



Boundaries, shading, and border ownership: A cusp at their interaction

Matthew Lawlor^a, Daniel Holtmann-Rice^b, Patrick Huggins^b, Ohad Ben-Shahar^c, Steven W. Zucker^{a,b,*}

^a Program in Applied Mathematics, Yale University, New Haven, CT 06520-8285, United States

^b Department of Computer Science, Yale University, New Haven, CT 06520-8285, United States

^c Department of Computer Science, Ben Gurion University of the Negev, P.O. Box 653, Beer-Sheva 84105, Israel

ARTICLE INFO

Keywords:

Edge analysis
Shading analysis
Cusp
Differential geometry
Co-circularity
Long-range horizontal connections

ABSTRACT

The association of borders with “figure” rather than “background” provides a topological organizing principle for early vision. Such global influences have recently been shown to have local effects, with neuronal activity modulated by stimulus properties from well outside the classical receptive field. We extend the theoretical analysis of such phenomena by developing the geometry of interaction between shading, boundaries, and boundary ownership for smooth surfaces. The purely exterior edges of smooth objects enjoy a fold-type relationship between shading and boundary, due to foreshortening, while the background is cut off transversely. However, at cusp points in the image mapping the exterior boundary ends abruptly. Since such singular points are notoriously unstable, we conjecture that this process is regularized by a natural quantization of suggestive contours due to physiological boundary-detection mechanisms. The result extends a theorem about how contours must end to one that characterizes surface (Gaussian) curvature in the neighborhood of where they appear to end. Apparent contours and their interaction with local shading thus provide important monocular shape cues.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

Nearly a century ago Rubin (1915) established the distinction between *figure* and *ground* as an organizing principle for early vision, together with the correlative notion of *border ownership*, or the tendency to assign the border to the “figure” rather than the “ground.” Both are illustrated with his powerful vase demonstration: as the figure shifts from the vase to the faces, the boundary appears to shift as well. This effect is illustrated in Fig. 1 (left).

The subsequent Gestalt analysis of figure/ground sought to relate image properties to topological notions. The “one-sided or asymmetrical function of the contour, as Koffka (1935, p. 184) puts it, “can also be described by saying that contours have an ‘inside’ and an ‘outside.’ . . . The figure depends for its characteristics upon the ground on which it appears.” In modern terms we speak of boundaries separating image or textured regions, and of these following from occlusion in the scene domain.

More recent research in the neurophysiology of border ownership and figure/ground relationships has revealed its signature in both V1 and V2 (see Zhou et al., 2000; Lamme, 1995; Zipser et al., 1996; Lee et al., 1998). There exists a local effect from global pattern; in effect, cells respond more vigorously to edge stimuli when they are part of the figure than when they are part of the

ground, even though the portion of the stimulus within that cell’s receptive field remains unchanged. Clearly there is an integrative action-at-a-distance effect here (see Fig. 2). Our goal is, in effect, to study certain signature patterns for the figure as it approaches such an edge.

The modeling of border ownership involves building circuits that not only involve how the elements of the border fit together, e.g. by considering edge orientation, but also the requirement that the figure remains consistently on the correct side of the boundary. An early attempt to capture both of these requirements was a system for dot pattern organization based on the Gestalt concepts of (i) orientation good continuation, which dictated that boundaries were to be smooth and (ii) closure, which dictated that the boundary surrounded the cluster (see Zucker and Hummel, 1979). More recent modeling efforts follow this line of thinking, with the figure requirement in the computation reduced to edge polarity. That is, the circuit enforces the constraint that the direction of contrast across the edge is consistent (Sakai and Nishimura, 2004; Sakai and Nishimura, 2006). Such consistency in edge contrast is also relevant to shadow detection (Cavanagh and Leclerc, 1989).

But there is much more to figure/ground organization than is revealed by uniform colors or sign of contrast. The examples in Fig. 1 indicate that surface as well as topological properties should be considered, and recent research in stereo and its effect on organization confirms this (see e.g., Qiu and von der Heydt, 2005; Bakin et al., 2000; Orban, 2008). However much of this analysis thus far

* Corresponding author. Address: Program in Applied Mathematics, Yale University, New Haven, CT 06520-8285, United States.

E-mail address: steven.zucker@yale.edu (S.W. Zucker).

has been limited to those scenes with step edges in depth and sharp occlusions. Moreover, the complexity of natural scenes is often such that contrast can reverse along an edge.

1.1. Shape and shading

For smooth surfaces shading provides a powerful shape cue. While shape-from-shading is a classical computation in computer vision (Horn, 1975; Horn and Brooks, 1989), it is a global computation and involves solving a partial differential equation. Obtaining stable, reliable solutions is a delicate numerical matter and presumes reliable information at the boundaries or elsewhere (Oliensis, 1991). Texture has similar structural issues (Garding, 1992). Local computer vision methods are very approximate and can be highly biased (Pentland, 1984).

Nevertheless, humans are capable of inferring qualitative estimates of shape information from shaded images (see e.g., Ramachandran, 1988; Todd and Mingolla, 1983; Mingolla and Todd, 1986). Some data in this regard, from Koenderink's laboratory (Kang and Koenderink, 2007), are reproduced in Fig. 3. Clearly subjects were able to infer shape from shading information for this furrow-shaped object, although they are not perfect. While there have been neurobiological models of shape-from-shading (Lehky and Sejnowski, 1988) and the psychophysical perception of shadows (e.g., Knill et al., 1997), they neither achieve this level of per-

formance nor explain the biases. We seek to make a step in that direction in this paper.

1.2. Shape, boundaries and Koenderink's theorem

Line drawings are another source of shape information, and research in computer vision has also attempted recovery from them (see e.g., Barrow and Tenenbaum, 1981; Malik, 1987). Again the problem is delicate and not yet properly posed for our purposes. It is helpful to consider this in more conceptual detail.

While edge contours can arise from many different surface/viewer constraints (Fig. 4), we begin by considering those points in images that arise from the projection of places on a smooth, solid shape at which the tangent plane folds away from the viewer; in particular, at points on the surface where the viewing vector is orthogonal to the surface normal. This is a classical definition of the edge of an object. The corresponding curve on the surface is called the contour generator. Special structure arises when the tangent to the contour generator aligns with the view vector: these are the places at which the edge contour in the image appears to end, because the solid body of the object obstructs the contour generator from the viewer.

In a seminal series of papers (reviewed more formally later), Koenderink showed how the curvature of the (edge) contour in an image, its apparent curvature, carries information about surface curvature. His central result is that the sign of the apparent curva-



Fig. 1. Examples of boundaries around objects. (left) A Kanizsa-like display in which figure/ground shift spontaneously. Notice how the boundary tends to belong to the figure, whether this is seen as the dark or the white columns. (middle) When shading is introduced, much of the ambiguity is removed. Now smooth surfaces appear almost everywhere, even though they occlude one another. The perception of this type of interaction between smooth surfaces, in projection, is studied in this paper. Notice, in particular, how certain contours appear to end. (right) The mathematical possibilities for smooth surfaces are illustrated by this image of a Klein bottle, in which the exterior contour proceeds to the interior of the shape. The exterior lobe, on the left side, has much in common with the thumb in the middle image.

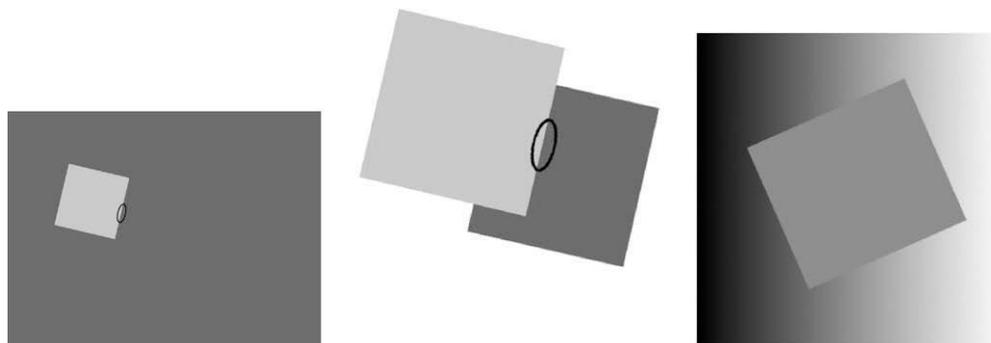


Fig. 2. Border ownership and border contrast. To illustrate the stimuli used in neurophysiological experiments, a bright/dark edge at the same orientation is shown with an elliptical receptive field. This edge can be part of the figure, as illustrated by (left) Gestalt closure considerations, while (middle) shows an apparent occlusion relationship. Notice that the sign of edge contrast changes for the foreground square because it covers two background surfaces. However, when non-constant contrasts are used such contrast reversals can occur (right). Since this is common in natural images, we seek to extend the relationship between shading and boundaries beyond absolute contrast comparisons.

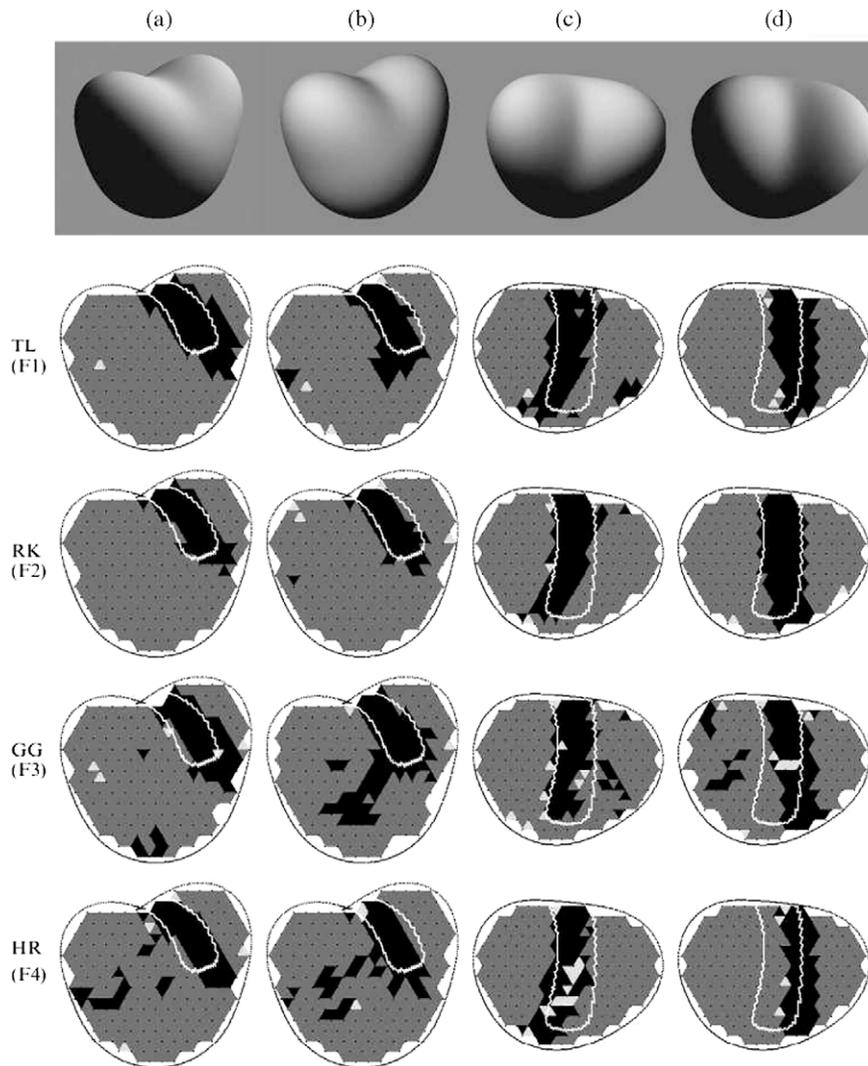


Fig. 3. Human performance in inferring shape from shading information. The gray areas indicate portions of the surface that are elliptic, or locally egg-shaped, while black areas indicate areas that are hyperbolic, or saddle-shaped. Such measures relate to surface curvature and will be important in our subsequent analysis. The white (parabolic) curves separate these two domains on the surface. Notice that, when the figure is in certain orientations, subjects can provide good qualitative estimates of shape, but there are biases. The perceived furrow (a and b), for example, often extends out of the hyperbolic region. Furthermore, when viewed straight on, the furrow indentation appears to shift with lighting (c and d). The theoretical challenge, therefore, is not only to predict the qualitative ability at estimating shape but to also explain the apparent biases. Data and figure from Kang and Koenderink (2007).

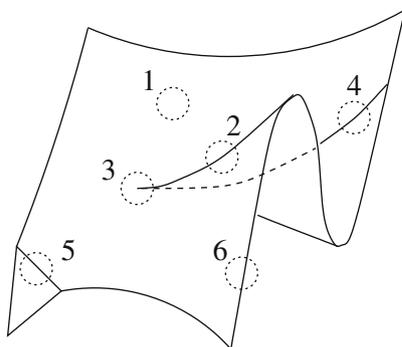


Fig. 4. Categories of points on a surface as they could project into an image: (1) a regular point, (2) a fold point, (3) a cusp, (4) a shadow point, (5) a crease point, (6) a boundary point. The viewpoint is taken to be at the upper left. If the surface is viewed from this position and a single light source also placed there, the fold (solid line) and the fold shadow (dashed line) appear aligned. In this paper we concentrate on points (2), folds, and (3), cusps, and study the interaction between the shading and the boundary in their neighborhood.

ture is the same as the sign of the Gaussian curvature for the surface at that point. This, basically, is what defined the gray and black regions in Fig. 3. Regions of positive Gaussian curvature are ellipsoidal, or shaped like a portion of an egg; regions of negative Gaussian curvature are saddle-shaped. The parabolic lines that separate them, under projection into the image, become points of inflection when they land on visible contours.

If we define image contours to be ideal edges; i.e., to be precisely those contours that arise when the surface normal is orthogonal to the vector pointing toward the viewer, then there is an important corollary to Koenderink's result that dictates the way contours should end, provided this is in an area of negative Gaussian curvature: they should end in a concavity that corresponds to the change in curvature caused by entering the saddle-shaped area. They should not, in particular, end by remaining strictly convex.

