Fragment Grouping via the Principle of Perceptual Occlusion

Jonas August Kaleem Siddiqi Steven W. Zucker
Centre for Intelligent Machines, McGill University, 3480 University Street, Montréal, Canada H3A 2A7
E-mail: august@cim.mcgill.ca

Abstract

Bounding contours of physical objects are often fragmented by other occluding objects. Long-distance perceptual grouping seeks to join fragments belonging to the same object. Approaches to grouping based on invariants assume objects are in restricted classes, while those based on minimal energy continuations assume a shape for the missing contours and require this shape to drive the grouping process. We propose the more general principle that those fragments should be grouped whose fragmentation could have arisen from a generic *occluder. The gap skeleton is introduced as a representation of this virtual occluder, and an algorithm for computing it is given.*

1 Introduction

Mother Nature can appear vindictive, or at least obstructionist. Such complaints are most frequent at the intermediate levels of computer vision, where, for example, object contours are interrupted by occlusion (Figure 1). If only objects were not obscured, we would not have to invent procedures for linking across the missing boundaries. We claim that this property of Nature is not obstructionist, but rather provides the key to long-distance perceptual grouping. The insight motivating this claim is that long-distance perceptual grouping is a kind of inverse to physical occlusion:

Principle of Perceptual Occlusion (i) Forward: *Suppose we are given a collection of (opaque) objects distributed in space. Occlusion will arise generically under projection (either perspective or orthogonal) onto images. The result is a composite bounding contour which, by transversality, will contain discontinuities at the occlusion points. The result is that only* fragments *of the bounding contours of occluded objects will be present in the image.*
(ii) Inverse: *Suppose we are given a collection of contour fragments. Long-distance perceptual grouping is the association of fragments that could belong together in the following sense: if there were a* generic *occluder, then the given fragment arrangement would ensue.*

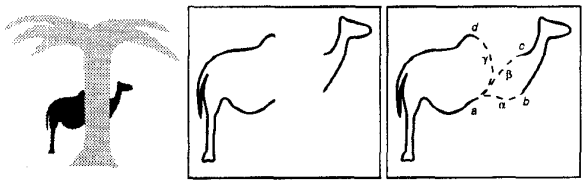

Figure 1: (LEFT) The bounding contour of a camel is broken by a foreground palm tree. (CENTER) Curve fragments remaining after depth separation using T-junctions. (RIGHT) Standard approaches do not link the fragments correctly. "Optimal" continuations α, β, and γ, joining end point a to ends b, c, and d, respectively, all have similar energies based on length and total squared curvature.

A classical consequence of the forward principle (i) is the formation of "T"-junctions. In the "blocks world", the existence of aligned T-junctions is sufficient to perform the inverse inference, because all projected sides are straight lines [8]. Connecting T-junctions with straight lines has emphasized a boundary-completion view of grouping, and modern approaches generalize polyhedral objects to those with sides obeying minimal energy or convexity principles [16, 20, 11]. Natural objects may not conform to such principles, and counter-examples abound (Figure 1(RIGHT)). Furthermore, whether these "optimal" continuations are necessary for recognition is unclear. The other currently popular approach is to use algebraic or differential invariants, but these only apply to extremely limited classes of shapes, such as surfaces of revolution [21].

Our approach amounts to implicitly hypothesizing a virtual occluder, with both an outline and an interior region. We wish the process to be generic, so we do not make assumptions limited to particular shapes or particular energy forms. Rather, our process is based on a classical aspect of shape description—Blum's skeleton [3]—and on modern methods for calculating it. Our contribution is to unify a subset of the Blum skeleton with end points via the above occlusion principle. This defines the gap skeleton, which in turn provides a description of an implicit occluder, and explicitly indicates which fragments to group.

