed shape must be exactly the transformed parts of the original shape. This constraint of *invariance* leads to computations that are *intrinsic* to the shape. The second class of changes are localized to certain portions of the shape, for example those due to partial occlusion, movement of parts, *etc.* It is natural to demand that such *local* deformations of the shape do not affect those parts that are remote to the change. This constraint of *robustness* leads to computations that are *local* to the part under consideration. Finally, changes in shape can be *global*, *e.g.*, due to changes in viewpoint, changes in viewing direction, growth, *etc.* When such changes to the two-dimensional shape are slight, we expect that the change in the "part-structure" is also slight. This *stability* constraint leads to nonbinary part computations.

A general-purpose decomposition of shape should be based on the interaction between two parts, rather than on their shapes. It is appropriate, then, to explicitly focus on the interface between parts:

Definition 1 *A* PART-LINE *is a curve whose end points rest on the boundary of the shape, which is entirely embedded in it, and which divides it into two connected components.*

This definition is general in that it allows for schemes that favor boundary features as well as those that are based on regional properties. A collection of proposed part-lines that divides a shape into connected parts forms a partitioning scheme:

Definition 2 *A* PARTITIONING SCHEME *is a mapping of a connected region in the image to a finite set of connected regions separated by part-lines.*

We are now in a position to formally develop the notions

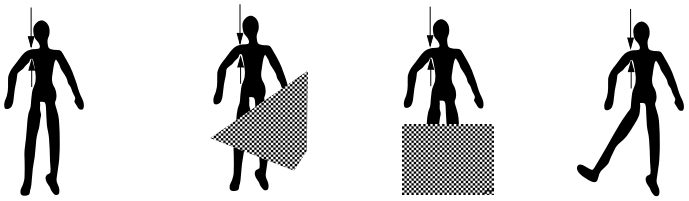

Fig. 5. The part-structure of the original image (left) should not be affected in areas remote to occlusion, movement of parts, *etc.*

of invariance, robustness, and stability:

Definition 3 *A partitioning scheme is* INVARIANT *if the part-lines of a shape that is transformed by a combination of translations, rotations, and scalings, are transformed in exactly the same manner.*

For example, a translated shape's part-lines should translate by the same amount.

Second, we demand that variations in the scene and viewing geometry, in particular the occlusion of an object by another, cause no change to part-structure in regions remote to such change. To illustrate this point further, observe that the part-structure of the visible portion of a doll, *e.g.*, head, hands, torso, *etc.* remains invariant to the occlusion or movement of the feet, as in Figure 5. Let the size of a part-line be the distance between its two end points and define the LOCAL-NEIGHBORHOOD N of a part-line as a disk centered on the part-line, whose radius R is a constant proportion of the length L of the part-line: $R = \alpha L$. Based on qualitative experience gained from psychophysical experiments, we select $\alpha = 2.5$. We should stress, however, that the operation of the algorithm in Section V is not sensitive to the exact choice of α. A part-line is CONTAINED in some region if its local-neighborhood is contained in the region. Then,

Definition 4 *A partitioning scheme is* ROBUST *if for any two shapes, $\mathcal{S}_1$ and $\mathcal{S}_2$, which are exactly the same in some neighborhood, N, the part-lines $\{P_i\}$ contained in N for $\mathcal{S}_1$ and $\mathcal{S}_2$ are exactly equivalent.*

Informally, robust schemes must therefore determine part-lines "locally", namely, based on information that is obtained from a local neighborhood of the part-line. However, since the part-line size may vary over the shape, this neighborhood also varies in size.

The third constraint relates global changes in the shape to the structure of its parts. We require that slight changes in the boundary of the object, which may arise from changes in viewpoint and viewing direction, deformation of the boundary, *etc.* , lead to slight changes in the parts and their relations. Informally, two similar objects must have similar parts. Formally,

Fig. 6. As an object is viewed from larger distances, or when the object is in a different plane of focus, its detailed structures may no longer be distinguishable. The part-structure of the shape should remain stable with variations in viewing distance.

Definition 5 *A partitioning scheme is* STABLE *if slight deformations of the boundary of a shape cause only slight changes in its part-lines.*

Specifically, slight deformations in the shape are measured by metrics for curves, *e.g.*, the Hausdorff metric, while changes in part-lines include changes in their center, size, angle, their disappearance, or their appearance.

Having dealt with the issue of the behavior of parts under various transformations, we now address two further requirements for recognition, namely, the need for stability with changes in resolution, and the need for a hierarchy of parts. First, it is important that part-structure remains stable with changes in resolution, *e.g.*, as an object is seen from varying distances, Figure 6. The parts at one resolution ought to have some correspondence with those at another, specifically, the parts at a coarse resolution must remain at finer resolutions. Second, the need for a hierarchy of parts is well recognized; this allows for a taxonomical representation of objects and efficient recognition via coarse-to-fine matching. What is required here is a notion of **scale** for parts that is compatible with changes in resolution, and which leads to a hierarchical representation. Formally,

Definition 6 *A partitioning scheme is* SCALE-TUNED *if when moving from coarse to fine scale, part-lines are only added, not removed, leading to a hierarchy of parts.*

We now proceed to develop a general purpose partitioning scheme that is consistent with these requirements.

III. A Partitioning Proposal

Objects in the three-dimensional world project to two-dimensional entities, and in the process, much information about their part-structure is lost. Our task is to define a two-dimensional notion of parts based on evidence that remains in this lower-dimensional space that leads to the recovery of three-dimensional parts. The recovery of structure, however, is not only a function of the information available in the retinal image but also of the assumptions one makes about properties of the visual world [12, 17]. Hence, in recovering parts we rely both on properties of objects and on the nature of visual projection, leading to a proposal for two kinds of parts, those that are *limb-based* and those that are *neck-based* [39]. The notions of necks and limbs were developed as a result of an interaction among psychophysical, "ecological" and computational considerations, in the context of a shock-based representation for shape [20]. While a summary of the psychophysical and ecological support for necks and limbs (discussed in detail in [39]) appears below, in this paper our chief objective is to establish computational support for them, as is related to the recognition of objects in realistic situations involving partial occlusion, the movement of parts, *etc.* It is essential to note that parts are computable from shocks; Section VI; the focus of the present work is to establish computational support for our notion of parts.

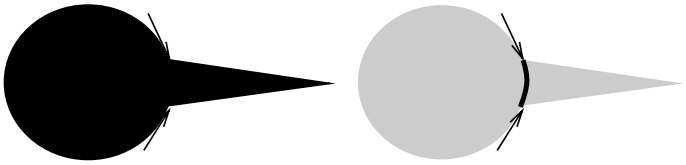

Fig. 7. Limb-based parts are the result of partitioning at a part-line whose end points are negative curvature minima, *and* whose endpoint's tangents continue smoothly from one to the other (co-circular tangents [31]).

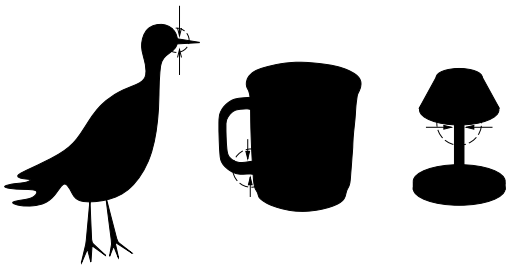

Fig. 8. Examples of limb-based parts for various shapes.