Various examples of ending contours are shown in Fig. 5. They have been obtained by a variety of techniques, from strictly local "edge detector" computational analysis (Iverson and Zucker, 1995), which works for very smooth, artificial surfaces, to a curve-inference system thought to model V1 (Ben-Shahar et al.,

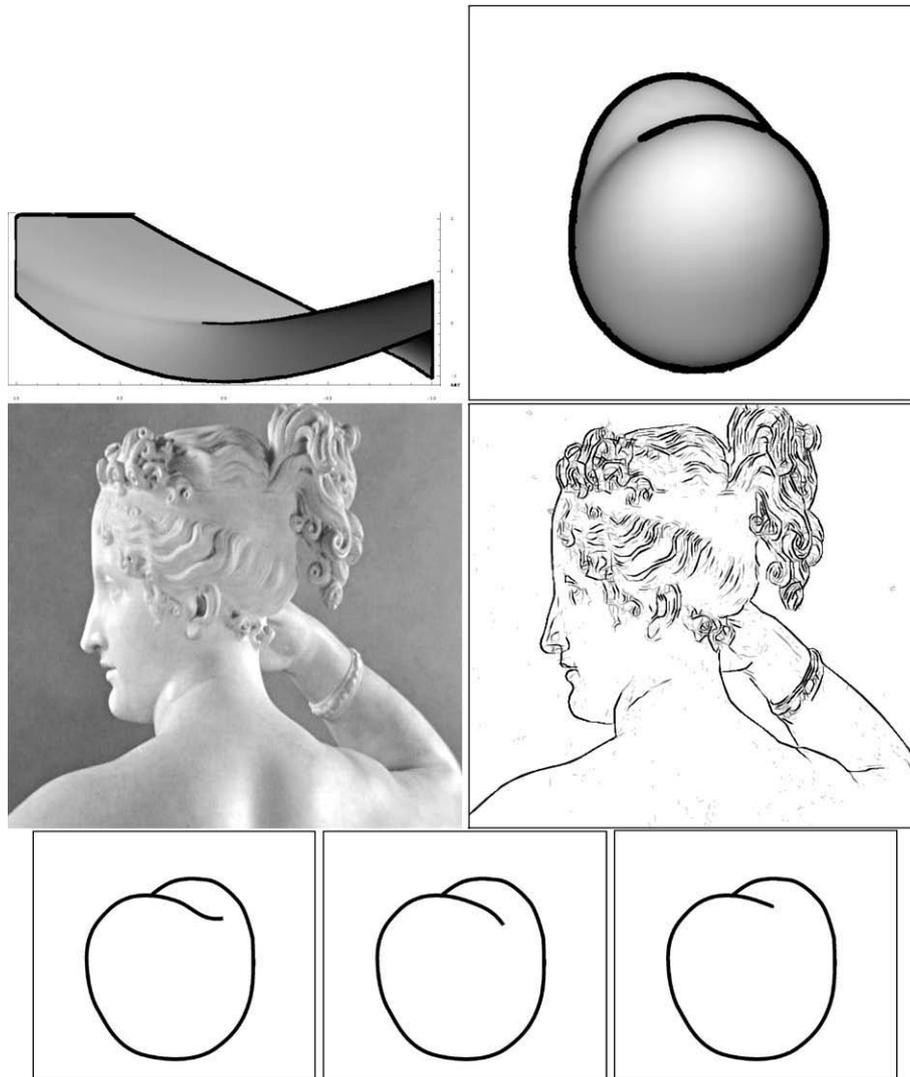


Fig. 5. Example images containing edge contours that end. They were obtained by different methods. On the top row, edge contours (in black) were computed with a local edge detector. In the folded cloth (upper left) the contour appears to end in a manner consistent with Koenderink's corollary, but for the furrow shape (upper right) it appears to end in an inconsistent manner. (This furrow shape is a rotated version of the one in Fig. 3.) In the Paolina edge map (middle), the contours were computed by a model for curve detection that is consistent with long-range horizontal connections in visual cortex. Note that now there is a mixture: some contours obey the Koenderink corollary and others do not. (bottom) Three example drawings of contours indicating the furrow shape which end properly (left) or improperly (middle, right) among these bottom drawings, which ones appear incorrect? Correct? What is the effect of varying the length? Our goal in this paper is to understand which aspects of visual shape are informed by such computed and drawn contours.

2003; Ben-Shahar and Zucker, 2003). From these it is clear that computed contours may differ substantially from the requirements of Koenderink's Corollary: sometimes they appear to terminate at a direct convexity.

Such deviations could, of course, simply be numerical and quantization artifacts. It may be that the contours are ending properly, except that their extension into the image is so short that the pixelization process collapses the apparent curvature of this portion of the contour to the wrong value. But this somehow goes against our intuitions, because artists so rarely end contours "correctly."

Another possibility is that the contour is not ending at the correct location. Since Fig. 5 (top, left) was composed from a mathematically-defined surface, we know exactly where the contour should end: at the cusp point where the surface just stops occluding itself. However, and most importantly, only rarely do computed contours end at this point. To the extent that these computed contours match perceived contours, we can conclude that further analysis is necessary to move the contour ending

away from the cusp and into the shape. This is precisely our goal in this paper.

1.3. Overview of the paper

We seek to integrate the processing of surface shading information with apparent contour information to derive constraints that could be active in shape inference. Such integration of edge and shading information is classical in computational vision, provided the world is simple (e.g., blocks (Waltz, 1975)). We seek to accomplish this integration in a neurobiologically-plausible fashion for smooth surfaces. As such, it is necessary to combine several seemingly disparate pieces together.

The key to combination is geometrical analysis. This follows a line of research in which orientation selectivity in visual cortex is abstracted in differential-geometric terms, with preferred orientation at a point corresponding to the tangent at that point. If we are dealing with boundaries, this tangent is to a curve; if shading, this tangent is to an isophote (a level set in the intensity distribution.)

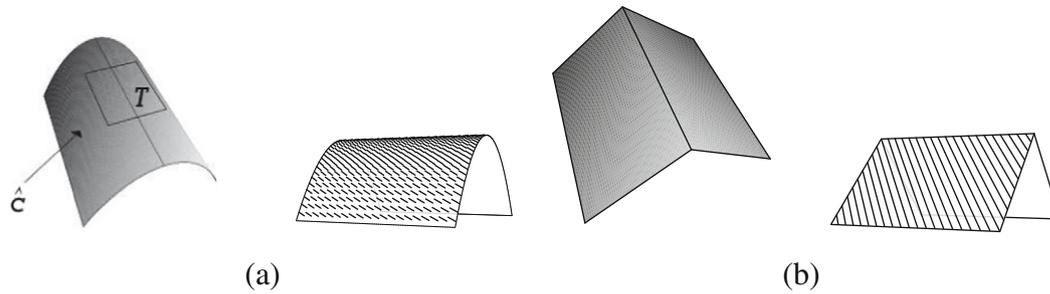


Fig. 6. Illustration of shading flow in the neighborhood of an edge. When a shaded surface is viewed such that an edge appears, the shading flow field takes on different appearances depending on the nature of the edge. A fold occurs (a) when the surface bends smoothly away from the viewer (the typical occlusion case), and the shading flow field appears tangent to the edge. A fold occurs when the view direction (\hat{c}) lies in the tangent plane (T) to the surface. At a cut (b), the surface is discontinuous (or occluded), and shading flow is generally non-tangent to the edge.

For shading, then, there is a field of tangents, called the shading flow field. While these constructs are defined more carefully in the next Section, for now we illustrate their power with a few examples.

In Fig. 6 we illustrate how the shading flow must approach the edge of a smooth surface in a tangential fashion; for a non-smooth break in the surface orientation, or for a portion of the background which has been occluded, the shading flow approaches the boundary in a transverse fashion. These claims have been confirmed psychophysically (Palmer and Ghose, 2008) and are developed in detail in the next Section.

While both contours and shading distributions have a lift into (position, orientation)-space, or orientation columns, they enjoy very different spatial frequency representations. Contours are essentially high-frequency; although extended in length, their energy in the other direction is concentrated within small spatial extents. Shading distributions, on the other hand, are more spread out in two-dimensions; they are much lower in their spatial frequency content. Thus we exploit the large spatial frequency range of representations in cat (Issa et al., 2000) and primate (Silverman et al., 1989). It has been observed that individual neurons can respond with substantial selectivity to stimulus position, orientation, and spatial frequency (Geisler and Albrecht, 1997; Bradley et al., 1987); we take this to be the basis for the representation of contours and shading flows. In Fig. 7 we demonstrate that this kind of representation can guide perceived surface structure and appearance.

Since local information in natural images is ambiguous about both contours and shading, models for integrating such information over local, overlapping neighborhoods are developed according to the principle that their neural substrate consists in the visual cortical machinery of columnar organization with long-range horizontal connections. Such models, taken individually,

have predicted connection structure through second order (Ben-Shahar and Zucker, 2003). However, because the shading flows involve area connections rather than linear ones, they can spread into background areas unless properly contained. We therefore show how to introduce boundary conditions from the contour system to constrain the spread of shading flows. This permits them to grow to the occluding boundary but no further (unless warranted by strong contextual support).

Thus far the material described is a review of earlier work, integrated around the theme of border ownership and the relationship between boundaries and shading for smooth surfaces. But as noted earlier, external boundaries may move into a shape and end at cusp points, and their analysis is the new contribution in this paper. In particular, viewing edge contours as the boundary conditions for shading, cusps are then those places where the edge contour ends; ideal edge contours, then, cannot limit the shading beyond the cusp. But there is a relationship between the shading and the contour dictated by the underlying surface geometry, and it suggests thinking about image contours in a broader sense. We now explain.

Stereoscopic vision provides information about surface structure by exploiting slight differences in viewpoint (Li and Zucker, 2006), and dynamically changing viewpoints is a classical technique in active computer vision (e.g., Kutulakos and Dyer, 1994). A related idea has now been developed in a clever computer graphics application to produce non-photorealistic renderings of scenes (DeCarlo et al., 2004). In particular, given a complete three dimensional model of an object, once the viewing position is given the edge contours can be computed exactly as the projection of those surface points at which the viewing vector is orthogonal to the surface normal. However, in practice this edge map is very unsatisfying – it neither produces a firm impression of surface structure nor resembles artists' renderings. These ideal physical edges just seem too abbreviated.

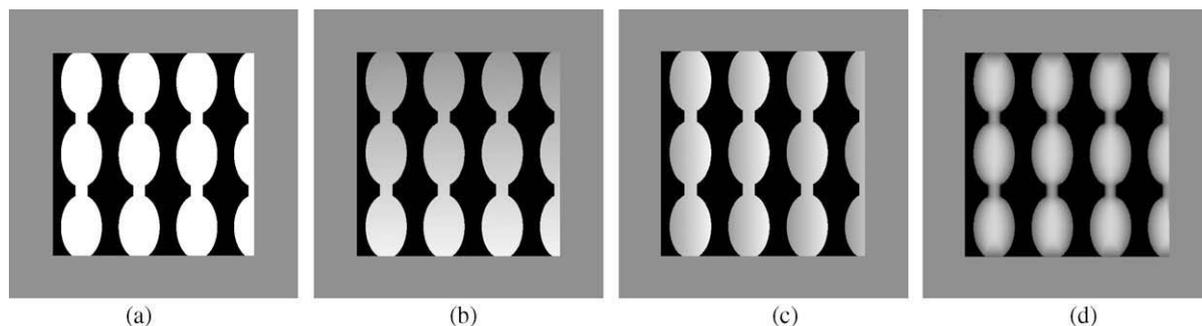


Fig. 7. Three examples of the Kanizsa-type display with shading, increasing in three-dimensionality from left to right. (a) Figure/ground reversals are common with no shading. (b) The white area is shaded so that it meets the boundary transversally – notice how flat it appears. (c) The white area is shaded with a vertical gradient. This meets the boundary tangentially at some points. (d) The shading flow in the white area meets the boundary tangentially; this appears the most rounded.

De Carlo et al.'s (2004) idea is to extend the exact edges with those *suggestive edges* that would be visible from certain nearby viewpoints, as if the viewer had moved slightly or the surface had rotated slightly. Technically this amounts to relaxing the orthogonality requirement between the view vector and the surface normal to something “close” to orthogonal. (Appropriate definitions are in Section 6.1.) The result is much more pleasing graphically.

The concept of a suggestive edge is relevant in the neighborhood of a cusp point, because it provides for a regularization – an extension – of the contour ending at it. But two questions arise and both are addressed (see Fig. 8).

First, how far should the suggestive contour be extended? Mathematically this is a difficult question to answer, since it amounts to asking how small a number should be so that it is “equivalent” to zero. (That is, since the surface normal is no longer precisely orthogonal to the view direction, how much can it be allowed to deviate?) We instead treat the question operationally, by evaluating these quantization limits through the edge detection process. Within surfaces of the sort we are illustrating, for a suggestive contour to exist there must still be some surface curvature, so that the orthogonality requirement can be met with a

small movement. Our edge detection systems in effect “detect” this foreshortening even though the formal edge condition is not met. As examples show, the results are quite natural (Fig. 9). Our thesis in this paper is that such computed lines and edges provide information about the geometry of the underlying surface, and it is this geometry that is exploited by the visual system.

The second question asks what surface information can be inferred in the neighborhood of these extended edges? Working from a local model, we show how a corollary to Koenderink's result can be formulated across the suggestive contour.

Although the basics of the models we describe could be implemented in the earlier cortical visual areas, they may certainly involve higher-level areas and feedback and feedforward as well as lateral, horizontal connections. There is certainly evidence that higher-level areas could be involved ((Caplovitz and Tse, 2007; Pasupathy and Connor, 2002). And, while influences beyond those considered here may matter, such as haptics (Gepshtein and Banks, 2003), attention (Craft et al., 2007)) and others (Kersten and Schrater, 2002), the relationship between viewing geometry and surface geometry is fundamental.

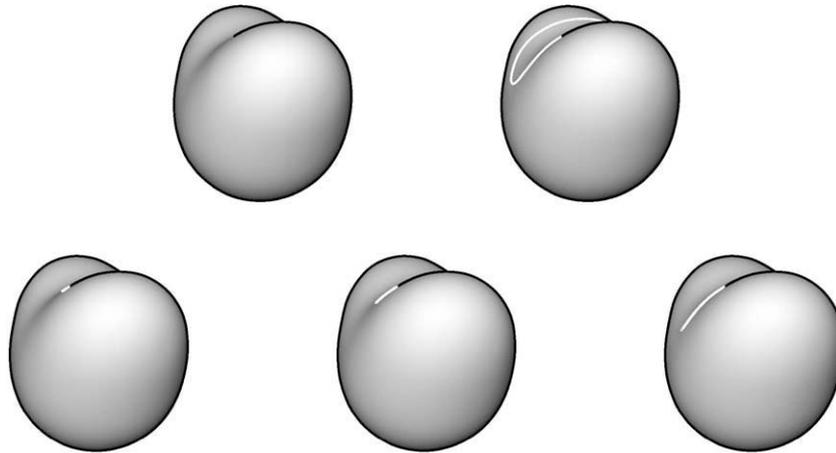


Fig. 8. How far should a suggestive contour extend? (top left) The idea edge contour (black) on the furrow shape, given by those positions at which the view vector and the surface normal are orthogonal. Notice the deep shading that extends from the edge; and the fact that it is line-like. (top right) The actual edge contour extended with the parabolic curve (in white). The cusp point is where they meet for this example. (bottom) Three extensions of the suggestive contour. Which one appears “correct?” What do they suggest about the underlying surface? Although this example shows the suggestive contour coinciding with the parabolic curve, this does not generally happen. The Gaussian curvature is what is important and we study it later in the paper (for this example it is zero at that point).