1015-4651/96 $5.00 © 1996 IEEE
Proceedings of ICPR '96

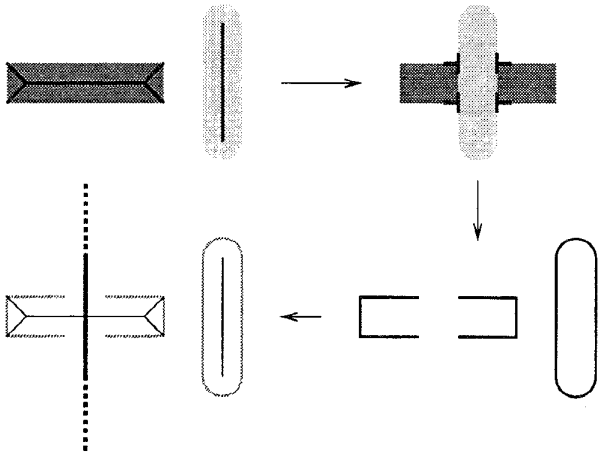

Figure 2: Grouping using the gap skeleton. (TOP LEFT) Two objects with their skeletons. (TOP RIGHT) The two objects in an occlusion relationship; observe the formation of T-junctions. (BOTTOM RIGHT) The bounding contours are separated by a partial ordering in depth induced by T-junctions: the fragments of the rectangle are "farther" than the bounding contour of the occluder. (BOTTOM LEFT) The skeletons of the fragments are shown as thin lines, while the gap skeleton is the thick vertical line. Note the similarity between the gap skeleton and the skeleton of the occluder; consider the gap skeleton as representing a "virtual occluder".

2 Overview of approach

The minimal universe in which contour fragmentation arises from occlusion is that of contrasting paper cutouts arranged with a partial ordering in depth. This 2.1-D world [16] gives rise to T-junctions when objects occlude, and we begin with a decomposition at T-junctions as shown for the pair of objects in Figure 2. Thus fragments explicitly arise, and our main goal is to recover a linking of such fragment endpoints. We seek the groups that most naturally account for the data in the sense of the above principle of perceptual occlusion. As this simple example illustrates, the linking is dictated by the gap skeleton (shown in bold, bottom left, Figure 2). In particular, two curve endpoints are grouped only if they have a gap skeleton. This also provides a rough (skeletal) approximation to an object that, if it were present, would have broken the rectangle into the fragments shown. The main content of the paper is the introduction of the gap skeleton and an algorithm for computing it.

3 Skeleton and gap skeleton

Recall the definition of the skeleton [3] of a closed set $A \subset \mathbf{R}^2$ in terms of MAXIMAL OPEN DISCS, where an open disc D is maximal if and only if there exists no other open disc contained in the complement of A that properly contains D:

Definition 1 (Blum) *The* SKELETON S *of* A *is the set of centers of maximal open discs contained in the complement of* A.

An equivalent definition of the skeleton introduces the notion of the projection of a point to the nearest point(s) in A. The distance $\rho(x)$ from the point $x \in \mathbf{R}^2$ to the set A is the minimum distance from x to any point of A, or $\rho(x) \triangleq \min\{\|x - p\| : p \in A\}$. The PROJECTION $\pi(x)$ is the set of points in A closest to x, or $\pi(x) \triangleq \{p \in A : \|x - p\| = \rho(x)\}$. The ray beginning at point p and passing through q and the line segment joining points p and q will be denoted $\overrightarrow{pq}$ and $\overline{pq}$, respectively. Using these terms, Calabi and Hartnett [5] then defined the skeleton S of A as the set of points $q \notin A$ satisfying:

$$\rho(q) = \|q - x\| + \rho(x) \qquad \forall x \in \overline{pq},$$
$$\rho(x') < \|x' - q\| + \rho(q) \qquad \forall x' \in \overrightarrow{pq} \setminus \overline{pq},$$

where $p \in \pi(q)$.

The RADIUS FUNCTION r is the restriction $r \triangleq \rho|_S$ of ρ to the skeleton S and so $r(q)$ is the the radius of the maximal disc at q. The projection $\pi(q)$ to A is the set of points at which the maximal disc at q "touches" A. In fact, the intersection of the closure of the maximal disc at q and A is precisely $\pi(q)$. Except at a finite number of points of S, the maximal disc will touch A at two points.

To formalize the notion of the set of curve fragments to be grouped, consider A as the disjoint union of traces of a finite number of simple, piecewise smooth curves, either open or closed. We introduce the gap skeleton through the more general notion of pregap skeleton.

Definition 2 *The* PREGAP SKELETON *with respect to endpoints* a *and* b *is the set* $\mathrm{pre}(a, b) \triangleq \{q \in S : \pi(q) = \{a, b\}\}$.

Note that $\mathrm{pre}(a, b)$ may be empty (see Figure 3) and $\mathrm{pre}(b, a) = \mathrm{pre}(a, b)$.