Psychophysical support is provided by a series of experiments in which subjects were required to partition a set of shapes, using part-lines, into what the subjects perceived to be their components. Two interesting results emerged: (1) a high degree of inter-subject and intra-subject consistency, and (2) an overwhelming overlap between perceived and computed parts. The subjects showed a high degree of consistency when the same shapes were presented in separate trials. Similarly, a high degree of consistency was observed when the same shapes were presented to different subjects. Also, there was a close correspondence between the perceived part-lines and computed necks and limbs: a large majority of part-lines perceived were computed as necks or limbs (*e.g.*, 90% for nonsense shapes), and a large majority of computed necks and limbs were perceived as part-lines.

Ecological support for necks and limbs is derived from a response to the question "Why do objects have form at all?" Observe that objects do not exist in isolation, rather, they must interact with their environment and with other objects in it. An object's two-dimensional form is influenced by both the nature of its *interaction* with other objects in a three-dimensional world and the nature of its *projection* onto the retinal image. First, as a result of this interaction, biological and man-made entities, whether by

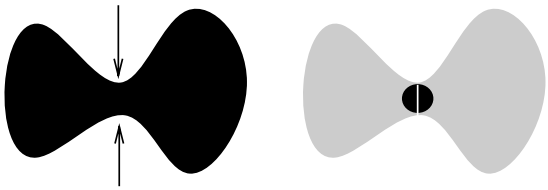

Fig. 9. Neck-based parts are the result of partitioning at the narrowest regions, namely, at part-lines whose lengths are a local minimum, *and* which are the diameter of an inscribed circle. These necks correspond to the second-order shocks of [20].

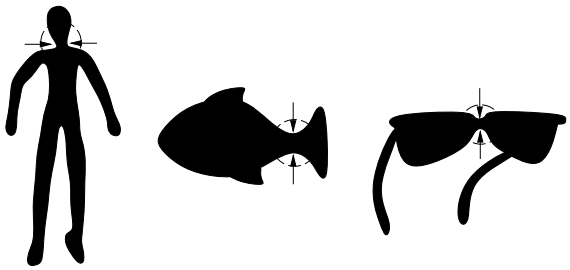

Fig. 10. Examples of neck-based parts for various shapes.

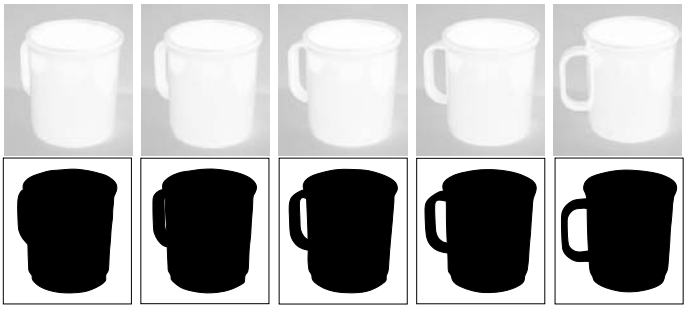

Fig. 11. This figure illustrates a sequence of views of a mug and its corresponding two-dimensional form. Note the change in perception from a protrusion to a part.

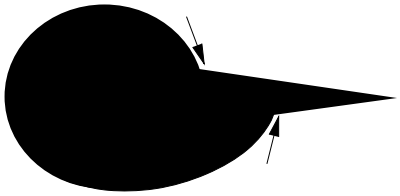

Fig. 12. In the absence of "good continuation", the decomposition is no longer intuitive.

evolution or by design, give rise to parts which specialize in their *function*. When an object specializes function independently in two different, but connected regions, the result is often a sharp change in the three dimensional surface of the object, *e.g.*, the join between the beak of a bird and its head. The projection of these sharp changes yields a pair of high curvature points, where the tangent at one point smoothly continues to the tangent at the other point. This is a restatement of the Gestalt principle of "good continuation" for parts, and provides the first type of parts, the *limb-based* parts, Figure 7. Examples of limb-based parts are shown in Figure 8. In contrast, two portions of an object may specialize function independently, but may be connected through a symbiotic relationship. Requirements such as space for articulation and economy of mass at the connection lead to a narrowing of the join between them, *e.g.*, the tail of a fish; this provides the second type of parts, the *neck-based* parts, Figure 9. Examples of neck-based parts are shown in Figure 10. The above arguments in support of limb-based and neck-based parts together constitute the "**form-from-function**" principle which maintains that form results from an *interaction* among objects and between objects and their environment. Second, this principle is complemented by an assumption about the nature of *projection* of objects which also leads to limb-based parts: a region of an object may partially occlude another, giving rise to a pair of negative curvature discontinuities where the tangent at one discontinuity "smoothly" continues to the tangent at the other discontinuity, Figure 11. Thus projection may also lead to limb-based parts, whose pairs of negative curvature minima are in fact concave discontinuities.

Formally, the strongest form of smooth continuation is co-circularity:

Definition 7 (Limb) *A* LIMB *is a part-line going through a pair of negative curvature minima with co-circular boundary tangents on (at least) one side of the part-line, Figure 7.*

Definition 8 (Neck) *A* NECK *is a part-line which is also a local minimum of the diameter of an inscribed circle, Figure 9.*

Whereas limb-based parts are reminiscent of Hoffman and Richard's parts from transversality, there are two important differences. First, the partitioning is through the region of the shape, via a part-line, and not along its boundary. Second, it is important that the boundary tangents at the pair of negative curvature minima show "good continuation". Observe that when this property does not hold, Figures 3 (bottom) and 12, negative curvature minima may not lead to intuitive parts. However, we are in general agreement with Hoffman and Richards that parts must derive from a regularity of nature. In fact, the "form-from-function" principle complements and generalizes the transversality principle. Limb-based parts are also related to the detection of "co-linear" T-junctions, part of an early heuristic proposal by Guzmán to group regions that belong to the same object [13]. These "co-linear" T-junctions arise in the special domain of line drawings of polyhedral objects, when one object occludes another.

We now establish computational support for limbs and necks by showing that they meet the constraints of invariance, robustness and stability, and lead to a hierarchy of parts for recognition. First, the detection of limbs and necks is based on the computation of curvature extrema, co-circularity, and distance function minima. Curvature is

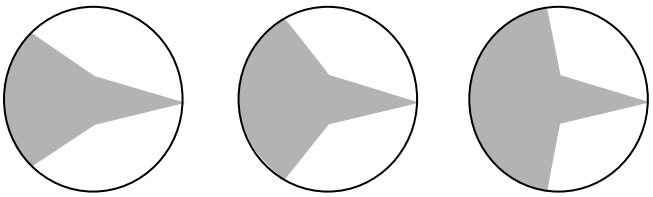

Fig. 13. Salience of a limb increases as total curvature associated with the tangents on the supporting side of the limb decreases.

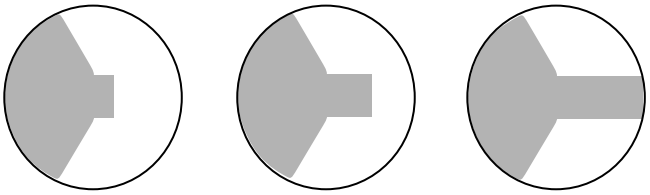

Fig. 14. Salience of a limb increases as extent increases.

an intrinsic quantity and is hence invariant under rotation and translation; its extrema are in addition invariant under scaling transformations. Co-circularity is based on relative orientation and relative position. As such, it is invariant under rotation, translation, and scaling. Euclidean distance is invariant under rotation and translation; distance minima are in addition invariant under scaling transformations.