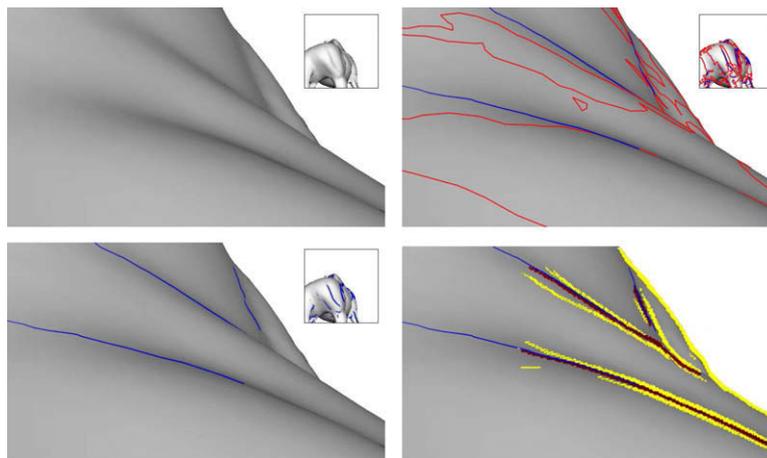


Fig. 9. Example on horse image with suggestive and computed contours. (top left) Zoom on region of horse image rich in ending contours (see insert). (top right) Suggestive contours in blue and parabolic lines in red. Notice how the parabolic lines meander over the surface, although there are places at which they agree well with the suggestive contours. (bottom left) Suggestive contours: where should they end? (bottom right) Computed edges (green) and lines (red). Notice how their location highlights aspects of the agreement between suggestive and parabolic lines. More generally, such computed lines and edges provide information about the underlying surface structure. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2. Folds

Occlusion edges arise from discontinuities in depth. In the case of smooth objects these discontinuities arise from singularities in the projective mapping from object surfaces to the image plane (Koenderink and van Doorn, 1976).

Let Σ be a smooth (C^2) surface, i.e., locally $\Sigma : X \subset \mathbb{R}^2 \rightarrow Y \subset \mathbb{R}^3$, where X is the space of parameters of the surface and Y is a portion of Euclidean space in which the surface is embedded, and let $\hat{\mathbf{c}} \in \mathbb{S}^2$ (the unit sphere) be the viewing direction. If $\Pi_{\hat{\mathbf{c}}}$ is orthographic projection ($\Pi_{\hat{\mathbf{c}}} : Y \rightarrow Z \subset \mathbb{R}^2$, where Z is the image domain), then the surface to image mapping is $\Pi_{\hat{\mathbf{c}}} \circ \Sigma$, which is from \mathbb{R}^2 to \mathbb{R}^2 (note that for what follows, the choice of orthographic projection can easily be generalized).

This implies, via a theorem of Whitney (1955), that the only stable singularities (i.e., points where the Jacobian is not of full rank) are folds (type 2 points in Fig. 4) and cusps (type 3 points in Fig. 4). In this section we consider folds and then, with this in place, move on to cusps.

The *fold* is the set of points in the image where:

- the mapping $\Pi_{\hat{\mathbf{c}}} \circ \Sigma$ is singular and
- Σ is smooth.

Complementary to folds are *cuts*, or the set of image points where image intensity is discontinuous and the projective mapping is regular (non-singular). This includes image edges arising from surface boundaries (e.g., the edge of a sheet of paper), surface normal discontinuities (e.g., polyhedral edges), reflectance changes and shadows.

Edges around smooth objects thus have a structure which is fold on one side (the occluding side) and cut on the other (the occluded side). This difference between folds and cuts results in an observable photometric difference (Rieger, 1990; Koenderink and van Doorn, 1993; Koenderink, 1990). If we define the *shading flow field* to be the vector field of tangents to isoluminance contours in the image (Breton and Zucker, 1996), we can differentiate between the fold and cut neighborhoods of edges as follows. On the fold side of an edge, the shading flow field is tangent to the occluding contour, while on the cut side of an edge the field is transverse (non-tangent) (see Fig. 6). Furthermore it has been proven that this is the generic case (Dufour, 1983; Huggins and Zucker, 2001).

At an edge point p of occluding contour γ in an image we can define two semi-open neighborhoods, N_p^A and N_p^B , where the surface to image mapping is continuous in each neighborhood. To enumerate the different arrangements around a fold (Huggins and Zucker, 2001) we classify p as follows:

1. *fold-cut*: The shading flow is tangent to γ at p in N_p^A and the shading flow is transverse to γ at p in N_p^B , with exception at isolated points.
2. *cut-cut*: The shading flow is transverse to γ at p in N_p^A and in N_p^B , with exception at isolated points.
3. *fold-fold*: The shading flow is tangent to γ in N_p^A and in N_p^B , with exception at isolated points.

In general, occlusion edges are the only edges that will appear as *fold-cut*, where the *fold* side corresponds to the foreground and the *cut* side to the background. Other edge types will appear as *cut-cut*, including edges such as shadows and reflectance discontinuities; these may be further classified (Witkin, 1982; Breton and Zucker, 1996; Zhang and Bergholm, 1997). We note that the third class, *fold-fold*, is accidental, or non-generic (Freeman, 1994), in that it requires a precise alignment of two folds.

This categorization, while useful, is qualitative. Appearance over light-source variation is quantitative.

2.1. The appearance of folds for a model surface patch

The local appearance of a fold in an image can be modeled given simple (though easily generalized) assumptions about surface geometry, reflectance, and lighting; this enables us to analytically compute the relation of the shading flow field to the occluding contour. We are especially concerned with how this appearance varies as a function of lighting. Our results show that the shading flow is not only tangent to the occluding contour, but is also on average parallel to the occluding contour in a neighborhood around it. This parallelism is stable with respect to illumination variation. Furthermore we show that the magnitude of the image gradient also serves to indicate the presence of a fold. Related image gradients will be important later when we move on to the analysis of cusps and suggestive contours.

Consider a surface patch (Σ) viewed such that it exhibits a fold. Let the surface patch be approximated by its second-order Taylor expansion about the origin, where the z -axis direction is the surface normal direction at the origin, the x -axis direction ($\hat{\mathbf{x}}$) and the y -axis direction ($\hat{\mathbf{y}}$) are the principal directions of the surface at the origin, and a and b are the corresponding principal curvatures with $a > b$ (Σ is a Monge patch)

$$z = \frac{1}{2}ax^2 + \frac{1}{2}by^2 + O((x,y)^3). \quad (2.1)$$

Let the origin appear as a fold in the image; i.e., we assume that the view direction ($\hat{\mathbf{c}}$) lies in the x - y plane and subtends an angle of θ with $\hat{\mathbf{x}}$.

Let the surface be Lambertian with constant albedo (without loss of generality we set the albedo to unity), and let the light source be a point light source at infinity in the direction \mathbf{l} (these assumptions are easily generalized). The radiance of a point on the surface is then given by

$$L = \hat{\mathbf{n}} \cdot \mathbf{l}, \quad (2.2)$$

where $\hat{\mathbf{n}}$ is the unit surface normal at a point, defined as

$$\hat{\mathbf{n}} = \left(-\frac{\partial z}{\partial x}, -\frac{\partial z}{\partial y}, 1 \right) / \left(1 + \frac{\partial z^2}{\partial x^2} + \frac{\partial z^2}{\partial y^2} \right)^{\frac{1}{2}}.$$

Given the framework set up previously, we now compute the shading flow field (\mathbf{s}); recall the shading flow field is defined as the vector field of tangents to the isoluminance contours in the image, i.e., if I is image irradiance, then \mathbf{s} is orthogonal to ∇I . To do this we first compute the gradient of scene radiance, expressed in terms of a local coordinate frame on the surface. We define $\hat{\mathbf{u}}$ and $\hat{\mathbf{v}}$ to be the principal directions (if \mathbf{v} is a vector, then $\hat{\mathbf{v}} = \mathbf{v}/\|\mathbf{v}\|$)

$$\begin{aligned} \hat{\mathbf{u}} &= \left(1 - \frac{a^2}{2}x^2 \right) \hat{\mathbf{x}} - \left(\frac{a^2b}{|a-b|}xy \right) \hat{\mathbf{y}} + (ax)\hat{\mathbf{z}}, \\ \hat{\mathbf{v}} &= \left(-\frac{b^2a}{|b-a|}yx \right) \hat{\mathbf{x}} - \left(1 - \frac{b^2}{2}y^2 \right) \hat{\mathbf{y}} - (by)\hat{\mathbf{z}}. \end{aligned} \quad (2.3)$$

The corresponding principal curvatures, κ_u and κ_v , are

$$\begin{aligned} \kappa_u &= a - \frac{3}{2}a^2x^2 - \frac{1}{2}ab^2y^2, \\ \kappa_v &= b - \frac{3}{2}b^3y^2 - \frac{1}{2}ba^2x^2. \end{aligned} \quad (2.4)$$

The radiance gradient in the $\hat{\mathbf{u}}-\hat{\mathbf{v}}-\hat{\mathbf{n}}$ coordinate frame is then

$$\nabla L = l_u \kappa_u \hat{\mathbf{u}} + l_v \kappa_v \hat{\mathbf{v}}, \quad (2.5)$$

where $\mathbf{l} = (l_u, l_v, l_n)$ in the $\hat{\mathbf{u}}-\hat{\mathbf{v}}-\hat{\mathbf{n}}$ frame.

We now assume a view centered coordinate frame such that $\hat{\mathbf{x}} = \hat{\mathbf{c}}$. The transformation from the Monge patch coordinate system to the view centered coordinate system is simply a rotation by $-\theta$ about the z -axis.

Assuming radiometric calibration of the lens, the image irradiance at a point p in the image can be found as the projection of the surface radiance onto the y - z plane

$$I(p) = L(\Pi_{\hat{\mathbf{x}}}^{-1}(p)). \quad (2.6)$$

The shading flow field in the image is the projection of the vector field that is orthogonal to the radiance gradient onto the y - z plane

$$\begin{aligned} \mathbf{s} &= \Pi_{\hat{\mathbf{x}}}((\nabla L)^\perp) \\ &= (-l_v \kappa_v u_y + l_u \kappa_u v_y) \hat{\mathbf{y}} + (-l_v \kappa_v u_z + l_u \kappa_u v_z) \hat{\mathbf{z}}, \end{aligned} \quad (2.7)$$

where $(\nabla L)^\perp$ is the vector orthogonal to ∇L on the surface. This expression gives the shading flow field in terms of our model.

To compare the shading flow field orientation to the edge orientation, we compute the slope of the shading flow field in the image (computing slope as opposed to orientation simplifies the analysis). The slope of the image shading flow field is

$$\omega = \frac{-l_v \kappa_v u_z + l_u \kappa_u v_z}{-l_v \kappa_v u_y + l_u \kappa_u v_y}. \quad (2.8)$$

Since the view direction is in the x - y plane, the slope of the occluding contour at the origin is zero.

In later computations it will be convenient to express l_v in terms of ω and l_u

$$l_v(\omega, l_u) = l_u \begin{pmatrix} \kappa_u \\ \kappa_v \end{pmatrix} \begin{pmatrix} v_z - v_y \omega \\ u_z - u_y \omega \end{pmatrix}. \quad (2.9)$$

Now that we have computed a description of the shading flow field, we consider its appearance under variable illumination. We are interested in the orientation of the shading flow field relative to the orientation of the occluding contour, specifically in whether the tangency described in Section 2 is observable, and, if so, whether or not it is stable.

A straightforward numerical computation yields the mean and variance of the shading flow field at every point on a surface patch,

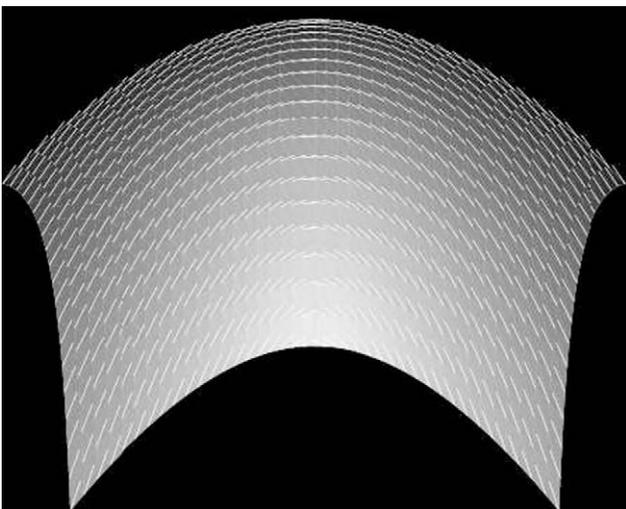


Fig. 10. The statistics of a fold. A plot of the mean orientation of the shading flow field (shown as vectors) and its variance (indicated by intensity), for one instance of our model. Observe how the shading flow aligns with the fold (located at the top of the image) as the fold is approached and how the variance of this alignment decreases as well, indicating that this alignment is a stable image feature.

per our model, assuming the light source direction \mathbf{l} is a random point on the hemisphere centered on $\hat{\mathbf{z}}$ (see Fig. 10). The computation shows that the mean shading flow field is not only tangent to the occluding contour, but also parallel to it in a large neighborhood. Furthermore, the variance of the shading flow field is small close to the contour and gradually increases away from it. Thus the observed parallelism is stable in the neighborhood of the contour with respect to variable illumination.

Given the observed parallelism of the shading flow to the occluding contour, we restrict our analysis to image points normal to the edge point under consideration. In what follows we take the origin to be our occluding edge and we consider the statistics of the shading flow field at image points normal to the edge, i.e., surface points in our model satisfying $y = 0$.

If in a given image we observe an edge and the shading flow field normal to the edge, we would like to be able to use the relative orientation of the shading flow field to the edge as evidence of the edge's type: e.g., if the shading flow field on one side of the edge is roughly parallel to the edge we might label that side of the edge as a fold.

Towards this, we compute the probability density of shading flow field slope under variable illumination. We assume a given surface viewed from a given direction, with a distribution on light sources that is uniform over light source directions and falls off in magnitude. Choosing simply a Gaussian, the probability density of light sources is

$$\rho_1(\mathbf{l}) = \frac{1}{(\sqrt{2\pi}\sigma)^3} e^{-\frac{1}{2\sigma^2}(l_x^2 + l_y^2 + l_z^2)}. \quad (2.10)$$

We acknowledge that there are many choices for $\rho_1(\mathbf{l})$, some of which may arguably be more appropriate. In the context of examining appearance under variable illumination, we choose our distribution so as not to favor any particular light source direction.

The probability density of the slope of the image shading flow field can be obtained by combining this result with Eq. (2.9)

$$\rho_\omega(\omega; z, a, b, \theta) = \frac{1}{2\pi\sigma^2} \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}(l_u^2 + l_v(\omega, l_u)^2)} \left| \frac{\partial(l_u, l_v)}{\partial(\omega, l_u)} \right| dl_u, \quad (2.11)$$

where z is the distance to the occlusion edge in the image, a and b are the principal curvatures of the Monge patch, and θ is the angle between the viewing direction and the first principal direction.

The resulting family of probability density functions is plotted in Fig. 11a. The most prominent feature of ρ_ω is the ridge close to $\omega = 0$, which indicates that the shading flow is on average parallel to the occluding contour. The sharpness of the ridge indicates the stability of this parallelism with respect to illumination variation. As expected from the tangency of the shading flow to the contour, the function ρ_ω approaches a Dirac delta function $\delta(\omega)$ when $z \rightarrow 0$. The observed ridge is always present in ρ_ω , though its spread and position are modulated by the view direction and the curvatures.

Since we do not know the viewing angle in advance, it is also important to look at the marginal density of the slope $\rho_\omega(\omega; z, a, b, \theta)$ integrated over the viewing angle θ

$$\rho_\omega^0(\omega; z, a, b) = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \rho_\omega(\omega; z, a, b, \theta) d\theta. \quad (2.12)$$

We present numerical results in Fig. 11b. ρ_ω^0 has the same qualities as ρ_ω ; the observed ridge is centered on $\omega = 0$ and remains quite sharp, especially near the occluding contour.