Proposition 1 $\mathrm{pre}(a, b)$ *is contained in the perpendicular bisector of the line segment joining* a *and* b.

The proof is by geometric construction. Since the circle of radius $r(q)$ at any $q \in \mathrm{pre}(a, b)$ intersects A at a and b, q is the apex of an isosceles triangle whose base is the line segment joining a and b. The result follows.

For proofs of the following, see [2].

Proposition 2 $\mathrm{pre}(a, b)$ *is connected.*

Let θ_a be the signed angle *from* the tangent extending out of endpoint a *to* the line segment $\overline{ab}$. Define θ_b similarly.

Proposition 3 *If* $\mathrm{pre}(a, b)$ *is nonempty and* $|\theta_a + \theta_b|$ *is not equal to* $180°$, *then* $\mathrm{pre}(a, b)$ *has nonzero length.*

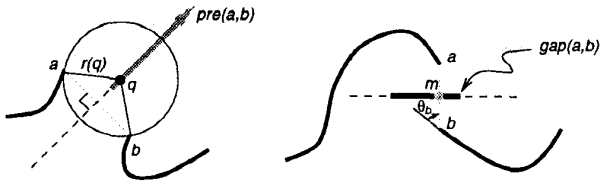

Figure 3: (LEFT) The definition of pregap illustrated: only endpoints a and b are a distance $r(q)$ away from the skeletal point q, so $\pi(q) = \{a, b\}$. All such skeletal points form $\text{pre}(a, b)$, and extend along a ray on the perpendicular bisector of the endpoints. Note here that the gap skeleton with respect to a and b is empty because the midpoint is not in $\text{pre}(a, b)$. (RIGHT) Here there is nonempty gap skeleton with respect to a and b, and the radius function r obtains a local minimum at m, the midpoint between a and b.

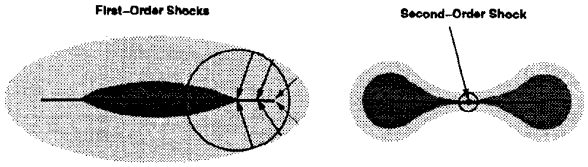

Figure 4: Some shock types which occur during curve evolution with β = constant. At two shock points the related maximal disc is shown. (LEFT) A FIRST-ORDER SHOCK is a discontinuity in orientation of the boundary of a shape. (RIGHT) A SECOND-ORDER SHOCK is formed when two distinct non-neighbouring boundary points join, but none of their immediate neighbours collapse together.

Coupled with these properties, we see that "almost always" a nonempty pregap skeleton is a "chunk" of the perpendicular bisector between its related endpoints.

Definition 3 *The* GAP SKELETON *with respect to a and b is the pregap skeleton with respect to a and b provided that r achieves a local minimum in* $\text{pre}(a, b)$*:*

$$\text{gap}(a, b) \triangleq \begin{cases} \text{pre}(a, b), & \text{if } r \text{ has a local min. in } \text{pre}(a, b); \\ \varnothing, & \text{otherwise.} \end{cases}$$

Conjecture 1 *The midpoint between a and b lies in* $\text{pre}(a, b)$ *if and only if r has a local minimum in* $\text{pre}(a, b)$.

Intuitively, gap skeleton arises when maximal discs touch a given pair of endpoints and become smallest "between" those endpoints. Thus there will be nonempty gap skeleton with respect to a pair of endpoints when no curve enters into the disc whose diameter is the line segment joining the endpoints. Note that, for a given pair of endpoints, nonempty pregap skeleton does not imply nonempty gap skeleton, as shown in Figure 3. In addition, if the pregap skeleton is empty with respect to all pairs of endpoints of a given set of curves, then so is the related gap skeleton.

For the task of grouping, we use the gap skeleton with respect to two endpoints as a cue for linking. Interestingly, others have used the related notion of Voronoi tesselation for grouping [1, 9], but in these approaches no special status is given to endpoints. Endpoints and the discontinuities in boundary orientation that give rise to them are critical for unit formation in human psychophysics [12].