Remark 1 *The "Limbs-and-Necks" partitioning scheme is invariant.*

Second, observe that the limb-neck computations are based on information that is constrained to a *local neighborhood* N of the part-line. As such, variations in "remote" portions of the shape, namely those outside N, do not affect the resulting part-lines. In contrast, for partitioning schemes with a priori defined primitives, each part-line is the result of an optimization which describes shape as the best fitting combination of primitives. As such, globally-optimized primitive-based parts are a result of a global computation and are not robust in the above sense.

Remark 2 *The "Limbs-and-Necks" partitioning scheme is robust.*

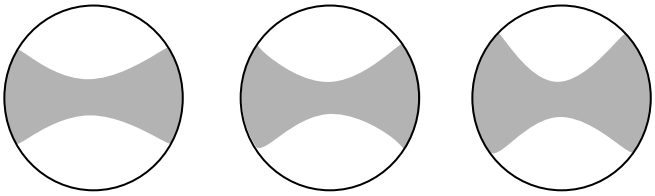

Fig. 15. Salience of a neck increases as curvature disparity increases.

Fig. 16. Parts of equal salience may occur at different scales.

Next, the issue of stability for our scheme is more subtle and requires a further development. Note that so far the notions of necks and limbs are binary in that a part-line is either a neck or not; is either a limb or not. However, two factors motivate the need for non-binary part computations. First, consider the behavior of parts on the continuums of deformations represented by the sides of the shape triangle, Figure 2. Observe how the perception of the shapes depicted on the *part-protrusion* axis changes smoothly from that of a "sausage with parts" to that of a "lasagna noodle with no parts" [19]. A rigid binary inference of parts would lead to instability because the part-based representation of the shape would change abruptly with a very slight change at some point on the axis; a similar argument holds for the *part-bend* axis. Second, recall that our goal is the recovery of three-dimensional parts through intermediate two-dimensional parts. Since evidence for three-dimensional parts is often degraded and may not be conclusive, *e.g.*, the sequence of views of a mug as it is rotated, Figure 11, a part hypothesis must reflect the uncertainty and have an associated measure of strength. As such, a graded partitioning scheme, one which associates a measure of **salience** with each part-line, reflects a slight change in the shape as a slight change in the strengths of its part-lines.

Measures of salience for limbs and necks were psychophysically motivated. For limbs, two factors affect salience: *total curvature* and *extent*. Total curvature is the actual amount a tangent at one negative curvature minima has to bend to align with the tangent at the second negative curvature minima. This measure describes how likely it is for one contour to continue across the limb (via no longer necessarily a circular arc) to the other contour, Figure 13. Extent is a measure of how likely it is that each side is an independent part, as indicated by the mass across the part-line, as computed in a local neighborhood of it, Figure 14. Similarly, the salience of a neck is affected by the *curvature disparity* across the part-line, Figure 15; a measure of how "thin" the neck is. Note that the notion of salience is independent of scale, Figure 16; see Section IV. Now, with the above measures of salience, the limb-based and neck-based part computations become stable: when a shape undergoes a slight deformation its negative curvature minima move slightly, leading to small changes in co-circularity, total curvature, extent and hence limb salience; similarly, neck location and salience are only slightly changed. This allows for a smooth and stable traversal of the sides of the shape triangle:

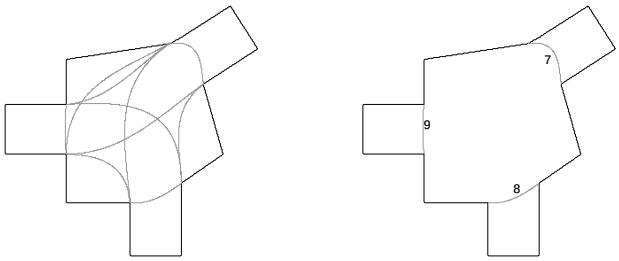

Fig. 17. LEFT: Local computations can lead to global conflicts between part-lines. RIGHT: Conflicts are resolved using salience. Here the selected part-lines occur at about the same scale but differ in salience.

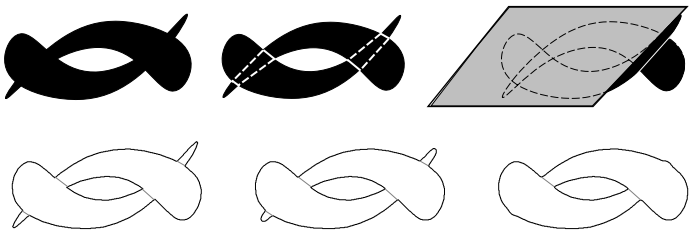

Fig. 19. TOP: This Kanizsa figure is seen as two shapes occluding one another. The solid lines are the selected part-lines, while the dashed-lines indicate a grouping of two parts based on part-lines not selected. Note that the hidden limbs can become selected part-lines if the object is occluded (right). BOTTOM: Part computation from coarse to fine scale (from right to left).

Remark 3 *The "Limbs-and-Necks" partitioning scheme is stable.*

Remark 4 *The "Limbs-and-Necks" partitioning scheme is scale-tuned.*

IV. Scale and Dynamic Partitioning

Next, we address the notion of scale for parts. This is closely related to a notion of smoothing for shape; as shapes are smoothed, small details are removed, including parts, while significant features remain. A smoothing space for shape, the *entropy scale-space*, was introduced in [18], Figure 27. Two processes underlie the approximation of shape in this space: *reaction* rigidly breaks off parts, while *diffusion* melts the boundary into a smoother shape. What is important here is that parts are sorted into a hierarchy of scale using this smoothing process: reaction removes the information locally, while diffusion operates globally [18]. In what follows we show how scale and salience interact, leading to a partitioning scheme that is *scale-tuned*.

Observe that local computations (for robustness to local deformations) can lead to global conflicts, *e.g.*, two part-lines may intersect or may use the same boundary tangents for continuation, Figure 17 (left). Both salience and scale play a role in resolving conflicts. The role for salience is straightforward; when two part-lines are in competition with one another the stronger one is selected, Figure 17 (right). For the scheme to be scale-tuned, parts at coarse scales must remain present at all finer scales. To merge these ideas we propose a **dynamic partitioning** procedure which results in a globally consistent hypothesis, as illustrated by Figure 18. First, at the coarsest scale the strongest part-line is selected and the corresponding part is "broken off." All part-lines in conflict with the chosen part-line are then removed, and the remaining part-lines are considered in order of their salience. This selection process is then repeated for the remaining scales, in order of coarse to fine, until all part-lines are accounted for as either being selected or discarded. As a result, the set of part-lines at any scale is a superset of the set of part-lines at any coarser scale, leading to a hierarchy for parts. Hence,

As a further illustration, consider the Kanizsa Figure 19 (top left), where the thicker portion of each shape is perceived as occluding the thinner portion of the other. Our partitioning scheme finds eight candidate limbs of about equal salience, but which differ in scale, Figure 19 (top middle). The smaller part-lines (solid), however, disappear due to the smoothing process, and at the same time remove evidence for the existence of the larger part-lines (dashed). Note that the discarded part-lines revive if the competing part-lines are removed by partial occlusion, Figure 19 (top right). In addition, while the accepted limbs are the occluding contours, the discarded ones are the occluded contours. A pair of these *hidden limbs* whose endpoints are common with visible limbs signal a single object in occlusion, Figures 19 (top middle), and 24 (bottom middle). Formally,

Definition 9 (Hidden Limb) *A* HIDDEN LIMB *is a discarded limb which shares endpoints with two selected limbs, but uses different boundary tangents for continuation.*

It is intriguing to note that our partitioning scheme is consistent with our visual perception of parts in Figure 20, where local evidence for parts overrides our cognitive knowledge of the visual scene [17].