A further observation about folds is that the image irradiance gradient magnitude depends on foreshortening: a radiance gradient on the surface will be amplified by the foreshortening that occurs as the fold is approached. In the limit, the foreshortening will give rise to an infinite image irradiance gradient. This foreshorten-

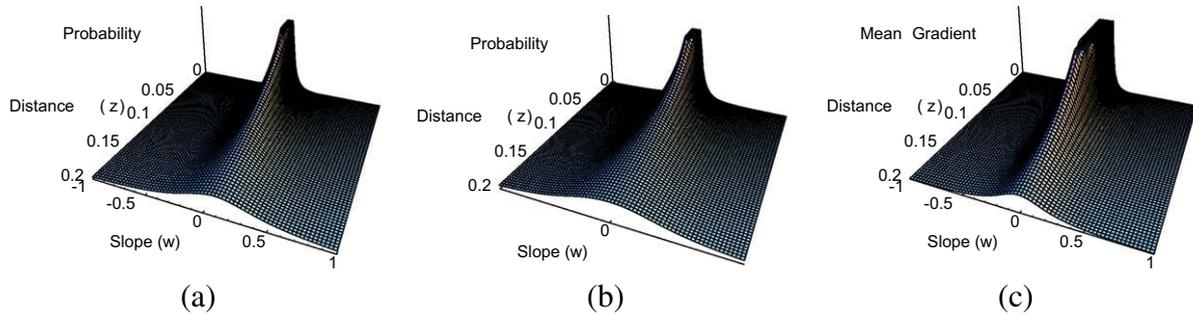


Fig. 11. (a) The family of probability density functions for the slope of the shading flow field ($\rho_{\omega}(\omega)$). (b) $\rho_{\omega}^0(\omega)$, (a) integrated over viewing directions. (c) The expectation of the image gradient magnitude ($\mathbb{E}[\|\nabla I\|]$). In all plots *Distance* is measured in the image plane normal to the contour. The surface parameters are $a = 1$ and $b = 0.3$. In (a) and (c) $\theta = -\pi/6$. The plots are clipped for display purposes.

ing is maximal when the radiance gradient is in the direction of the tilt of the surface near the fold, i.e., when the shading flow field is parallel to the occluding contour. Thus we expect the image irradiance gradient magnitude as a function of shading flow field slope to correlate with the probability density function of slope.

The image irradiance gradient is

$$\nabla I = \frac{\|\nabla L\|}{\left\| \Pi_{\hat{\mathbf{x}}} \left(\frac{\nabla L}{\|\nabla L\|} \right) \times \hat{\mathbf{s}} \right\|} \hat{\mathbf{x}} \times \hat{\mathbf{s}}. \quad (2.13)$$

Using Eqs. (2.13) and (2.10), we calculate the expectation of the magnitude of the image gradient

$$\mathbb{E}[\|\nabla I\|] = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} \|\nabla I\| e^{-\frac{1}{2\sigma^2}l_u^2} dl_u, \quad (2.14)$$

where l_v is given by Eq. (2.9).

Fig. 11c shows the mean magnitude of the image irradiance gradient. It is clearly correlated with the observed parallelism of shading flow field. In this sense, the gradient magnitude can be seen as an indicator of where in the distribution (Fig. 11a) a particular instance of shading flow field orientation lies.

The above analysis suggests that the alignment of the shading flow with the contour of the fold is stable in the neighborhood of the occluding contour. It is insensitive to the lighting conditions, the geometry of the surface, and the view direction. We now proceed to verify the conclusions of our model with real images.

2.2. Measuring the appearance of folds

We need to evaluate the results of the previous section in the context of real images. The model discussed in the previous section does not capture all of the variability present in real images: albedos vary, surfaces are often non-Lambertian, small scale surface structure can significantly affect image measurements, and the imaging process is noisy. Given these factors, how robust is the parallelism between the shading flow field and the occluding contour, which we observed in the model, in real images?

We show two experiments in answer to this question. First, for a very simple object, we evaluate the statistics of the shading flow field and the gradient in images taken under variable illumination. For simple objects we show that these agree well. However in viewing a scene one does not have access to the statistics over illumination, so we need to consider what a *single* image yields for shading flow. This, it turns out, is not well behaved when it is measured only locally.

2.2.1. Average appearance over lighting

We measured the statistics of a simple object (a toy duck) under variable illumination. Sixty-four images were taken, each illuminated by a single distinct flash. The flashes were uniformly distrib-

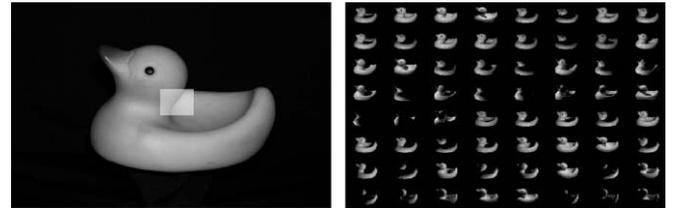


Fig. 12. (left) An experimental subject: a toy rubber duck. (right) Images of the duck under illumination from 64 different point sources.

uted over the hemisphere centered over the object. This array of point light sources is effectively a sampling of the theoretical uniform distribution we considered in the previous section. Fig. 12 shows the images captured for this object.

We computed the mean and variance of the shading flow field and the irradiance gradient in the image. Some care needs to be taken in these computations. Consider the shading flow field measured at a point in image i illuminated by light source direction $\mathbf{l}_i = (l_u, l_v, l_n)$; if the light source illuminating image j is $\mathbf{l}_j = (-l_u, -l_v, l_n)$, then the shading flow vectors in images i and j will cancel one another at that point, even though in some sense their orientation is the same. Thus we define orientation to be π -periodic in our computations.

The expected shading flow field is then

$$\mathbb{E}[\hat{\mathbf{s}}] = \angle_{\times \frac{1}{2}} \left(\int_{\mathbf{l} \in \mathbb{S}^2} \angle_{\times 2}(\hat{\mathbf{s}}) \|\mathbf{l} \times \hat{\mathbf{z}}\| d\mathbf{l} \right), \quad (2.15)$$

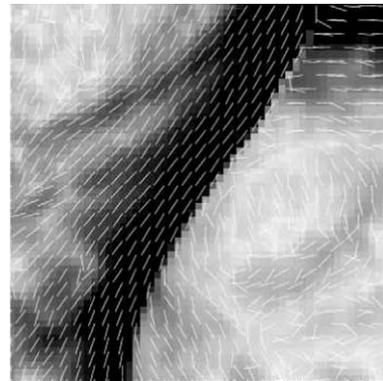


Fig. 13. The mean shading flow field, $\mathbb{E}[\hat{\mathbf{s}}]$, in the neck region of the duck (highlighted region of Fig. 5), shown as vectors, superimposed on its variance, $\mathbb{V}[\hat{\mathbf{s}}]$. Observe the parallelism between the shading flow field on the fold side of the occluding contour and the contour itself, and the accompanying low variance. Compare with Fig. 10.

where $\angle_{\times 2}$ and $\angle_{\times \frac{1}{2}}$ map vectors to double- and half-angled vectors, respectively. We define the variance of the shading flow field as

$$\mathbb{V}[\hat{\mathbf{s}}] = \int_{\mathbf{l} \in \mathbb{S}^2} \|\angle_{\times 2}(\hat{\mathbf{s}}) - \angle_{\times 2}(\mathbb{E}[\hat{\mathbf{s}}])\|^2 \|\mathbf{l} \times \hat{\mathbf{z}}\| d\mathbf{l}. \quad (2.16)$$

Fig. 13 shows the expected shading flow field for the duck image set (Fig. 12), in a region of the image centered on the back of the duck's neck. The alignment of the shading flow field with the occluding contour on the fold side of the edge is clear. In contrast, the shading flow field on the cut side of the edge exhibits no such alignment. Furthermore, on the fold side of the edge the variance is low, thus the observed parallelism is robust with respect to illumination variation, while the variance is high on the cut side, and so even if there exist lighting conditions where the shading flow field here appears parallel to the occluding contour, those situations occur with low probability. These statistics agree nicely with those of our model (Fig. 10).

Fig. 14a shows $\mathbb{V}[\hat{\mathbf{s}}]$ for the rubber duck. The regions of low variance are where the shading flow field is stable; these are most notably the folds and the reflectance edges (the eye of the duck). In regions where the fold is smooth (i.e., the curvature of the surface in the view direction is low), such as the wing of the duck, the shading flow is stable over a greater extent.

As we saw in the previous section, the magnitude of the image gradient is usually largest at the occluding contour and, on the fold side of the contour, decreases as the distance from the contour increases. We define the expected gradient similarly to the expectation of the shading flow field (Eq. (2.15))

$$\mathbb{E}[\nabla I] = \angle_{\times \frac{1}{2}} \left(\int_{\mathbf{l} \in \mathbb{S}^2} \angle_{\times 2}(\nabla I) \|\mathbf{l} \times \hat{\mathbf{z}}\| d\mathbf{l} \right), \quad (2.17)$$

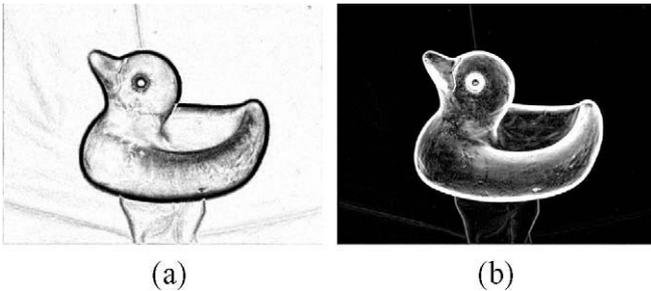


Fig. 14. Statistics for the rubber duck images. (a) Shading flow field variance, $\mathbb{V}[\hat{\mathbf{s}}]$, where dark/bright indicates low/high variance. The shading flow field is clearly stable at folds. The gradient magnitude amplifies the appearance of the folds, as can be seen in (b) the magnitude of the mean image irradiance gradient, $\mathbb{E}[|\nabla I|]$.

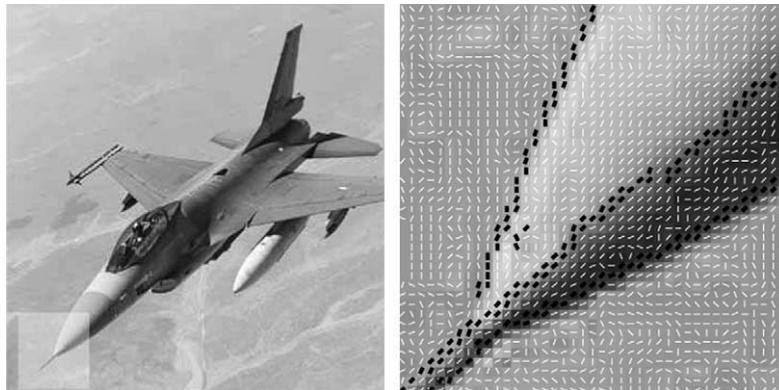


Fig. 15. Local estimate of the shading flow for a single image. The estimate was obtained by evaluating the image gradient over a small neighborhood and rotating; this is equivalent to a linear model for a simple cell, tuned to a low spatial frequency, providing maximum response (over all possible orientations). Notice how noisy and imperfect the local shading flow is. White: local shading flow vectors. Black: local edge vectors.

In Fig. 14b we show $\mathbb{E}[|\nabla I|]$. The folds appear accentuated as the gradient magnitude gives more weight to aligned shading flow vectors, as discussed in Section 2.1. The results suggest that the gradient magnitude is a useful measure of the significance of a particular orientation measurement.

2.2.2. An individual image

The previous analysis showed that information for diagnosing fold-type edges exists in the shading, but when averaged across lighting for an artificial object. In the natural world reflectances are rarely as consistent as for manufactured plastics; and dirt, dust, and detritus may lead to individual variation. We therefore examine a more natural scene in Fig. 15. It is clearly the case that local information does not suffice. Since we do not have access to multiple lighting images to compute averages, we seek to exploit contextual interactions within the single image.

3. Inferring shading flow fields

Given that the initial measurements of a visual flow field may contain spurious or missing values (see Fig. 15), we refine the flow field to counteract these effects. Interpolating and fitting (Rao and Jain, 1992), smoothing (Perona, 1998), and diffusing (Tang et al., 2000) the orientation function $\theta(x, y)$ corresponding to the flow are commonly used approaches to achieving this goal, but they are also prone to affect the underlying geometry of the flow in undesirable ways. In particular, they can distort flow singularities that must be preserved to correctly interpret visual scenes (Ben-Shahar and Zucker, 2003).

To overcome this problem we process the visual flow by enforcing local coherence, that is, by ensuring that each local measurement of the flow field is consistent with its neighboring measurements. Toward this we first examine the differential geometry of visual flows.

3.1. The differential geometry of visual flows

A natural representation of a visual flow which highlights its intrinsic geometry is its frame field (O'Neill, 1966). A local frame $\{E_T, E_N\}$ is attached to each point q of the flow, with E_T tangent and E_N normal to the flow. Small translations in direction V from the point q rotate the frame, a change which is characterized through the covariant derivatives $\nabla_V E_T$ and $\nabla_V E_N$ of the underlying pattern (see Fig. 16). Cartan's connection equation (O'Neill, 1966) expresses these covariant derivatives in terms of the frame field itself:

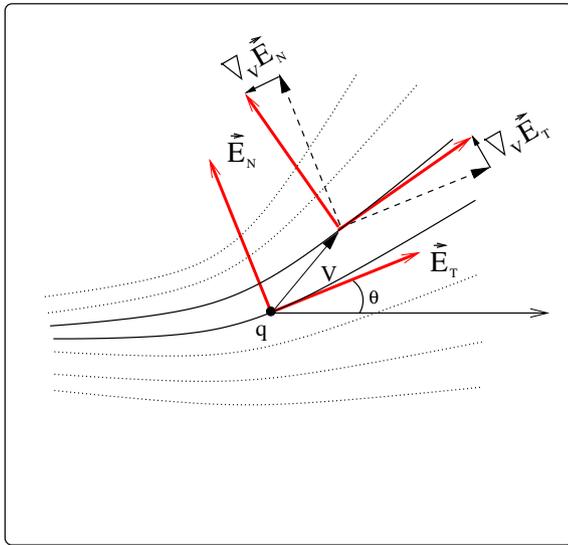


Fig. 16. The frame field representation of visual flows. The local behavior of the frame is described by its covariant derivatives $\nabla_V E_T$ and $\nabla_V E_N$ which are always normal to E_T and E_N , respectively. Since the connection form – the operator that describes the frame’s rotation for any direction V – is linear, it is fully characterized by two numbers computed as projections on two independent directions. In the basis of the frame this yields the curvatures κ_T and κ_N .

$$\begin{pmatrix} \nabla_V E_T \\ \nabla_V E_N \end{pmatrix} = \begin{bmatrix} 0 & w_{12}(V) \\ -w_{12}(V) & 0 \end{bmatrix} \begin{pmatrix} E_T \\ E_N \end{pmatrix}. \quad (3.1)$$

The connection form $w_{12}(V)$ is a linear function of the tangent vector V and can thus be represented by two scalars at each point. In the basis $\{E_T, E_N\}$ these scalars are defined as $\kappa_T \triangleq w_{12}(E_T)$ and $\kappa_N \triangleq w_{12}(E_N)$, which we call the *tangential curvature* and the *normal curvature* – they represent the rate of change of the flow’s dominant orientation in the tangential and normal directions, respectively. In terms of $\theta(x, y)$ and its differential, these curvatures become:

$$\begin{aligned} \kappa_T &= d\theta(E_T) = \nabla\theta \cdot E_T = \nabla\theta \cdot (\cos\theta, \sin\theta), \\ \kappa_N &= d\theta(E_N) = \nabla\theta \cdot E_N = \nabla\theta \cdot (-\sin\theta, \cos\theta). \end{aligned} \quad (3.2)$$

Knowledge of E_T , E_N , κ_T , and κ_N at a point q enables us to construct a local approximation to the flow which has the same orientation and curvatures at q ; we call such an approximation an *osculating flow field*. The osculating flow field is important in that it predicts flow values in the neighborhood of q . Comparing these predictions to the measured flow values indicates how consistent the measured values of the flow at q are with those at its neighbors and suggests how to update them to be consistent. This is analogous to refining curve measurements using cocircularity (Parent et al., 1989; see also Sarti et al., 2008).