4 Skeletons and shocks

In Blum's *grassfire* formulation, the skeleton is obtained as quench points of a fire front moving parallel to the shape's boundary. Here there is an explicit ordering on the locus of skeletal points in time: they appear in the direction of grassfire flow, in the direction of increasing object width. An evolutionary approach to shape description supports and complements this view by introducing a partial differential equation [4, 13]: $\frac{\partial \mathcal{C}}{\partial t} = \beta \vec{N}$, with the initial condition $\mathcal{C}(s, 0) = \mathcal{C}_0(s)$, where $\mathcal{C}(\cdot, t)$ is the closed curve describing the boundary of the shape after a deformation by time t, $\vec{N}$ is the outward normal, s is the arc-length parameter, and β is an arbitrary function of the local geometry of $\mathcal{C}(\cdot, t)$. The special case where β is a constant is precisely the Blum grassfire.

Curves evolving according to the above equation develop *shocks* [15], or entropy-satisfying singularities, which are points where some information is lost during deformation. The shocks are classified into four types [13], two of which directly relate to the gap skeleton (see Figure 4). Numerical algorithms have recently been developed for detecting, classifying, and grouping shocks into higher level data structures [19]. We use this algorithm to detect gap skeletons because: (i) the approach provides shock locations with sub-pixel resolution; in principle the speed of the propagating front can be made arbitrarily small, (ii) the algorithm provides accurate estimates of the radius function as well as shock velocities, both of which are are critical for the computation of the projection $\pi(q)$ of a shock point q, and (iii) the detected shocks are *already* grouped into shock branches, along the direction of shock flow.

As we show, first- and second-order shocks are extremely important. First, however, we note that the curve evolution approach is designed for a collection of *closed* curves in the plane having a topologically well-defined inside (the computation being carried out on an associated embedding surface [17]). Open curves, however, are slightly more subtle since they arise from edges for which the orientation is known modulo 180 degrees: the polarity of the orientation is undefined (see Figures 1 and 2). Thus both polarities must be considered, which we effect by dilating the curve an ϵ-

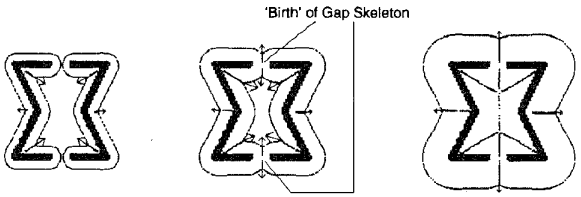

Figure 5: (LEFT) The bold figure is the dilation of the broken bounding contour of an hourglass. A front during curve evolution is overlaid, with first-order shock points forming in high curvature regions. Arrows denote the direction and speed of first-order shocks. End points propagate as semi-circular arcs. (CENTER) Gap skeleton is "born" at both at the top and bottom: two first-order shock branches emerge from a collision of fronts from endpoints. (RIGHT) The front continues to propagate outwards, with first-order shocks tracing out the skeleton.

amount. The boundary of the dilation is a closed curve with interior, and evolution can now proceed as usual (see Figure 5).[1] These ideas can be extended to objects with holes by observing how the optical projection of the tangent to the surface defines an inside and outside to a closed curve: this generalized "inside" may be the space outside the closed curve.

We are now in a position to see the relationship between the shocks formed during curve evolution and the definition of gap skeleton. The locus of points through which the shocks migrate corresponds to a Blum skeleton. In Figure 5 [7], note how the endpoints propagate outwards as semicircular arcs (Huygens' principle). At the top of the figure, the collision of the two arcs midway between the endpoints causes a second-order shock, which instantaneously becomes two high-velocity first-order shocks heading in opposite directions. Since the time of shock formation is equivalent to the radius of the maximal disc at the shock (Figure 4), the gap skeleton of a pair of endpoints is the set of shocks of the corresponding semicircular arcs, provided that this set includes a second-order shock. In summary, we have:

Proposition 4 *A second-order shock q satisfies $\pi(q) = \{a, b\}$ if and only if $\mathrm{gap}(a, b) \neq \varnothing$.*

Having related shocks to the gap skeleton, we now proceed to demonstrate the detection of gap skeleton for real examples.

[1] This informal argument is supported by the topological observation that a curve is the boundary of a set with empty interior; the ϵ-dilation realizes this connection in discrete domains.

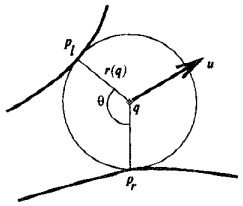

Figure 6: The geometry of the computation of the projection $\pi(q) = \{p_l, p_r\}$ of skeletal point q. The unit vector u points in the direction of the first-order shock.