V. Algorithms

In this section we present algorithms for computing limb-based and neck-based part-lines[1] by outlining the procedure to process a binary shape, *e.g.*, a region of 1's on a background of 0's, to obtain a set of part-lines. We wish to emphasize two points before describing the algorithms. First, the method as described below involves a number of parameters whose values are chosen very liberally, keeping in mind that it is not the selection of arbitrary parameter

[1] The code will be available via ftp from the authors.

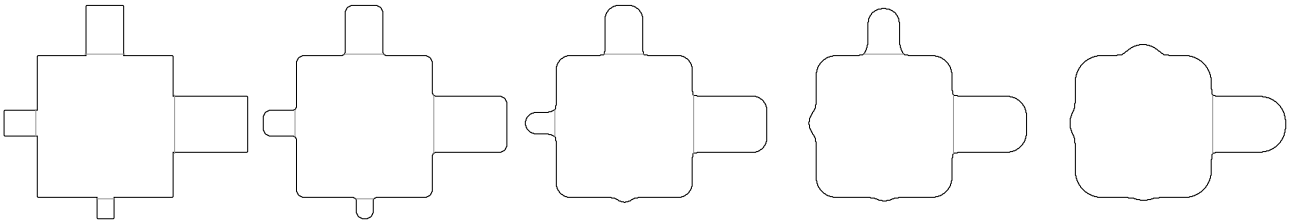

Fig. 18. Part computation from coarse to fine scale (from right to left) leads to a hierarchy of parts. Here the selected part-lines have about equal salience but differ in scale.

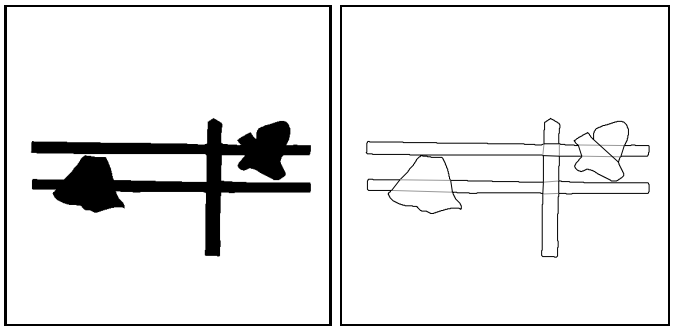

Fig. 20. This figure, taken from [17], is an example of how a local partitioning scheme may be in conflict with more global information. It is clear from the position of the feet and the occlusion of various body parts that the man and the woman are behind the fence. However, our partitioning scheme applied to the areas formed by the the man's jacket and the woman's skirt, which are merged with the fence, yields short vertical part-lines (solid) perpendicular to the fence. The discarded part-lines (dashed) are long horizontal lines parallel to the fence and are considered hidden limbs. Together, they imply that the man's jacket, and the woman's skirt, occlude the fence, and hence must be in front of it!

values that will rule out extraneous part-lines, but the interaction between part-lines. As such, all parameter values were fixed and then used for all the examples, including biological as well as nonsense shapes, Section VII. Second, in our implementation the emphasis has been on the development and illustration of ideas. Hence, many of the steps are not optimal and much can be done to improve their computational efficiency.

ALGORITHM FOR COMPUTING LIMBS

We assume the availability of an ordered list of the shape's

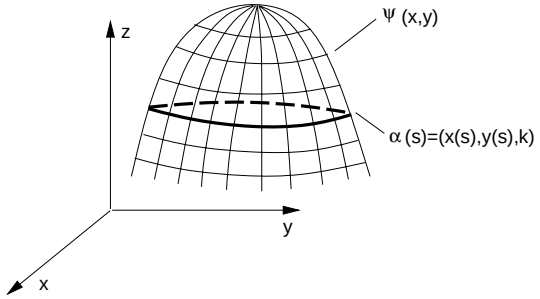

Fig. 21. This figure depicts a surface $z = \Psi(x, y)$ with its level-set $z = k$.

contour points in *contour-list*.

Detecting Negative Curvature Minima: The detection of negative curvature minima is based on computing changes in orientation, where orientation is computed from the distance transform of the binary shape. First, the distance transform of the shape is computed [10, 20]. In this representation, the shape is the zero level set of a surface $z = \Psi(x, y)$, Figure 21. Second, the direction of the tangent in the plane $z = 0$ is given by $\arctan(\frac{-\partial\Psi/\partial x}{\partial\Psi/\partial y})$, which gives the orientation at each point on the shape's contour. This computation is motivated by a larger framework for computing shocks [20]. We find that it gives a robust and accurate estimate of orientation, although other techniques may be used instead. Third, local extrema in curvature are computed as the local extrema in the difference in average orientation between two neighboring points[2]. Finally, for a clockwise trace of the shape's boundary, the extrema where the tangent is turning to the left are identified and stored in *negative-minima-list*. In addition, to rule out "weak" curvature extrema that may be present due to discretization but do not represent significant changes in orientation, the magnitude of the local change in orientation is required to exceed a low threshold (approximately 10 degrees).

Extracting Candidate Limbs: Candidate limbs are extracted by 1) considering each pair of negative curvature minima through which a part-line passes, 2) recomputing

[2] Alternatively one could first compute curvature at each contour point, based on the expression:

$$\kappa(x, y) = -\frac{(\Psi_{xx}\Psi_y^2 - 2\Psi_{xy}\Psi_x\Psi_y + \Psi_{yy}\Psi_x^2)}{(\Psi_x^2 + \Psi_y^2)^{3/2}}, \quad (1)$$

and then select all local extrema of curvature. However, this involves the computation of second derivatives and is less robust than the approach we adopt.