There are an infinite number of possible osculating flow fields to choose from. However, there exist criteria for “good” osculating

flow fields. One such criterion is the minimization of the harmonic energy $E[\theta] = \int \int \|\nabla\theta\|^2 dx dy$ associated with the orientation function of the flow, as is used in orientation diffusion (Tang et al., 2000). Viewing the orientation function as a surface in $\mathbb{R}^2 \times \mathcal{S}^1$, however, suggests that the osculating flow field should minimize the surface area $A[\theta] = \int \int \sqrt{1 + \theta_x^2 + \theta_y^2} dx dy$. Finally, the duality of the curvatures κ_T and κ_N suggests that the osculating flow field should exhibit unbiased curvature covariation. Surprisingly, there is a unique osculating flow field which satisfies all of these criteria simultaneously (Ben-Shahar and Zucker, 2003). In the space $\mathbb{R}^2 \times \mathcal{S}^1$ of orientations over the image plane it takes the form of a *right helicoid* (Fig. 17): In particular, we have the following result (Ben-Shahar and Zucker, 2003).

Assume (w.l.o.g) that a visual flow $\theta(x, y)$ satisfies $q = (0, 0)$ and $\theta(q) = 0$, $\kappa_T(q) = K_T$, and $\kappa_N(q) = K_N$. Of all possible orientation functions $\theta(x, y)$ around q that satisfy these constraints, the only one that extremizes both $E[\theta]$ and $A[\theta]$, and has curvature functions that covary *identically* (i.e., $\frac{\kappa_T(x, y)}{\kappa_N(x, y)} = \text{const} = \frac{K_T}{K_N}$) is the right helicoid $\theta(x, y) = \tan^{-1} \left(\frac{K_T x + K_N y}{1 + K_N x - K_T y} \right)$.

With this model of the local structure of visual flow we are in a position to compute a globally coherent flow, the procedure for which is described in the next section.

3.2. Computing coherent visual flows from a single image

The advantage of having a model for the local behavior of visual flow lies in the ability to assess the degree to which a particular measurement is consistent with the context in which it is embedded. This, in turn, can be used to refine noisy measurements, remove spurious ones, and fill in “holes” so that global structures become coherent. A framework in which one can pursue this task by iteratively maximizing the average local consistency over the domain of interest is relaxation labeling (Hummel and Zucker, 1983). We have developed such a network for the organization of coherent visual flows (Ben-Shahar and Zucker, 2003). The following is a short overview of that system.

A direct abstraction of the relaxation process for visual flow should involve an image-based 2D network of nodes $i = (x, y)$ (i.e., pixels) whose labels are drawn from the set

$$\Lambda = \{\text{no-flow}\} \cup \left\{ (\theta, \kappa_T, \kappa_N) \mid \theta \in \left(-\frac{\pi}{2}, \frac{\pi}{2} \right), \kappa_T, \kappa_N \in [-K, K] \right\}$$

after the appropriate quantization. To allow for the representation of either “no-flow” or multiple flows at a pixel, we replace this abstraction with a 5D network of nodes $i = (x, y, \theta, \kappa_T, \kappa_N)$ whose labels are either *TRUE* (T) or *FALSE* (F). For each node i , $p_i(T)$ denotes the probability that a visual flow of orientation θ and curvatures κ_T , κ_N passes through pixel (x, y) . Since $p_i(F) = 1 - p_i(T)$ we need to maintain and update the probability of only one label at each node.

The geometrical compatibilities $r_{ij}(\lambda, \lambda')$ that drive our relaxation process are computed from the osculating flow field as defined

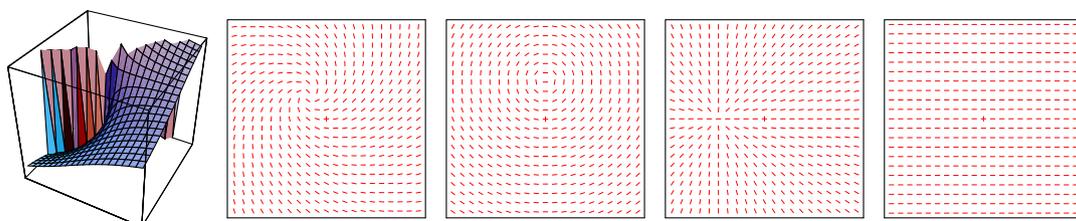


Fig. 17. Examples of right helicoidal visual flows, both in (x, y, θ) space (left) and the image plane. Note that tuning to different curvatures at the origin (point marked with a cross) produces qualitatively different coherent behaviors in its neighborhood.

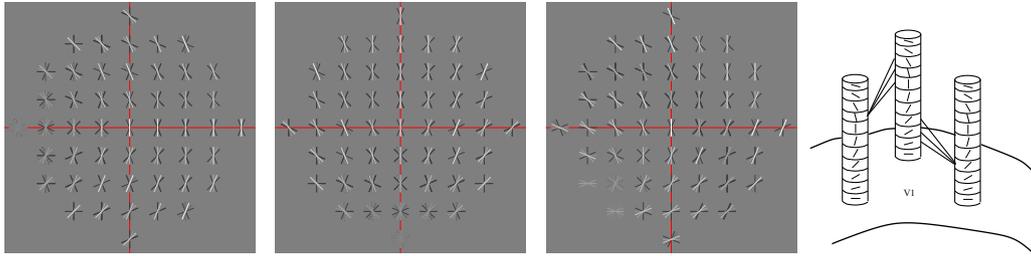


Fig. 18. Examples of compatibility structure (for different values of θ , κ_T and κ_N) projected onto the image plane (brightness represents degree of compatibility, black segments represent an inhibitory surround). As is illustrated on the right, these structures are closely related to long-range horizontal connections between orientation columns in V1 (Ben-Shahar and Zucker, 2003).

by the right helicoid. Measurement quantization implies that every possible node i represents an equivalence class of measurements, each of which induces a field of compatible labels in the neighborhood of i . In the continuum, the union of all these fields forms a consistent 5D “volume” that after quantization results in a set of excitatory labels (see Fig. 18).

With the network structure, labels, and compatibilities all designed, one can compute the support $s_i(\lambda)$ that label λ at node i gathers from its neighborhood. $s_i(\lambda)$ is typically the sum of the individual support of all labels j in the neighborhood of i

$$s_i(\lambda) = \sum_j \sum_{\lambda'} r_{ij}(\lambda, \lambda') p_j(\lambda'). \quad (3.3)$$

Having computed the support for a label, $s_i(\lambda)$ is then used to update the confidence $p_i(\lambda)$ by gradient ascent, followed by non-linear projection. Under the 2-label paradigm and the appropriate weighing of negative (F) versus positive (T) evidence (Ben-Shahar and Zucker, 2003), the projection operator takes a particularly convenient form and the update rule reduces to

$$p_i(\lambda) \leftarrow \Pi_0^1(p_i(\lambda) + \delta s_i(\lambda)), \quad (3.4)$$

where $\Pi_0^1(x)$ projects its operand to the nearest point on the interval $[0, 1]$ and δ is the step size of the gradient descent.

While the relaxation labeling network described is an abstraction to perform an inference based on the differential geometry of flow fields, it is motivated by the architecture of the primary visual cortex. The columnar structure of V1 clearly lends itself to the representation of orientation fields (Hubel and Wiesel, 1977), and

is capable of the necessary curvature computations (Dobbins et al., 1987). Considerable speculation surrounds the functional significance of long-range horizontal connections (Gilbert, 1992) between orientation columns; we posit that they may play a role not unlike the compatibility structures of our network (Fig. 18, right panel). In related work we have shown that, in fact, these compatibility structures are consistent with the population statistics of long-range horizontal connections through second order (Ben-Shahar and Zucker, 2004). Furthermore, the relaxation computation has been reduced to cliques of pyramidal cells (Miller and Zucker, 1992; Miller and Zucker, 1999), a construct recently found by Yoshimura et al. (2005).

3.3. Stability at discontinuities

In computing coherent visual flows it is important to respect its discontinuities, as these often correspond to significant physical phenomena. The relaxation process described above does not destroy these structures because in the high dimensional space in which it operates the flow structures that meet along a line discontinuity, either in orientation or curvature, are separated and thus do not interact. However, without proper tuning, the relaxation process will quickly shrink or expand the flow in the neighborhood of boundaries. It is this behavior we seek to suppress.

To achieve stability we normalize the compatibility function, and thus the support function $s_i(\lambda)$, to account for reduced support in the neighborhood of a discontinuity. Given the compatibility volume V_i which corresponds to a particular node i , we compute

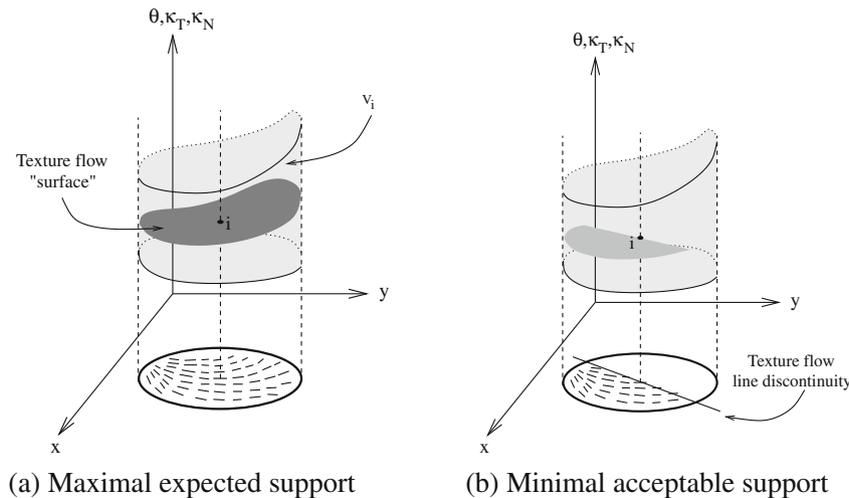


Fig. 19. Practical stability of the relaxation labeling process at line discontinuities in the flow can be achieved through the normalization of the support function. (a) At each node i , s_{max} is determined by integrating the support gathered from a full confidence, compatible flow that traverses the entire compatibility volume V_i . (b) The minimal accepted support s_{min} of a flow of some minimally accepted confidence $\rho_{min} < 1$ (depicted here by the brighter surface intensity) that terminates along a line that intersects i .

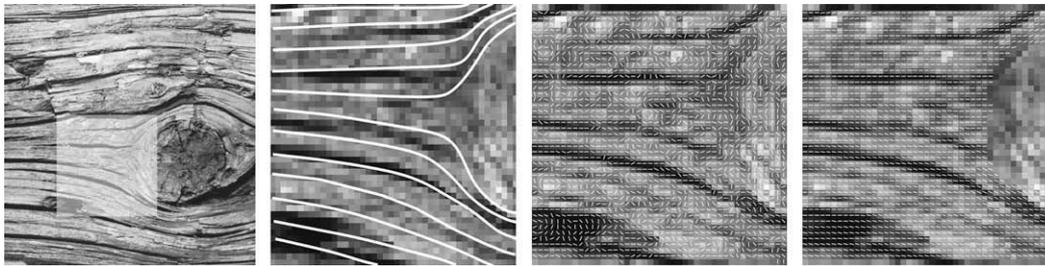


Fig. 20. Visual flow organization based on right helicoidal compatibilities. Shown (left to right) are: Tree bark image and a region of interest (ROI), perceptual structure (drawn manually), initial flow measurements (gradient based filter), and the relaxed visual flow after few iterations of relaxation labeling with the right helicoidal compatibilities. Compare the latter to the perceptual structure and note how the non-flow region was rejected altogether.

the maximal support a node can receive, s_{max} , as the integral of the compatibility coefficients assuming a consistent flow traverses V_i with all supporting nodes at full confidence (Fig. 19). It is clear that the closer i is to a flow discontinuity, the less context supports it. At the discontinuity, the flow should neither grow nor shrink, leading us to define the minimal level of support for which no change in confidence occurs, s_{min} . Observe that s_{min} depends on both the geometry of the discontinuity and the minimally accepted confidence of the supporting nodes. For simplicity we assume the discontinuity (locally) occurs along a straight line. The support from neighboring nodes of minimally accepted average confidence ρ_{min} (Fig. 19) can be approximated as $s_{min} = \frac{\rho_{min}s_{max}}{2}$. Normally ρ_{min} would be set to 0.5, which is the minimal confidence that cannot be disambiguated as the *TRUE* label. In the context of the two-label relaxation labeling paradigm and the gradient ascent update rule (Eq. (3.4)), a decrease in the confidence of a label occurs only if $s_i < 0$. Thus, it remains to normalize the support values by mapping the interval $[s_{min}, s_{max}]$ to the unit interval $[0, 1]$ via the transformation $s_i \leftarrow \frac{s_i - s_{min}}{s_{max} - s_{min}}$ before applying the update rule.

The result of the normalized relaxation process is usually very good (Fig. 20). Nevertheless, the fact that both the support function (Eq. (3.3)) and the normalization are linear creates a delicate balance: while better noise resistance suggests smaller s_{min} , it also implies that at discontinuities the flow will eventually grow uncontrollably. Some solutions to this problem are discussed in Ben-Shahar and Zucker (2003). However, in the case of shading flow fields, discontinuities are intensity edges and thus can be explicitly identified by edge detection. As we discuss below, this information can be directly embedded into the network to decouple the handling of discontinuities from the support normalization.

3.4. Edges as shading flow boundaries

We are now ready to incorporate edges into the shading flow computation. As we discussed earlier, the relationship between the edge and the shading flow field in the neighborhood of the edge can be used to identify the physical cause of the edge, so we need to be careful not to damage that relationship. For example, for a smooth surface the shading flow field is continuous across an edge caused by an abrupt albedo change (Breton and Zucker, 1996). We remain concerned with finding folds, where the shading flow field is generically tangent to the edge.

As discussed earlier, an edge can be thought of as dividing the shading flow field domain into distinct regions, implying that the computation of the shading flow on either side of the edge can and should be done separately. This is an intuitive but powerful argument: incorporating edges into the relaxation labeling network to regulate the growth of flow structure obviates the trade-off between high resistance to noise and strict stability along discontinuities we mentioned in Section 3.2.