5 Gap skeleton computations

To compute the gap skeleton, we first compute the projection $\pi(q)$ for all skeletal points q, then detect the pregap skeleton for pairs of endpoints, and finally search for gap skeleton, if any, within the pregap skeleton.

To compute the projection of q, we need not only the skeleton, but also local derivative properties. Specifically, the formulas for the two points, p_l and p_r, in the projection $\pi(q)$ are as follows:

$$p_l, p_r \triangleq q - r(q)\mathrm{Rot}(\mp\theta/2)u,$$

where $\theta = \arcsin(1/s)$, s is the speed of the first-order shock, and u is the unit vector in the direction of the shock, and $\mathrm{Rot}(\phi)$ is the two-dimensional rotation matrix through angle ϕ (Figure 6) [18]. Thus the we require the position, radius, direction, and speed of a skeletal point to compute its projection. Accurate estimates of all these quantities are provided by the algorithm for shock detection [19].

Our method for detecting pregap skeletal points exploits the uncertainty in the detection of endpoints of a curve. Given an end point uncertainty ϵ, the pregap with respect to endpoints a and b is the set of skeletal points whose projection falls within an ϵ-ball around these endpoints. The connectedness of the one-shock branches from the shock detection and the manner in which one-shock branches flow out of a two-shock branch greatly simplify the computation of the gap skeleton. Essentially we need only search for two pregap one-shock branches travelling in opposite directions, where each branch is induced by the same pair of endpoints.

We now illustrate the computation of gap skeleton with a variety of examples. For numerical as well as theoretical reasons, the curve flow $\frac{\partial \mathcal{C}}{\partial t} = \beta \vec{N}$, with β a constant, is formulated as the level set evolution of an evolving embedding surface [17]. In our simulations, the initial embedding surface is the signed distance function (slightly blurred to combat discretization) of the set of dilated curve fragments. Figure 7 depicts the process of gap detection. Note that as predicted the gap skeleton lies on the perpendicular bisector of the line joining two endpoints.

Figure 8 depicts our view of a complete system whereby grouped curves can be obtained from a real image. The

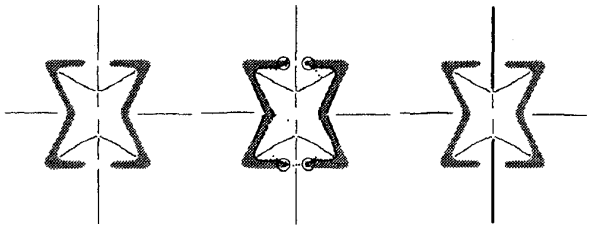

Figure 7: The process of gap detection. (LEFT) The initial level set is displayed in grey and the thin lines are first-order shock branches obtained from its outward evolution. Second-order shocks are not shown, giving a broken appearance to the skeleton. (MIDDLE) The ϵ-balls (circles) around the curve fragment endpoints (manually positioned) and the projection $\pi(q)$ (dots) of each skeletal point q are overlaid. (RIGHT) The detected gap skeleton (thickened lines) is shown connected due to Proposition 2.

curve fragments are dilated "curve-like" tangents [6] obtained from the discrete tangent map after logical/linear edge detection followed by relaxation labelling [10]. Figure 9 illustrates how the gap skeleton explains a visual illusion [14, p. 168, adapted from Köhler]. Figure 10 illustrates the computation of gap skeleton on curve fragments of a Kanizsa triangle. Finally, in Figure 11, curve fragments from the scene in Figure 1 are now appropriately grouped via the detected gap skeleton.

6 Conclusion

In this paper we have formally defined the gap skeleton and related it to the formation of shocks during curve evolution. The gap skeleton detected on a variety of examples groups fragments appropriately, and moreover has a physical interpretation as describing a virtual occluder. In Figure 12 we suggest three views of this virtual occluder. The "least commitment" virtual occluder is a disc centred where the radius function has a local minimum in the gap skeleton. When two gap skeletons align we gain a stronger impression of an occluder, and so we extend these discs along the gap skeleton, in a minimum area sense. Finally, from the perspective of a maximal possible region of the occluder, the third view is the other extreme.