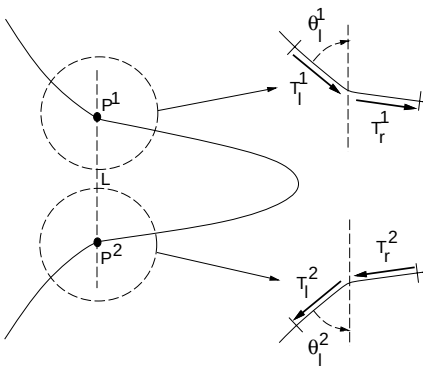

Fig. 22. The left and right neighborhoods at each endpoint of the part-line are proportional in size to the part-line length L and give rise to the left and right boundary tangents $T_l^1, T_l^2, T_r^1, T_r^2$.

the boundary tangents at each endpoint in a local neighborhood of the part-line, 3) computing co-circularity between boundary tangents, and finally 4) selecting those part-lines for which the continuation curve between the co-circular boundary tangents is entirely embedded within the shape. First, two negative curvature minima are paired when a part-line passes through them and divides the contour at each end into two neighborhoods. Second, the "left" and "right" neighborhoods, which are proportional in size to the part-line length L, give rise to the left and right boundary tangents $T_l^1, T_l^2, T_r^1, T_r^2$, Figure 22. Third, with one boundary tangent to be selected from each endpoint, there are four possible pairings of boundary tangents. The co-circularity of each pair is computed [31]. Finally, those part-lines with at least one pair of co-circular boundary tangents, for which the continuation curve lies entirely within the shape are stored in *candidate-limb-list*. The continuation curves used for the examples in this paper are cubic splines. This may be extended to the use of more appropriate subjective contours [40, 8, 27], *e.g.*, the shape of a bent thin beam, minimizing its stored energy [15]. Recently, we have learned of Nitzberg *et al.'s* proposal for the use of "elastica", or curves which minimize a sum of length and the square of curvature, as continuations [30].

Conflict Resolution: Conflicts between candidate limbs are resolved by computing a measure of salience for each. As discussed earlier, salience is a function of total (absolute) curvature and extent. First, the computation of total curvature is as follows. Consider the co-circular boundary tangents T_l^1, T_l^2 which form angles θ_l^1, θ_l^2, with the part-line, respectively, Figure 22. Over the space of curves that smoothly continue T_l^1 to T_l^2, the minimum total absolute curvature is easily seen to be $\theta_l^1 + \theta_l^2$, which has a maximum of π when the co-circular boundary tangents are parallel. Hence, total curvature is normalized as $\frac{\theta_l^1 + \theta_l^2}{\pi}$. Second, extent is computed as the ratio of major to minor axes of the largest ellipse that can be inscribed in the shape, with the part-line as its minor axis, but restricted to lie within the part-line's local neighborhood N. While we have used this crude method as a first order approximation, the results have proven sufficiently robust not to merit further improvement at this stage. Extent attains a maximum when the major axis of the inscribed ellipse coincides with the diameter of the local neighborhood's disk, leading to the normalization of extent as $\frac{extent}{2\alpha}$ (we used $\alpha = 2.5$, with αL as the radius of the disk). The limbs in *candidate-limb-list* are then ordered by salience, which is computed as the average of $(1 - \frac{\theta_l^1 + \theta_l^2}{\pi})$ and $\frac{extent}{2\alpha}$. When a limb is selected, it is placed in *selected-limb-list* and all conflicting limbs are discarded.

Hidden Limbs: When a discarded limb shares endpoints with two selected limbs, but uses different boundary tangents for continuation, it signals occlusion and is recovered as a hidden limb.

ALGORITHM FOR COMPUTING NECKS

We assume the availability of an ordered list of the shape's contour points in *contour-list*.

Extracting Candidate Necks: Candidate necks are extracted by 1) finding all locally shortest part-lines for which a circle with the part-line as its diameter can be inscribed in the shape, and 2) selecting those whose curvature disparity exceeds a low threshold. First, pairs of contour pieces are considered and the locally shortest part-line through each pair is found, if it exists (note that this brute force technique for computing necks is computationally expensive; a reimplementation of our algorithm – as outlined in Section VI – will use much more efficient and elegant second order shock computations). If there are two or more such shortest part-lines for a particular pair of contour pieces, their endpoints are averaged. Among these, part-lines for which a circle can be inscribed in the shape with the part-line as its diameter, are selected. Second, only those locally shortest part-lines across which the curvature disparity exceeds a low threshold are placed in *candidate-neck-list*. This is done because the procedure often results in a significant number of locally shortest part-lines that are present due to discretization, but which do not reflect a significant amount of thinning through the shape's region, Figure 24 (top middle).

Conflict Resolution: Conflicts between candidate necks are resolved by computing a measure of salience for each. As discussed earlier, salience is proportional to the curvature disparity across the neck; for invariance to scaling transformations the measure used is the product of curvature disparity and part-line length L, where curvature is coarsely computed as a function of the deviation of the boundary away from the direction of the tangent. The necks in *candidate-neck-list* are then ordered by salience. When a neck is selected, it is placed in *selected-neck-list* and all conflicting necks are discarded. This concludes the computation of necks.

The computation of limbs and necks begins at the coarsest scales and proceeds to finer scales, as outlined in Section IV. Limbs and necks that are essentially overlapping are grouped into single part-lines. The resulting set of limbs and necks is the algorithm's final output.

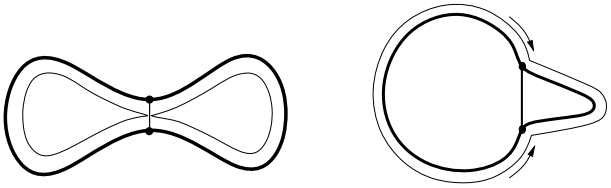

Fig. 23. LEFT: A neck-based part-line is in one-to-one correspondence with a second-order shock formed under inward reaction. RIGHT: Under outward reaction, the endpoints of a limb-based part-line produce a pair of first-order shocks with "co-circular" boundary tangents.

VI. Shock-based Computation of Parts in the Reaction-Diffusion Space

Thus far we have presented computational support for necks and limbs by showing that they are invariant, robust, stable, and scale-tuned. The investigations of parts as necks, however, first began in the reaction-diffusion space of a shape [20], Figure 4. Specifically, the necks give rise to *second-order shocks*: when a shape evolves due to a deformation applied in the normal direction to each boundary point, non-neighboring portions of the boundary may collide, giving rise to two cusps, Figure 23 (left). These discontinuities of curvature are *second-order shocks* which are in one-to-one correspondence with neck-based part-lines. For constant deformation, or reaction, the center of the part-line coincides with the location of the shock and its orientation is perpendicular to the boundary tangent at the shock, Figure 23 (left). The effect of curvature deformation, or diffusion, is to place a measure of salience on the second order shock.

Similarly, in the context of the present study, it is now appropriate to relate limb-based parts to the formation of shocks in the reaction-diffusion space as well. Note that when the boundary of a shape deforms by constant deformation, or reaction, each negative curvature minima produces a single discontinuity in orientation, Figure 23 (right), or a *first-order shock*. Under outward reaction, the endpoints of a limb-based part-line produce a pair of first-order shocks where the boundary tangent of one smoothly follows onto the boundary tangent of the other, Figure 23 (right). To a first approximation, these boundary tangents are co-circular. The detection of simultaneously formed co-circular first-order shocks is an alternate basis for the computation of limb-based parts.

The part computations can therefore be based on the computation of shocks in the reaction-diffusion space. The question naturally arises as to whether the shock-based computations present any advantage over the algorithms presented here. Our goal in departing from the "shape from deformation" framework [20] has been to focus on an *independent* study of parts, guided by constraints from recognition and chiefly motivated by psychophysics. The outcome is a confirmation and a refinement to include limb-based parts. In addition, two numerical advantages of working within a shock-based framework emerge. First, because the computation of shocks is integrated with computations in scale, it is not necessary to separate part computations and approximations of shape in scale. Second, the evolutionary shock computations are numerically more elegant and efficient than the algorithms presented here. It is our intention to return to part computations based on a representation of shape as shocks of the reaction-diffusion space.