To implement this idea we require a specialized set of interactions between edge nodes and nearby shading flow nodes. These interactions would block the flow input if it comes from across the edge. With this input blocked, and so long as s_{min} is positive, the flow on one side of the edge will not extend across the edge, because the total support contributed to the other side will never exceed zero. This frees the selection of s_{min} from stability considerations and allows us to determine it solely on the basis of noise resistance and structural criteria. A cartoon illustrating these interactions appears in Fig. 21. A non-linear veto mechanism that is reminiscent of the one proposed here also exists in the form of shunting inhibition (Borg-Graham et al., 1998).

We have tested this adaptive network on a variety of synthetic and natural images, two of which are shown in Fig. 22. We used the Logical/Linear (Iverson and Zucker, 1995) and the Canny (Canny, 1986) edge detectors and the shading flow fields were measured using standard differential operators.

4. Cusps

Apart from folds, cusps are the only other generic singularity in the mapping of smooth surfaces onto smooth surfaces; all other singularities are killed by small viewer motions. A cusp occurs when the view direction \hat{c} is tangent to the contour generator, i.e. when \hat{c} has second order contact with the surface. At such a point,

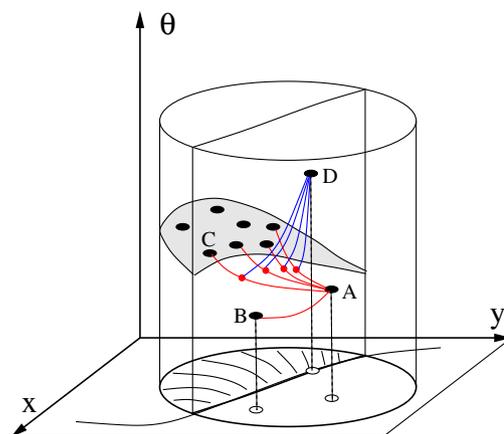


Fig. 21. Edge-flow interactions for boundary stability. Assume the flow structure in the image plane is bounded by the indicated edge. Flow cell A is connected to a set of other cells (B and C) which are a part of the same coherent flow. Although A is not active (there is no flow in its corresponding retinotopic position), its facilitatory interaction with the cells on the other side of the edge may eventually raise its activity level. To prevent cell C from affecting A, an active edge cell D blocks the facilitatory inputs from C, thus effectively limiting A's context to cell B only. Unless enough of these cells are also active, A will not reach its activation potential, and thus will not signal any flow.

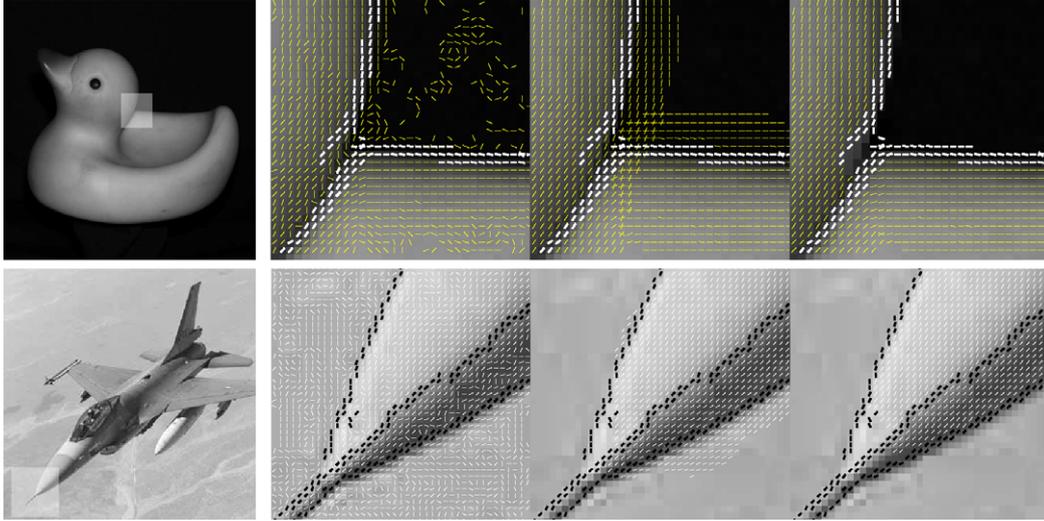


Fig. 22. Examples of shading flow field relaxation with edges as boundary conditions. Shown are (left to right) image and ROI, initial shading flow (thin segments) and edges (thick segments), relaxation without boundaries, and relaxation with boundaries. Note that while both relaxations compute correct flow structure, the one without boundaries extends the flow beyond the edge, making classification more difficult. On the other hand, edge classification and occlusion relationship is trivial based on the result using edges as boundary conditions.

the tangent vector to the contour generator grazes and penetrates the surface. Generically cusps only occur in regions of *non-positive* Gaussian curvature (Koenderink, 1990). To simplify our analysis, we concentrate on those surface patches that are well characterized by a third-order Monge patch description and leave out the mathematical oddities. At a cusp point, then, the view vector lies along an asymptotic line of the surface; there are no asymptotic lines in regions of positive Gaussian curvature.

Formally, cusps are the set of points where

- the mapping $\Pi_{\hat{\mathbf{c}}} \circ \Sigma$ is singular;
- Σ is smooth; and
- $\hat{\mathbf{c}}$ lies along an asymptotic line of Σ .

At a cusp the ideal edge contour ends; before this the contour is a *fold* and, as such, is defined by those points at which the surface normal is orthogonal to the view vector. In general, for smooth surfaces, if a contour ends then, generically, it either ends at a cusp or at a T-junction. On the contour side of the cusp, the usual fold statistics apply. On the other side of the cusp there is no contour in the formal geometric sense. However, the shading flow extends beyond the cusp in a stable manner and this can be characterized. It is related to a curve called the *suggestive contour* (DeCarlo et al., 2003).

4.1. Suggestive contours

Suggestive contours were introduced in computer graphics to improve line drawings for non-realistic rendering (DeCarlo et al., 2003). They share many properties with the ending contours found in artistic drawings and provide, in a sense, the natural extension of the apparent contour beyond the visual cusp. Intuitively they represent “almost contours,” to use a phrase from DeCarlo et al.

The suggestive contour is defined in terms of inflections of curvature in the direction of the view vector projected onto the surface (tangent plane); following DeCarlo we denote the projected view vector \mathbf{w} .¹

Formally, the suggestive contour is the locus of points such that the following two conditions hold. Denote the *radial curvature* by κ_r ;

$$\kappa_r \equiv \mathbf{w} \cdot \nabla_{\mathbf{w}} \hat{\mathbf{n}} = \mathbb{I}(\mathbf{w}, \mathbf{w}) = 0 \quad (4.1)$$

and

$$\nabla_{\mathbf{w}}(\kappa_r) > 0, \quad (4.2)$$

where $\mathbb{I}(\mathbf{u}, \mathbf{v})$ is the second fundamental form.

Equivalently the suggestive contour is all zeros of curvature in the plane spanned by the view vector and surface normal. This implies that the suggestive contour can only exist in regions of non-positive Gaussian curvature, although we focus on regions of negative Gaussian curvature (saddles). In those regions, the suggestive contour consists of all points where \mathbf{w} lies along one of the asymptotic directions of the surface.

An example is shown in Fig. 23. Most importantly, notice how the actual edges and lines computed from the image extend beyond the cusp point and convey similar surface properties as the extended contour.

4.2. Properties of suggestive contours

Suggestive contours, when they exist, have several properties of interest. First, they *extend* contours beyond their formal endpoint (DeCarlo et al., 2003). In fact, ending contours can be defined as the intersection of the contour generator and the suggestive contour. Second, they *anticipate* contours. A surface or viewer rotation can turn a suggestive contour into a cusp. Simply rotate the surface so that the projected view vector \mathbf{w} and the view vector $\hat{\mathbf{c}}$ are the same. This is a rotation of the surface normal about the the $\hat{\mathbf{c}} \times \hat{\mathbf{n}}$ plane. At the cusp, this rotation changes the cusp to a suggestive contour or it is occluded.

In theory, the movement of a cusp under viewer motion or surface rotation can be used to define a chart on the surface. By following them it is possible to characterize shape in a neighborhood around the cusp (Cipolla et al., 1995; Cipolla et al., 1997). However, in practice, it is quite difficult to find cusps precisely in images because cusps extend shading flow stability much like contours. In the case of cusps, this stability leads to *line-like* shading that extends the contour, often obscuring the cusp, and extends the apparent edge well beyond it; see the edge extension in Fig. 23 (right).

¹ There are several definitions for the suggestive contour, for full detail see DeCarlo et al. (2003).

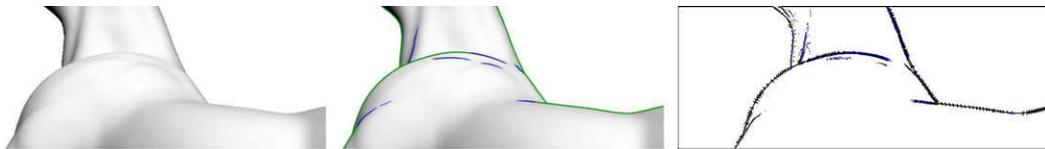


Fig. 23. Illustration of the suggestive contour and its relationship to edges. (left) A Lambertian horse illuminated with a single point light source. (middle) Lambertian horse with contours (green) and suggestive contours (blue). Note how the suggestive contours complement the actual edge contours and result in a more intuitive rendering of the underlying surface structure. (right) Edge detection (black) and line detection (blue) on horse image. Note how they agree with the suggestive contours and extend beyond the cusp. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4.3. Modeling the appearance of cusps

The shading flow exhibits stability around ending contours; i.e., around cusps of the visual map. On one side there is the stability around the fold approaching the cusp; on the other the shading flow is stable along an extension of the contour beyond the cusp. We analyze this by extending the earlier analysis of folds.

Let a surface patch (Σ) be viewed such that it exhibits a cusp. The patch is given by its third-order Taylor expansion at the origin, where the z -axis is the surface normal direction at the origin. Assume orthographic projection and let the view vector ($\hat{\mathbf{v}}$) point in the $-x$ -axis direction ($-\hat{\mathbf{x}}$). This restricts Σ to the following Monge patch form:

$$z = a_1xy + a_2y^2 + b_0x^3 + b_1x^2y + b_2xy^2 + b_3y^3 + O((x,y)^4). \quad (4.3)$$

As before let the surface be Lambertian with albedo of unity, and let the light source be a point source at infinity in the direction \mathbf{l} . Then the covariant derivative of the radiance on the surface is

$$\nabla_{\mathbf{w}}I = \nabla_{\mathbf{w}}(\hat{\mathbf{n}} \cdot \hat{\mathbf{l}}) = (\nabla_{\mathbf{w}}\hat{\mathbf{n}}) \cdot (\mathbf{l}_t + \mathbf{l}_n) + (\nabla_{\mathbf{w}}\hat{\mathbf{l}}) \cdot \hat{\mathbf{n}} = \mathbb{I}(\mathbf{w}, \mathbf{l}_t),$$

where \mathbf{l} is decomposed into \mathbf{l}_t in the tangent plane and \mathbf{l}_n which lies along the normal.

To find the gradient, as an orthogonal basis on the surface we choose the suggested view vector and any vector orthogonal to it. Therefore the surface gradient can be written as follows:

$$\nabla I = (\nabla_{\mathbf{w}}I)\hat{\mathbf{w}} + (\nabla_{\mathbf{w}\perp}I)\hat{\mathbf{w}}_{\perp}. \quad (4.4)$$

From the surface gradient, the statistical analysis of shading flow over lighting positions can be followed. However, since this stability manifests itself as line and edge-like shading, it is also possible to look at the stability of edge detectors beyond the cusp; see Figs. 24 and 25.

4.4. Motion of cusps under viewer motion

Both the contour generator and suggestive contour “slide” over the surface under changes in viewer position (Cipolla et al., 1995, 1997). From the motion of cusps it is possible to determine local surface information. However, this relies on the ability to accurately find the cusp, a task that we believe to be quite difficult. Any viewer motion can be broken into a component that leaves

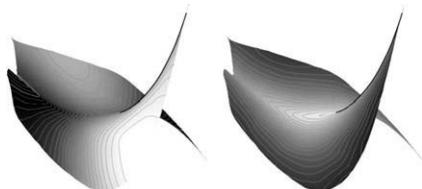


Fig. 24. Appearance of a cusp under different lighting conditions. While the surface geometry remains the same, the image appearance can shift drastically with changes in light position. Nevertheless, notice the relative stability of the ending contour and the isophotes around it.

$\hat{\mathbf{n}} \cdot \hat{\mathbf{c}}$ unchanged, and a component that leaves $\hat{\mathbf{w}} \cdot \nabla_{\mathbf{w}}\hat{\mathbf{n}}$ unchanged, plus an essentially irrelevant rotation. These are orthogonal viewer motions, and they represent rotations of the surface about $\hat{\mathbf{n}}$ and $\hat{\mathbf{n}} \times \hat{\mathbf{c}}$, respectively. Therefore the cusp slides along the (moving) suggestive contour and contour generator. It is possible to use this motion to set up a coordinate system on the surface, allowing shape information to be derived from known viewer motion. However, we are interested in the shape information contained in a static image.

It can be shown that the velocity of the cusp along the suggestive contour under viewer motion is proportional to $\frac{1}{\sqrt{K}}$ where K is the Gaussian curvature (Cipolla et al., 1995). However, at the cusp point it can also be shown that the derivative of the foreshortening along the suggestive contour is proportional to \sqrt{K} .² In regions of high foreshortening, small changes in viewing position will rapidly move the cusp along the contour. In other words, if small viewer motion would make a point a contour, it will look like a contour.

4.5. Cusps under special lighting conditions

There is one special lighting condition under which the shading flow is exceptionally well behaved. For a Lambertian surface and a point light source at infinity, the intensity at a given point is given by $I(p) = \hat{\mathbf{n}} \cdot \hat{\mathbf{l}}$ where $\hat{\mathbf{n}}$ is the surface normal and $\hat{\mathbf{l}}$ is a vector pointing along the light path. We let $\mathbf{w} = \text{proj}(\hat{\mathbf{c}})$ and choose the $(\hat{\mathbf{w}}, \hat{\mathbf{w}}_{\perp}, \hat{\mathbf{n}})$ frame. Then $\hat{\mathbf{l}} = (l_w, l_{w\perp}, l_n)$ and we calculate:

$$\begin{aligned} \nabla_{\mathbf{w}}I &= \nabla_{\mathbf{w}}(\hat{\mathbf{n}} \cdot \hat{\mathbf{l}}) = (\nabla_{\mathbf{w}}\hat{\mathbf{n}}) \cdot (l_n\hat{\mathbf{n}} + l_w\hat{\mathbf{w}} + l_{w\perp}\hat{\mathbf{w}}_{\perp}) + (\nabla_{\mathbf{w}}\hat{\mathbf{l}}) \cdot \hat{\mathbf{n}} \\ &= l_w\mathbb{I}(\mathbf{w}, \mathbf{w}) + l_{w\perp}\mathbb{I}(\mathbf{w}, \mathbf{w}_{\perp}) = l_w\kappa + l_{w\perp}\tau, \end{aligned} \quad (4.5)$$

where $\kappa = \mathbb{I}(\mathbf{w}, \mathbf{w})$ and $\tau = \mathbb{I}(\mathbf{w}, \mathbf{w}_{\perp})$. Following (DeCarlo et al., 2003) we call κ the *radial curvature* and τ the *geodesic torsion*. This radial curvature is an extension of Koenderink’s notion of radial curvature along a contour (Koenderink, 1984). Therefore when the light source lines up with \mathbf{w} then it is clear that $\nabla_{\mathbf{w}}I = 0$ when $\kappa = 0$ which is along the suggestive contour. Since the suggestive contour is an inflection of curvature in the direction of maximum foreshortening, this corresponds to a bright or dark peak in the intensity. Therefore we would expect to see a bright or dark line response that extends the contour. Note that \mathbf{l}_n is irrelevant for the covariant derivative of the shading flow. DeCarlo et al. exploit this property in their image space algorithm for finding suggestive contours (DeCarlo et al., 2003).