Acknowledgements: Thanks to Yves Bérube-Lausière, Shamez Alibhai, and Mike Kelly.

References

[1] N. Ahuja and M. Tuceryan. Extraction of early perceptual structure in dot patterns: integrating region, boundary, and component gestalt. *Computer Vision, Graphics and Image Processing*, 48:304–356, 1989.

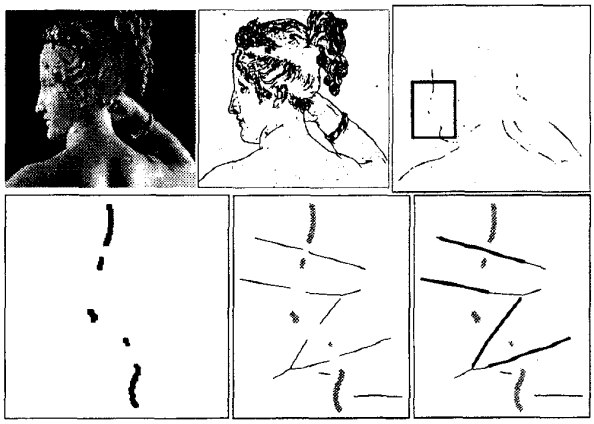

Figure 8: This figure depicts our view for obtaining grouped curves from a real image. (TOP LEFT) Original image. (TOP CENTER) Its discrete tangent map. (TOP RIGHT) We focus on a region of the "curve-like" tangents. (BOTTOM LEFT) The dilated "curve-like" tangents. (BOTTOM CENTER) Shock branches detected under outward evolution are overlaid. (BOTTOM RIGHT) The gap skeleton is shown with thickened lines.

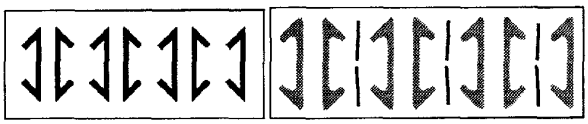

Figure 9: (LEFT) Illusion by Köhler. Note the groupings are seen across the gaps, and not between the parallel sides. Viewers may also see a subjective disc or ellipse (virtual occluder). (RIGHT) The detected gap skeleton (dark lines) selects a grouping of the three "bow-tie" pairs.

[2] J. August, K. Siddiqi, and S. W. Zucker. Fragemnt grouping via the principle of perceptual occlusion. Technical Report 96-02, McGill University, Centre for Intelligent Machines, 1996.

[3] H. Blum. Biological shape and visual science. *J. Theor. Biol.*, 38:205–287, 1973.

[4] R. Brockett and P. Maragos. Evolution equations for continuous-scale morphology. In *Proceedings of the IEEE Conference on Acoustics, Speech and Signal Processing*, San Francisco, CA, March 1992.

[5] L. Calabi and W. E. Hartnett. Shape recognition, prairie fires, convex deficiencies and skeletons. *American Mathematical Monthly*, 75:335–342, 1968.

[6] B. Dubuc and S. W. Zucker. Indexing visual representations through the complexity map. In *Fifth International Conference on Computer Vision*, pages 142–149. IEEE, 1995.

[7] J. Elder and S. W. Zucker. The effect of contour closure on the rapid discrimination of two-dimensional shapes. *Vision Research*, 33(7):981–991, 1993.

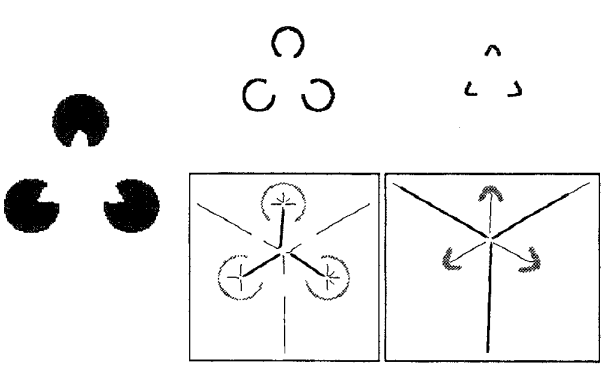

Figure 10: (LEFT) A Kanizsa triangle. (TOP RIGHT) Curve fragments from the Kanizsa triangle. (BOTTOM RIGHT) Shock branches obtained under outward evolution along with detected gap skeletons overlaid with thickened lines. Whereas each of the three circles form a separate group, and the triangle vertices form a single group, other fragment combinations cannot account for all fragment endpoints.