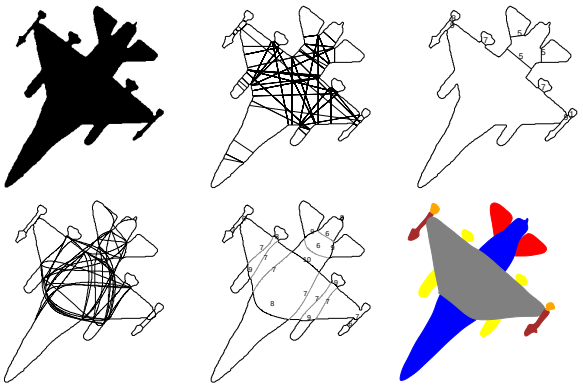

Fig. 24. Necks and Limbs for an F16. TOP, from left to right: the original image; locally shortest part-lines; selected necks with their strengths. BOTTOM, from left to right: all candidate limbs; selected limbs with their strengths, hidden limbs are shown as dashed lines; parts of the F16 depicted in different colors. Observe that the front and rear portions of the occluded rockets are grouped together as single parts.

VII. Examples

In this section we illustrate our partitioning scheme on several examples of man-made and biological objects. We emphasize that *the parameter values used were the same for all examples*. The choices were determined by qualitative experience with a large variety of shapes and the performance of the algorithm is not sensitive to their exact values.

Figure 24 demonstrates the process of computing parts for the shape of an F16 fighter jet. Specifically, note how salience resolves conflicts between local part-line hypotheses, resulting in a natural decomposition of the shape and a recovery of the hidden parts, *e.g.*, the rockets. Figure 25 demonstrates the generality of our proposal for partitioning; the decompositions are natural for a variety of biological and man-made objects. Figure 26 demonstrates the robustness of the process in the presence of occlusion. Portions of occluded and unoccluded versions of the objects with identical shapes have identical parts, as required by the robustness criterion. Figure 27 shows the computation of parts across scale for four pears [35]; note that the similarity in their parts is captured at the coarsest level. Figure 28 illustrates the process of dynamic partitioning for a tank which has been occluded by palm trees. Part computations begin at the coarsest scale, where the scene is described as a smaller blob on top of a larger one, Figure 28 (third row). At a slightly finer scale a third blob is added to the top of the tank, and so on, leading to a hierarchy of parts for the tank. Observe that the parts recovered for the occluded tank correspond closely to those recovered for an unoccluded version of it, Figure 28 (bottom row).

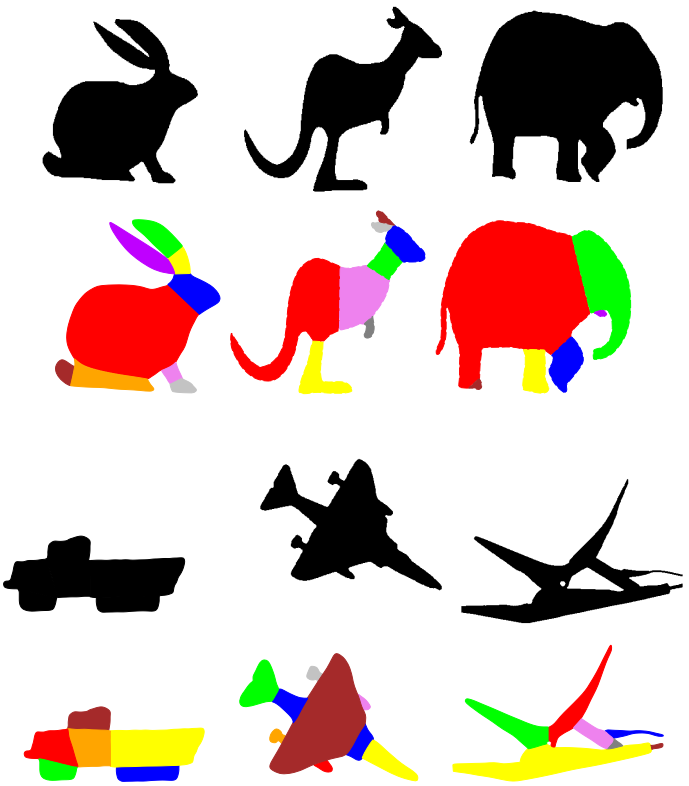

Fig. 25. This figure illustrates the partitioning of a variety of shapes by our algorithm. Note that the results appear intuitive, relating parts to functional portions of the original object, *e.g.*, the tail of the SKYHAWK.

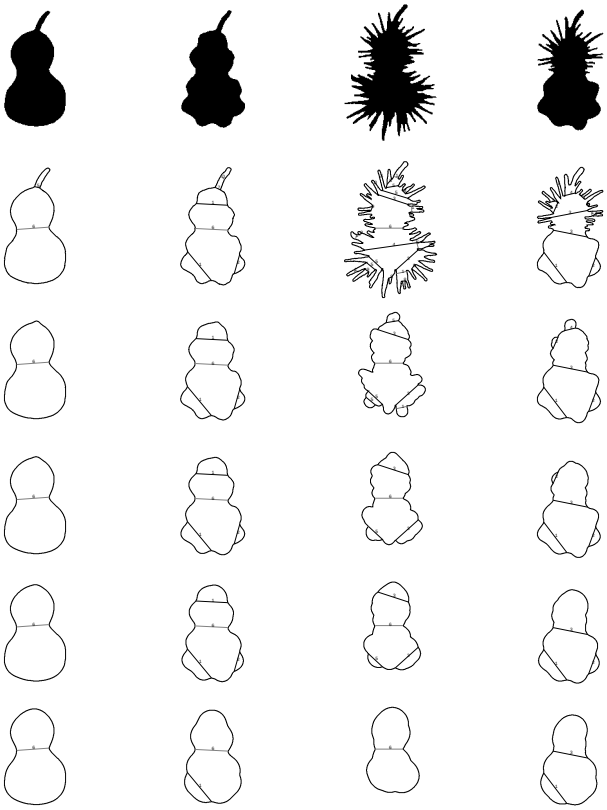

Fig. 27. Bottom to Top: Necks and Limbs computed from coarse to fine scale for four different pears. The top row consists of the original images. Observe that the similarity in their parts is captured at the coarsest scale.

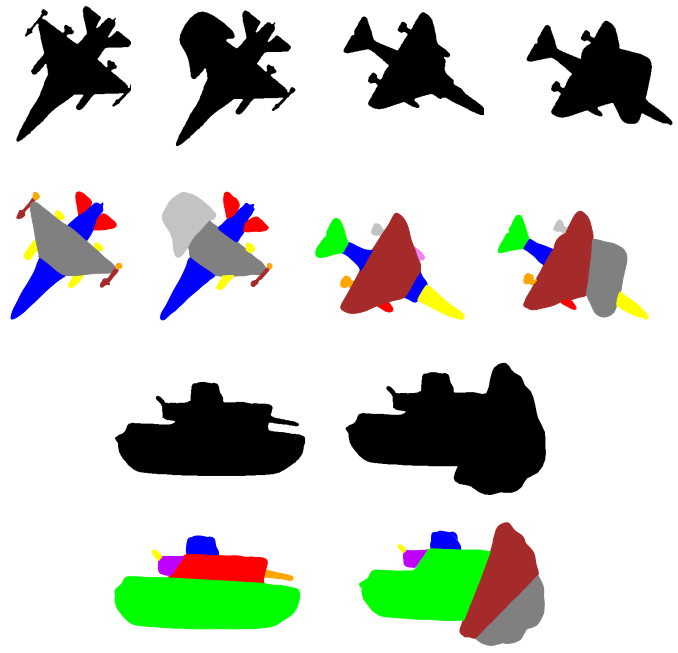

Fig. 26. Examples that illustrate robustness in the presence of occlusion. While partial occlusion affects the occluded parts, it does not interfere with the recovery of the visible ones.