4.6. Cusps under general lighting conditions

For general lighting conditions the shading flow is not as well behaved. Under general lighting conditions, the suggested contour is not a maximum or minimum of brightness. However, the behavior of isophotes around the suggested contour is dominated by the foreshortening effect. The shading flow is stable in much the same fashion that it is stable around a contour, and this stability extends

² (5.5), where $a_1 = \sqrt{K}$.

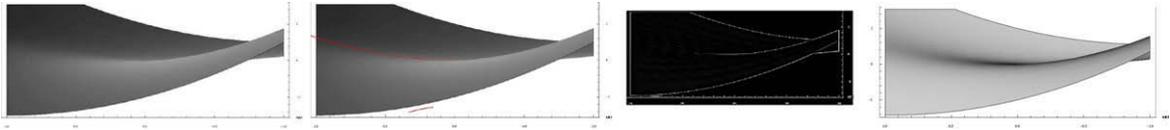


Fig. 25. Estimation of the probability of an edge over all lighting positions. Panels left to right: 1. Original surface patch with the light at a typical position. The cusp is located at 0.0 on the x-axis. 2. The suggestive contour in red. Notice that it begins at the cusp. 3. The probability of an edge at any location, calculated as before. Notice how the edge probability extends well beyond the cusp point. 4. Variance of surface vector orientation in image plane under orthographic projection. Notice the stability near the edges and beyond the cusp at $x = 0$.

beyond the cusp. This is the sense in which the suggestive contour is “almost” a contour.

Along with the shading flow stability the foreshortening boosts the gradient across the suggested contour. Every small patch of surface is compressed along the projection of the view vector onto the surface. The isophotes are surface curves, so they are similarly compressed. This boosts the gradient in the projection of the view vector, or equivalently, in the direction of tilt as the slant increases (Stevens, 1983). In the special case above, a maximum or minimum of brightness occurred along the suggested contour. This was compressed into a line-like image feature. In general, both edge-like and line-like features occur; which feature appears depends on the lighting direction.

5. Curvature of suggestive contour and gaussian curvature

The suggestive contour in the image plane provides useful information about the surface. First, for a cusp to arise the contour generator must be tangent to the view vector. The view vector then has third-order contact with the surface, which cannot occur in a region of positive Gaussian curvature. We focus on those neighborhoods with negative Gaussian curvature since, in these cases, the suggestive contour extends the (fold) edge contour. Furthermore, such suggestive contours often have a line-like appearance, since the shading flow is compressed in a neighborhood around it. Viewed through finitely-sized receptive fields, the suggestive contour is thus often indistinguishable from physical edges or lines. This raises the basic question of whether suggestive contours provides similar shape information to the physical edge contour.

We know, from a theorem of Koenderink (1984), that the sign of curvature of an apparent edge contour will give the sign of Gaussian curvature. We now show that the curvature of the suggestive contour varies with the derivative of Gaussian curvature across the suggestive contour. Since the suggestive contour tends to lie near parabolic lines (DeCarlo et al., 2003), this provides very similar information.

Assume orthographic projection (this assumption can be generalized). Our surface around the cusp can be locally parameterized as $\mathbf{r}(x, z) = (x, y, z(x, y))$ where $z(x, y) = a_1xy + a_2y^2 + b_0x^3 + b_1x^2y + b_2xy^2 + b_3y^3 + O((x, y)^4)$.³ As before assume a view vector of $-\hat{\mathbf{x}}$ (viewer at $+\infty$). Under this setup and at the cusp the second fundamental form becomes simply

$$\mathbb{I}(\mathbf{u}, \mathbf{v}) = [u_x u_{ry}] \begin{pmatrix} 0 & a_1 \\ a_1 & 2a_2 \end{pmatrix} \begin{bmatrix} v_{rx} \\ v_{ry} \end{bmatrix},$$

where \mathbf{u} and \mathbf{v} are vectors in the tangent plane, in terms of the basis $\mathbf{r}_x, \mathbf{r}_y$. Let the suggestive contour be the curve $\mathbf{s}(t)$ for some parameterization. We want to find $\frac{\dot{\mathbf{s}}(0)}{\|\dot{\mathbf{s}}(0)\|}$.

The suggested contour is defined as the locus of points

$$L = \{\mathbf{p} \in \Sigma : \mathbb{I}(\mathbf{w}_p, \mathbf{w}_p) = 0 \text{ and } \nabla_{w_p} \kappa_r > 0\},$$

where $\mathbf{w}_p = \hat{\mathbf{c}} - (\hat{\mathbf{c}} \cdot \hat{\mathbf{n}}_p)\hat{\mathbf{n}}_p$. In our Monge form at the cusp⁴

$$\nabla_{w_p} \kappa_r = \nabla_{w_p} \mathbb{I}(\mathbf{w}_p, \mathbf{w}_p) = 6b_0. \quad (5.1)$$

Since \mathbf{w} is a vector in the tangent plane it can be written in terms of $(\mathbf{r}_x, \mathbf{r}_y)$. In this basis the shape operator can be written as

$$\mathbb{I} = \begin{bmatrix} \mathbf{r}_{xx} \cdot \hat{\mathbf{n}} & \mathbf{r}_{xy} \cdot \hat{\mathbf{n}} \\ \mathbf{r}_{xy} \cdot \hat{\mathbf{n}} & \mathbf{r}_{yy} \cdot \hat{\mathbf{n}} \end{bmatrix}. \quad (5.2)$$

Therefore L is the locus of points such that:

$$[w_{r_x} w_{r_y}] \begin{bmatrix} \mathbf{r}_{xx} \cdot \hat{\mathbf{n}} & \mathbf{r}_{xy} \cdot \hat{\mathbf{n}} \\ \mathbf{r}_{xy} \cdot \hat{\mathbf{n}} & \mathbf{r}_{yy} \cdot \hat{\mathbf{n}} \end{bmatrix} \begin{bmatrix} w_{r_x} \\ w_{r_y} \end{bmatrix} = 0. \quad (5.3)$$

We can then use the implicit function theorem to take implicit derivatives of this equation with respect to y , and then solve at the origin. We find that at the origin, the tangent vector $\dot{\mathbf{s}}(0) \propto \begin{bmatrix} -\frac{b_1}{3b_0} & 1 \end{bmatrix} = \mathbf{t}$ in the basis r_x, r_y . Then:

$$\mathbb{I}(\mathbf{t}) = \begin{bmatrix} 0 & a_1 \\ a_1 & 2a_2 \end{bmatrix} \begin{bmatrix} -\frac{b_1}{3b_0} \\ 1 \end{bmatrix}, \quad (5.4)$$

$$= \begin{bmatrix} a_1 \\ 2a_2 - \frac{a_1 b_1}{3b_0} \end{bmatrix}. \quad (5.5)$$

Then the curvature in the image plane is

$$\mathbb{I}(\mathbf{t}, \mathbf{t}) = \begin{bmatrix} -\frac{b_1}{3b_0} & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ 2a_2 - \frac{a_1 b_1}{3b_0} \end{bmatrix} = 2a_2 - \frac{2a_1 b_1}{3b_0}. \quad (5.6)$$

We stress that \mathbf{t} is not a unit vector on the surface. Rather, we are interested in a unit step in the image plane.

We can therefore identify the first term of (5.5) with a rotation of the normal toward the viewer. Therefore the decrease in foreshortening is proportional to the Gaussian curvature. It can also be shown the curvature of the suggested contour in the image plane is related to the derivative of the Gaussian curvature in the viewer direction. In fact, from (5.3) and (4.3) one can show that

$$\nabla_{\hat{\mathbf{c}}} K = 12a_2 b_0 - 4a_1 b_1. \quad (5.7a)$$

By comparing (5.1), (5.6) and (5.7a) it is clear that

$$\kappa_{app} * \nabla_{\hat{\mathbf{c}}} \kappa_r = \left(2a_2 - \frac{2a_1 b_1}{3b_0}\right) * 6b_0, \quad (5.7b)$$

$$\nabla_{\hat{\mathbf{c}}} K = \kappa_{app} * \nabla_{\hat{\mathbf{c}}} \kappa_r. \quad (5.7c)$$

Since $\nabla_{\hat{\mathbf{c}}} \kappa_r > 0$ it is possible to determine the sign of $\nabla_{\hat{\mathbf{c}}} K$ from the apparent curvature of the suggested contour. This result can be extended along the entire suggested contour by taking the third-order Taylor approximation along the suggested contour.⁵

For illustration, see Figs. 26 and 27.

³ We can always rotate the surface to kill off the a_0x^2 term. We must do so to bring one asymptotic line of the surface in line with the x-axis, where we assume the view vector lies. One could choose to line up the other asymptotic line, but it needlessly complicates the analysis.

⁴ Note this puts a further restriction on the form of our Monge patch, $b_0 > 0$.

⁵ This will add a x term to the Taylor approximation to the cusp. This makes the calculations significantly more unpleasant but (5.7c) holds.

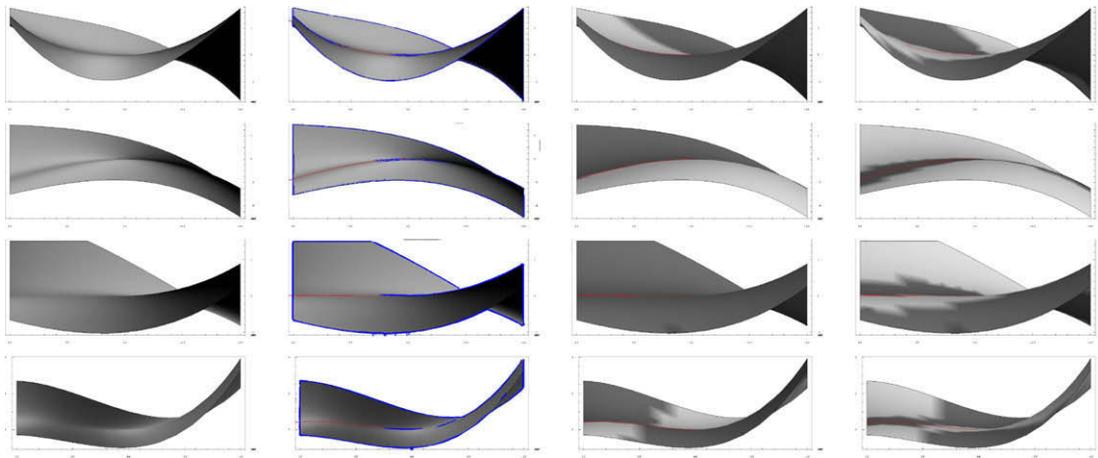


Fig. 26. Several cusps shown with (columns starting at left and proceeding right) Lambertian lighting, suggestive contour (in red), w/Logical/Linear edge detection (in blue), Gaussian curvature (K) and $\nabla_w K$. The suggestive contour has positive, negative, and zero apparent curvature in the first three rows, while the last row of images has an inflection point. Notice that $\kappa_{app} > 0 \rightarrow \nabla_w K > 0$. Notation: In this last two columns, dark intensity indicates negative and light intensity positive values. Quantization artifacts courtesy of Mathematica, v.6. (For interpretation of the references to colours in this figure legend, the reader is referred to the web version of this paper.)

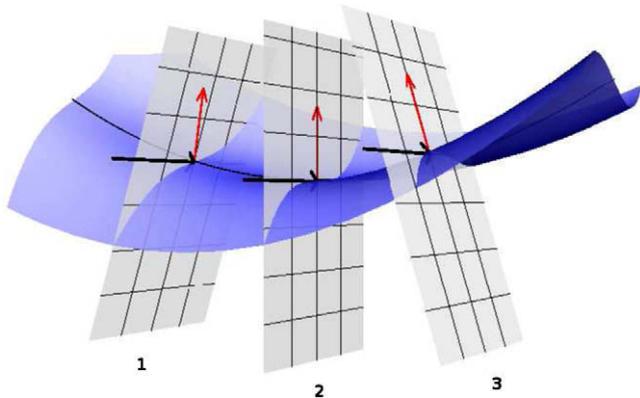


Fig. 27. The view-vector/normal (\hat{c}, \hat{n}) plane along the suggestive contour (1), at the cusp (2), and along the contour (3). Notice in (1) that the view vector is not orthogonal to the surface normal. The suggestive contour runs along the inflection points of the plane/surface intersections.

6. Connections to contour theorems

The suggested contour is connected to two theorems of Koenderink. The first (and most famous) relates the apparent curvature of the contour to the Gaussian curvature at the contour. In orthographic projection the theorem reduces to

$$K = \kappa_r * \kappa_{app}, \quad (6.1)$$

where K is the Gaussian curvature, κ_r is the surface curvature in the view vector–normal plane, and κ_{app} is the curvature of the projection of the contour into the image plane. Note that our definition of the radial curvature agrees with this definition along the contour, but extends it beyond. One could take the directional derivative of the Koenderink result along the projected view vector to derive the following:

$$\nabla_{\hat{c}} K = (\nabla_{\hat{c}} \kappa_r) \kappa_{app} + \kappa_r \nabla_{\hat{c}} \kappa_{app}. \quad (6.2)$$

However, this is not particularly enlightening along the contour, as only κ_{app} is clearly accessible from an image. Naively extending this result to the suggestive contour is significantly more instructive as $\kappa_r = 0$. Then (6.2) and (5.7b) take the same form, with κ_{app} representing the curvature of the contour and suggestive con-

tour in the image plane. This provides the strongest evidence that the suggestive contour is in a sense the most natural extension of the contour.

One consequence of this result is that the sign of curvature of the contour gives the sign of Gaussian curvature along the contour generator. We found that the curvature of the suggestive contour gives information about the derivative of Gaussian curvature across the cusp. In addition, since the suggestive contour hugs the parabolic lines, the derivative of Gaussian curvature very nearly gives the sign change across the parabolic lines. Therefore the suggestive contour holds similar information to the contour.

A corollary result of Koenderink states that contours always end in a convexity (Koenderink and van Doorn, 1982). This is a consequence of contours ending in regions of negative Gaussian curvature. Locally this is a hyperbolic patch, so contours adopt the convexity of the patch. The suggestive contour shares no such restriction. Since the suggestive contour continues the contour in a smooth (C^1) fashion, it extends the actual contour beyond the cusp visually; nearby cusps it follows line-like features in the projected shading. It is therefore a very plausible explanation for why artists choose to draw “impossible” ending contours. These contours accurately reflect the way that contours appear to end. Furthermore, they contain important curvature information. It is easy to draw ending contours that have the “wrong”, or mathematically impossible curvature, but appear correct and reasonable. In fact, the “correct” curvature can seem quite odd, or can drastically change the apparent shape. Under these circumstances, the suggested contour curvature seems to match our intuitive notion of the curvature of ending contours, particularly where the suggestive contour is near the parabolic lines and therefore extends quite far.