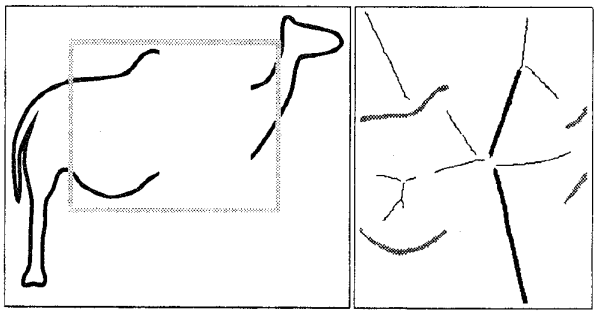

Figure 11: (LEFT) We focus on the region of fragmentation of the camel. (RIGHT) Observe that the gap skeleton (thickened lines) groups the fragments appropriately, appealing to the notion of a virtual occluder; in this case the actual occluder is a tree (see Figure 1).

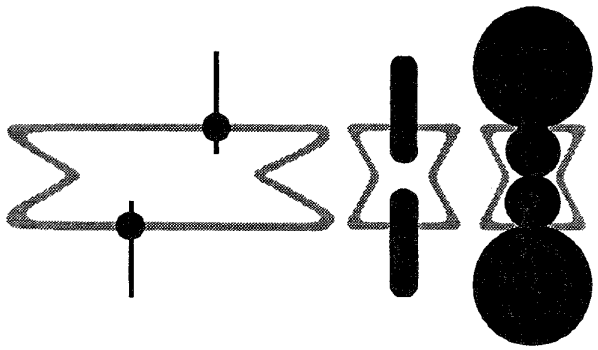

Figure 12: Virtual occluders. Curve fragments are in grey. (LEFT) A disc of radius $r(m)$ (in black) placed where the radius function takes on a local minimum m in the gap skeleton (thick line) is a "least commitment" occluder. (CENTRE) The aligned gap skeletons induce a stronger virtual occluder, in a minimum region sense. (RIGHT) A maximum region virtual occluder is described by a disc at each gap skeletal point q of radius $r(q)$.

[8] A. Guzman. *Decomposition of a visual scene into three-dimensional bodies*. Academic Press, 1969.

[9] M. Ilg. Knowledge-based understanding of road maps and other line images. In *Proceedings, 10th International Conference on Pattern Recognition*, volume 1, pages 282–284.

[10] L. A. Iverson and S. W. Zucker. Logical/linear operators for image curves. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1995.

[11] D. Jacobs. Grouping for recognition. MIT AI Lab Memo No. 1177, MIT, 1989.

[12] P. J. Kellman and T. F. Shipley. A theory of visual interpolation in object perception. *Cognitive Psychology*, 23:141–221, 1991.

[13] B. B. Kimia, A. R. Tannenbaum, and S. W. Zucker. Shapes, shocks, and deformations, I: The components of shape and the reaction-diffusion space. *International Journal of Computer Vision*, 15:189–224, 1995.

[14] K. Koffka. *Principles of Gestalt Psychology*. Harcourt, Brace & World, Inc., 1963.

[15] P. D. Lax. *Hyperbolic Systems of Conservation Laws and the Mathematical Theory of Shock Waves*. SIAM Regional Conference series in Applied Mathematics, Philadelphia, 1973.

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

[17] S. Osher and J. Sethian. Fronts propagating with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulations. *Journal of Computational Physics*, 79:12–49, 1988.

[18] D. Shaked and A. M. Bruckstein. On symmetry axes and boundary curves. In C. A. et. al., editor, *Aspects of Visual Form Processing*, pages 497–506, London, 1994. World Scientific.

[19] K. Siddiqi and B. B. Kimia. A shock grammar for recognition. Technical Report LEMS 143, LEMS, Brown University, September 1995.

[20] L. R. Williams and D. W. Jacobs. Stochastic completion fields: A neural model of illusory contour shape and salience. In *Fifth International Conference on Computer Vision*, pages 408–415. IEEE, IEEE, 1995.

[21] A. Zisserman, J. Mundy, D. Forsyth, and J. Liu. Class-based grouping in perspective images. In *Fifth International Conference in Conputer Vision*, pages 183–187. IEEE, 1995.