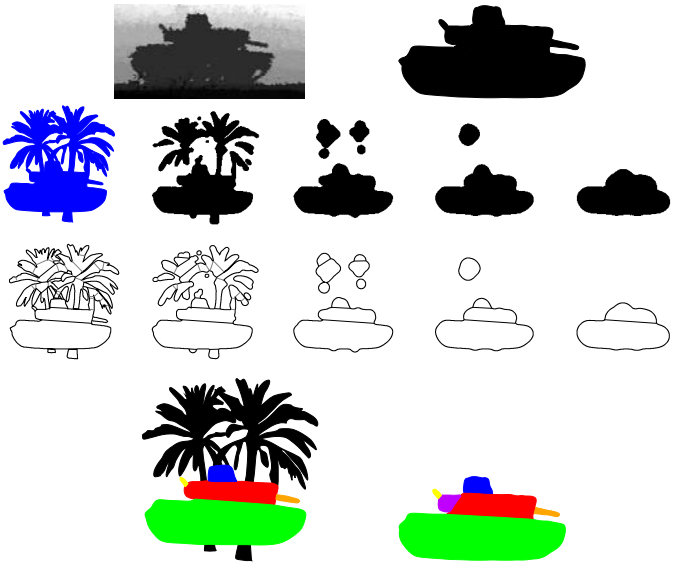

Fig. 28. Top Row: A LADAR range image of a tank and a segmented version of it. Second Row: An image of the tank occluded by palm trees, and samples of its entropy scale space. Third Row: Part computation from coarse to fine scale (from right to left). At the coarsest scale the scene is described as a smaller blob on top of a larger one. At a finer scale a third blob is added to the top of the tank, and so on. Bottom Row: Observe the correspondence between the parts of the occluded tank (left) and those of an unoccluded version of it (right).

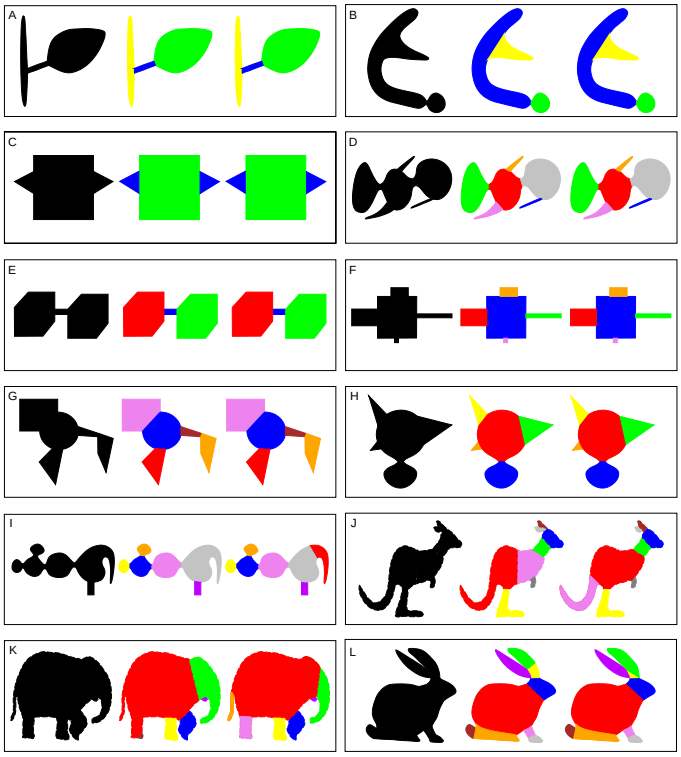

Fig. 29. A comparison of computed parts and perceived parts for a variety of biological and nonsense shapes. The shapes are a representative subset of those used for the psychophysical experiments reported in [39]. Each box depicts the original shape (left), the parts computed by applying our algorithm (middle), and the parts perceived by a majority of the 14 subjects (right). Note that for shapes (A) through (H), the computed and perceived parts are in exact agreement. Shapes (I), (J), and (K) illustrate discrepancies that occur due to the existence of bent limbs, *e.g.*, those manifested as the kangaroo's tail and the elephant's trunk. Shape (L) illustrates the limits of the algorithm's performance when parts of low salience are admitted; here a "weak" neck which breaks off the top part of the rabbit's front ear is computed, but is not perceived.

VIII. Discussion

The validity of our partitioning scheme can be measured against the principles it sought to satisfy, as well as against human performance. Whereas we have previously discussed the former, Figure 29 illustrates the latter. Despite a high degree of correspondence between computed parts and perceived parts, we have observed two minor classes of discrepancies between them: 1) those due to part-bend interactions, and 2) those due to cognitive knowledge of the underlying object. First, consider the part-bend axis of Figure 2. Observe how the perception of the leftmost shape as a "sausage" with four parts changes continuously to one of a snake with a single bent part. Now consider Figure 30; whereas perceptual evidence for "trunks" and "tails" as parts is strong for the shapes on the right of each box, leading to clearly partitioned limbs, the evidence is greatly diminished for the shapes on the left, due to bending. Our psychophysical experiments indicate that in such situations, whereas subjects continue to place one endpoint of a part-line at the negative curvature minimum, the position of the second endpoint is somewhat arbitrary [39]. Such parts can only be recovered under a more comprehensive framework [19]. Second, cognitive knowledge influences part perception; *e.g.*, familiarity with the underlying object for a recognizable shape, and the existence of a semantic vocabulary for describing its various components may cause a subject to break off parts, even when perceptual evidence is weak [39].

IX. Conclusion

In conclusion, we comment on the relationship between partitioning and recognition. Thus far, we have assumed the availability of a 2D shape, that which comes from the projection of the occluding contour of an object. However, it is well recognized that under general conditions the segmentation of an image into regions corresponding to the projections of distinct objects is not an easy task. In the face of this difficulty, how can partitioning proceed? It is clear that since segmentation is a difficult task, a partitioning scheme should be able to handle errors in the segmentation process, partially correct segmentations, *etc.* To this end the limbs-and-necks scheme, being robust to local deformations and stable with slight global deformations, is appropriately designed. A more complete answer, however, lies in viewing parts as an intermediate representation that allows for the flow of bottom-up as well as top-down information. Consider that since part computations are local, edges of the appropriate polarity can interact to form necks and limbs *prior* to obtaining a segmentation of the object, Figures 13, 14 and 15, leading to a "*parts receptive field*". This constitutes the bottom-up flow of information, *i.e.*, from local edge hypotheses to the more global part hypotheses. Now, part hypotheses can in turn play an integral role in the segmentation process through the top-down flow of information, *i.e.*, a combination of likely parts can lead to an object hypothesis, followed by a segmentation hypothesis for the image. Such a notion of parts may be key in resolving the bottom-up/top-down bottleneck of recognition.

Acknowledgement

The authors thank David Cooper, Allan Dobbins, James Elder, Danny Keren, Frédéric Leymarie, Francene Reichel, Jayashree Subrahmonia, Bill Warren and Steve Zucker for discussions and comments. We are grateful to Fridtjof Stein of the University of Southern California for providing us with the animal shapes. The support of NSF grant IRI-9305630 and a Sun Microsystems Academic Grant is gratefully acknowledged.

References

[1] H. Asada and M. Brady. The curvature primal sketch. *IEEE PAMI*, 8:2–14, 1983.
[2] F. Attneave. Some informational aspects of visual perception. *Psych. Review*, 61:183–193, 1954.
[3] R. Bajcsy and F. Solina. Three-dimensional object representation revisited. In ICCV1987 [16], pages 231–240.
[4] I. Biederman. Recognition by components. *Psych. Review*, 94:115–147, 1987.
[5] T. O. Binford. Visual perception by computer. In *IEEE Conference on Systems and Control*, December 1971.
[6] H. Blum. Biological shape and visual science. *J. Theor. Biol.*, 38:205–287, 1973.

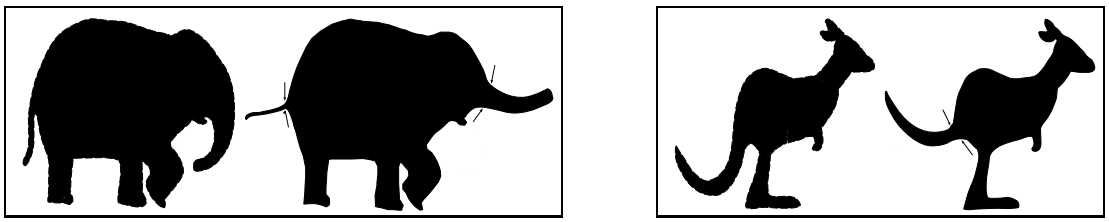

Fig. 30. A trunk or tail is a bent limb! Whereas perceptual evidence for "trunks" and "tails" as parts is strong for the shapes on the right of each box, leading to clearly partitioned limbs, the evidence is greatly diminished for the shapes on the left, due to bending.

[7] H. Blum and R. N. Nagel. Shape description using weighted symmetric axis features. *Pattern Recognition*, 10:167–180, 1978.

[8] M. Brady, W. E. L. Grimson, and D. J. Langridge. Shape encoding and subjective contours. *Proc. National Conf. on Artificial Intelligence*, pages 15–17, 1980.

[9] M. L. Braunstein, D. D. Hoffman, and A. Saidpour. Parts of visual objects: an experimental test of the minima rule. *Perception*, 18:817–826, 1989.

[10] P. Danielsson. Euclidean distance mapping. *Computer Graphics and Image Processing*, 14:227–248, 1980.

[11] H. Freeman. Shape description via the use of critical points. *Pattern Recognition*, 10:159–166, 1978.

[12] J. Gibson. *The Perception of the Visual World*. Riverside Press, 1950.

[13] A. Guzmán. Decomposition of a visual scene into three-dimensional bodies. In *Automatic Interpretation and Classification of Images*. Academic Press, 1969.

[14] D. D. Hoffman and W. A. Richards. Parts of recognition. *Cognition*, 18:65–96, 1985.

[15] B. K. P. Horn. The curve of least energy. *ACM Trans. Mathematical Software*, 9(4):442–460, December 1983.

[16] *First International Conference on Computer Vision, (London, England, June 8–11, 1987)*, Washington, DC., 1987. IEEE Computer Society Press.

[17] G. Kanizsa. *Organization in Vision: essays on Gestalt perception*. Praeger, 1979.

[18] B. B. Kimia, A. R. Tannenbaum, and S. W. Zucker. *Entropy Scale-Space*, pages 333–344. Plenum Press, New York, May 1991.

[19] —. The shape triangle: Parts, protrusions, and bends. Technical Report TR-92-15, McGill University Research Center for Intelligent Machines, 1992.

[20] —. Shapes, shocks, and deformations, I: The components of shape and the reaction-diffusion space. *International Journal of Computer Vision*, To Appear, 1994.

[21] J. J. Koenderink and A. J. van Doorn. The shape of smooth objects and the way contours end. *Perception*, 11:129–137, 1982.

[22] Y. G. Leclerc. Constructing simple stable descriptions for image partitioning. *International Journal of Computer Vision*, 3:73–102, 1989.

[23] M. Leyton. Symmetry-curvature duality. *CVGIP*, 38:327–341, 1987.

[24] —. A process grammar for shape. *Artificial Intelligence*, 34:213–247, 1988.

[25] —. Inferring causal history from shape. *Cognitive Science*, 13:357–387, 1989.

[26] —. *Symmetry, Causality, Mind*. MIT press, April 1992.

[27] A. Linnér. Steepest descent as a tool to find critical points of $\int \kappa^2$ defined on curves in the plane with arbitrary types of boundary conditions. In *Geometric Analysis and Computer Graphics*, pages 127–138, New York, 1988. Springer-Verlag.

[28] D. Marr and K. H. Nishihara. Representation and recognition of the spatial organization of three dimensional structure. *Proceedings of the Royal Society of London*, B 200:269–294, 1978.

[29] R. Nevatia and T. O. Binford. Description and recognition of curved objects. *Artificial Intelligence*, 8:77–98, 1977.

[30] M. Nitzberg, D. Mumford, and T. Shiota. *Filtering, Segmentation and Depth*. Springer-Verlag, 1993.

[31] P. Parent and S. W. Zucker. Trace inference, curvature consistency and curve detection. *IEEE PAMI*, 11(8):823–839, August 1989.

[32] A. Pentland. Recognition by parts. In ICCV1987 [16].

[33] —. Part segmentation for object recognition. *Neural Computation*, 1:82–91, 1989.

[34] U. Ramer. An iterative procedure for the polygonal approximation of plane curves. *CVGIP*, 1(3):244–256, 1972.

[35] W. Richards, B. Dawson, and D. Whittington. Encoding contour shape by curvature extrema. *Journal of Optical Society of America*, 3(9):1483–1489, 1986.

[36] W. Richards and D. D. Hoffman. Codon constraints on closed 2d shapes. *CVGIP*, 31(2):156–177, 1985.

[37] W. Richards, J. Koenderink, and D. Hoffman. Inferring three dimensional shapes from two-dimensional silhouettes. *Journal of Optical Society of America*, 4(7):1168–1175, 1987.

[38] L. Shapiro and R. Haralick. Decomposition of two-dimensional shapes by graph-theoretic clustering. *IEEE PAMI*, 1:10–20, 1979.

[39] K. Siddiqi, K. J. Tresness, and B. B. Kimia. Parts of visual form: Ecological and psychophysical aspects. Technical Report LEMS 104, LEMS, Brown University, June 1992.

[40] S. Ullman. The shape of subjective contours and a model for their generation. *Biological Cybernetics*, 25:1–6, 1976.

[41] D. M. Wuescher and K. L. Boyer. Robust contour decomposition using a constant curvature criterion. *IEEE PAMI*, 13(1):41–51, 1991.