6.1. Revisiting the psychophysics

In Fig. 3 we showed several examples of human shape estimates from different figures. With the material on cusps, the apparent extension of boundaries, and the curvature estimates along the suggestive contour, we are now in a position to revisit those data.

In Fig. 3a and b were two instances of a shape under different lighting configurations; in both cases all subjects extended the hyperbolic region well beyond the cusp. This is consistent with both the probability that an edge will be extended beyond a cusp (Fig. 25) and the computed edges (Fig. 23). It remains to make these predictions quantitative but for now their qualitative agreement can be seen in Fig. 28.

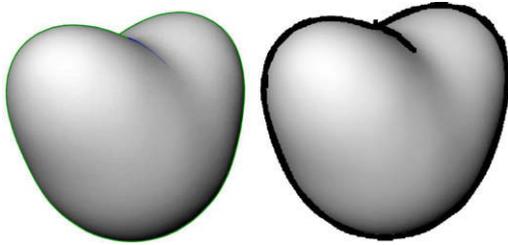


Fig. 28. Evaluating how contours influence shape. (left) The furrow image with true edges (green) and a length of suggestive contour (blue). The cusp point is where they meet. (right) Detected edges (black) superimposed on the furrow shape. Notice how it is extended well beyond the cusp. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

7. Conclusions

Contours and shading both carry information about shape, and integrating them into a coherent representation is a basic step in understanding their neural implementation. In this paper we described three aspects of this integration. First, regarding qualitative shape features, we defined folds and cuts in terms of how the shading flow field approached the boundary. This analysis was generic mathematically, and was shown to hold in a statistical sense over different lighting configurations. However for vision applications only single images are available, and local estimates of both boundaries and shading are notoriously noisy. Therefore our second contribution was in the inference of consistent shading flow fields, at which a second interaction between shading and boundaries emerged: edges serve as the boundary conditions for limiting the growth of shading (and other surface) properties. Finally, for many lobed shapes viewed from many viewpoints, cusp configurations arise in the surface and exterior edges move interior to the shape and end. This implies a fold configuration on one side of the edge, and a cut on the other; and the shading flow field has a singularity right there. We show how apparent contours arise there as well, often extending the edge beyond the precise cusp and suggesting further details of the shape.

Local orientation representations coupled through curvature-based connections underlie all of these computations. In this sense they are consistent with neural structure in several retinotopically-organized visual areas, which makes it difficult to predict their precise neural substrate. They may well be distributed across a number of visual areas. But our models do explain psychophysical data on the perception of surfaces, which encourages us to believe that their neural realization will eventually be identified.

Acknowledgements

We thank D. DeCarlo for the use of his program, *rtsc*, which was used to generate several figures; H. Chen and P. Belhumeur for participation in earlier stages of this research; and J. Koenderink and A. van Doorn for conversations. Portions of this material were presented at the following meetings and appeared in their proceedings: IEEE Conference on Computer Vision and Pattern Recognition, 2001; the Eighth International Conference on Computer Vision, 2001; The Fourth International Workshop on Visual Form, 2001; Workshop on Biologically Motivated Computer Vision, 2002; and the Stockholm Workshop on Computational Vision, 2008. They are brought together here for the first time. Research supported by AFOSR, ARO, DARPA, NGA, ONR and NSF. OBS is partially funded by the Psychobiology Young Investigator grant 207-07-08, the Israel Science Foundation (ISF) Grant No. 1245/08 and the generous support of the Frankel fund, the Paul Ivanier cen-

ter for Robotics Research and the Zlotowski Center for Neuroscience at Ben-Gurion University.

References

- Bakin, Jonathan S., Nakayama, Ken, Gilbert, Charles D., 2000. Visual responses in monkey areas V1 and V2 to three-dimensional surface configurations. *Journal of Neuroscience* 20 (21), 8188–8198.
- Barrow, H.G., Tenenbaum, J.M., 1981. Interpreting line drawings as three-dimensional surfaces. *Artificial Intelligence* 17 (1–3), 75–116.
- Ben-Shahar, O., Zucker, S.W., 2003. Geometrical computations explain projection patterns of long-range horizontal connections in visual cortex. *Neural Computation* 16, 445–476.
- Ben-Shahar, O., Zucker, S.W., 2003. The perceptual organization of texture flows: a contextual inference approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25 (4), 401–417.
- Ben-Shahar, O., Zucker, S.W., 2004. Geometrical computations explain projection patterns of long-range horizontal connections in visual cortex. *Neural Computation* 16 (3), 445–476.
- Ben-Shahar, O., Huggins, P.S., Izo, T., Zucker, S.W., 2003. Cortical connections and early visual function: intra- and inter-columnar processing. *Journal of Physiology (Paris)* 97, 191–208.
- Borg-Graham, L.J., Monier, C., Frégnac, Y., 1998. Visual input evokes transient and strong shunting inhibition in visual cortical neurons. *Nature* 393, 369–373.
- Bradley, A., Skottun, B.C., Ohzawa, I., Sclar, G., Freeman, R.D., 1987. Visual orientation and spatial frequency discrimination: a comparison of single neurons and behavior. *Journal of Neurophysiology* 57 (3), 755–772.
- Breton, P., Zucker, S.W., 1996. Shadows and shading flow fields. *Computer Vision and Pattern Recognition*, 782–789.
- Canny, J., 1986. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 8 (6), 679–698.
- Caplovitz, G.P., Tse, P.U., 2007. V3A processes contour curvature as a trackable feature for the perception of rotational motion. *Cerebral Cortex* 17 (5), 1179.
- Cavanagh, P., Leclerc, Y., 1989. Shape from shadows. *Journal of Experimental Psychology: Human Perception* 16 (4), 3–27.
- Cipolla, R., Fletcher, G., Giblin, P., 1995. Surface geometry from cusps of apparent contours. In: *Proceedings of the Fifth International Conference on Computer Vision*, p. 858.
- Cipolla, R., Fletcher, G., Giblin, P., 1997. Following cusps. *International Journal of Computer Vision* 23 (2), 115–129.
- Craft, Edward, Schutze, Hartmut, Niebur, Ernst, Heydt, Rudiger von der, 2007. A Neural Model of Figure–Ground Organization. *Journal of Neurophysiology* 97 (6), 4310–4326.
- DeCarlo, D., Finkelstein, A., Rusinkiewicz, S., Santella, A., 2003. Suggestive contours for conveying shape. *ACM Transactions on Graphics* 22 (3), 848–855.
- DeCarlo, D., Finkelstein, A., Rusinkiewicz, S., 2004. Interactive rendering of suggestive contours with temporal coherence. In: *Proceedings of the 3rd International Symposium on Non-photorealistic Animation and Rendering*, pp. 15–145.
- Dobbins, A., Zucker, S.W., Cynader, M.S., 1987. Endstopped neurons in the visual cortex as a substrate for calculating curvature. *Nature* 329 (6138), 438–441.
- Dufour, J.P., 1983. Familles de courbes planes différentiables. *Topology* 4, 449–474.
- Freeman, W.T., 1994. The generic viewpoint assumption in a framework for visual perception. *Nature* 368, 542–545.
- Garding, Jonas, 1992. Shape from texture for smooth curved surfaces in perspective projection. *Journal of Mathematical Imaging and Vision* 2, 630–638.
- Geisler, W.S., Albrecht, D.G., 1997. Visual cortex neurons in monkeys and cats: detection discrimination and identification. *Visual Neuroscience (Print)* 14 (5), 897–919.
- Gepshtein, Sergei, Banks, Martin S., 2003. Viewing geometry determines how vision and haptics combine in size perception. *Current Biology* 13 (13), 483–488.
- Gilbert, C.D., 1992. Horizontal integration and cortical dynamics. *Neuron* 9, 1–13.
- Horn, Berthold, 1975. *Obtaining Shape from Shading*. McGraw-Hill.
- Horn, B., Brooks, M.J., 1989. *Shape From Shading*. MIT Press.
- Hubel, D., Wiesel, T., 1977. Functional architecture of macaque monkey visual cortex. *Proceedings of the Royal Society of London Series B* 198, 1–59.
- Huggins, P., Zucker, S.W., 2001. How folds cut a scene. In: *Proc. Fourth Int. Workshop on Visual Form*, Capri.
- Huggins, Patrick, Zucker, Steven W., 2001. Folds and cuts: how shading flows into edges. In: *Eighth International Conference on Computer Vision*. Computer Society Press, Washington, DC, pp. 153–158.
- Hummel, R.A., Zucker, S.W., 1983. On the foundations of the relaxation labeling processes. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 5, 267–287.
- Issa, N.P., Trepel, C., Stryker, M.P., 2000. Spatial frequency maps in cat visual cortex. *Journal of Neuroscience* 20 (22), 8504.
- Iverson, L.A., Zucker, A.W., 1995. Logical/linear operators for image curves. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 17 (10), 982–996.
- Kang, Byung-Geun, Koenderink, J.J., 2007. Shape from shading from images rendered with various surface types and light fields. *Perception* 36 (8), 1191.
- Kersten, Daniel, Schrater, Paul, 2002. Pattern inference theory: a probabilistic approach to vision. *Perception and the Physical World*.
- Knill, David C., Mamassian, Pascal, Kersten, Daniel, 1997. Geometry of shadows. *Journal of the Optical Society of America A* 14, 3216–3232.

- Koenderink, J.J., 1984. What does the occluding contour tell us about solid shape. *Perception* 13 (3), 321–330.
- Koenderink, J.J., 1990. *Solid Shape*. MIT Press Cambridge, MA, USA.
- Koenderink, Jan J., van Doorn, A.J., 1976. Singularities of the visual mapping. *Biological Cybernetics* 24, 51–59.
- Koenderink, J.J., van Doorn, A.J., 1982. The shape of smooth objects and the way contours end. *Perception* 11 (2), 129–137.
- Koenderink, Jan J., van Doorn, A.J., 1993. Local features of smooth shapes: ridges and courses. *Proceedings of SPIE* 2031, 2–13.
- Koffka, K., 1935. *Principles of Gestalt Psychology*. Harcourt, Brace & World.
- Kutulakos, Kiriakos N., Dyer, Charles R., 1994. Recovering shape by purposive viewpoint adjustment. *International Journal of Computer Vision* 12, 113–136.
- Lamme, V.A.F., 1995. The neurophysiology of figure ground segregation in primary visual cortex. *Journal of Neuroscience* 15, 1605–1615.
- Lee, T.S., Mumford, D., Romeo, R., Lamme, V.A.F., 1998. The role of the primary visual cortex in higher level vision. *Vision Research* 38, 2429–2454.
- Lehky, Sidney R., Sejnowski, Terrence J., 1988. Network model of shape-from-shading: neural function arises from both receptive and projective fields. *Nature* 333 (2), 452–454.
- Li, G., Zucker, S.W., 2006. Surface geometric constraints for stereo in belief propagation. In: 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 2.
- Malik, J., 1987. Interpreting line drawings of curved objects. *International Journal of Computer Vision* 1.
- Miller, D.A., Zucker, S.W., 1992. Efficient simplex-like methods for equilibria of nonsymmetric analog networks. *Neural Computation* 4, 167–190.
- Miller, D.A., Zucker, S.W., 1999. Computing with self-excitatory cliques: a model and an application to hyperacuity-scale computation in visual cortex. *Neural Computation* 11, 21–66.
- Mingolla, E., Todd, J.T., 1986. Perception of solid shape from shading. *Biological Cybernetics* 53 (3), 137–151.
- Oliensis, J., 1991. Uniqueness in shape from shading. *International Journal of Computer Vision* 6 (2), 75–104.
- O'Neill, B., 1966. *Elementary Differential Geometry*. Academic Press.
- Orban, G.A., 2008. Higher order visual processing in macaque extrastriate cortex. *Physiological Reviews* 88 (1), 59–89.
- Palmer, S.E., Ghose, T., 2008. Extremal edges: a powerful cue to depth perception and figure-ground organization. *Psychological Science* 19 (1), 77–84.
- Parent, P., Zucker, S.W., 1989. Trace inference, curvature consistency, and curve detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11 (8), 823–839.
- Pasupathy, A., Connor, C.E., 2002. Population coding of shape in area V4. *Nature Neuroscience* 5 (12), 1332–1338.
- Pentland, Alex, 1984. Local shading analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 6, 170–187. March.
- Perona, P., 1998. Orientation diffusion. *IEEE Transactions on Image Processing* 7 (3), 457–467.
- Qiu, F., von der Heydt, R., 2005. Figure and ground in the visual cortex: V2 combines stereoscopic cues with gestalt rules. *Neuron* 47, 155–166.
- Ramachandran, V.S., 1988. Perception of shape from shading. *Nature* 331 (6152), 163–166.
- Rao, A.A., Jain, R.C., 1992. Computerized flow field analysis: oriented texture fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 17 (7), 693–709.
- Rieger, J.H., 1990. The geometry of view space of opaque objects bounded by smooth surfaces. *Artificial Intelligence* 44, 1–40.
- Rubin, E., 1915. *Synsoplevede Figurer*, Denmark.
- Sakai, Ko, Nishimura, Haruka, 2004. Determination of border ownership based on the surround context of contrast. *Neurocomputing* 58, 843–848.
- Sakai, Ko, Nishimura, Haruka, 2006. Surrounding suppression and facilitation in the determination of border ownership. *Journal of Cognitive Neuroscience* 18 (4), 562–579.
- Sarti, A., Citti, G., Petitot, J., 2008. The symplectic structure of the primary visual cortex. *Biological Cybernetics* 98 (1), 33–48.
- Silverman, M.S., Grosf, D.H., Valois, R.L.D., Elfar, S.D., 1989. spatial-frequency organization in Primate Striate Cortex. *Proceedings of the National Academy of Sciences* 86 (2), 711–715.
- Stevens, K.A., 1983. Slant-tilt: the visual encoding of surface orientation. *Biological Cybernetics* 46 (3), 183–195.
- Tang, B., Sapiro, G., Caselles, V., 2000. Diffusion of general data on non-flat manifolds via harmonic maps theory: the direction diffusion case. *International Journal of Computer Vision* 36 (2), 149–161.
- Todd, J.T., Mingolla, E., 1983. Perception of surface curvature and direction of illumination from patterns of shading. *Journal of Experimental Psychology: Human Perception Performance* 9 (4), 583–595.
- Waltz, D., 1975. *Understanding Line Drawings of Scenes with Shadows*. McGraw-Hill.
- Whitney, H., 1955. On singularities of mappings of euclidean spaces i, mappings of the plane into the plane. *Annals of Mathematics* 62, 374–410.
- Witkin, A., 1982. Intensity-based edge classification. *AAAI* 82 (1), 36–41.
- Yoshimura, Y., Dantzker, J.L.M., Callaway, E.M., 2005. Excitatory cortical neurons form fine-scale functional networks. *Nature* 433, 868–873.
- Zhang, W., Bergholm, F., 1997. Multiscale blur estimation and edge type classification for scene analysis. *IJCV* 24, 219–250.
- Zhou, H., Friedman, H.S., von der Heydt, R., 2000. Coding of border ownership in monkey visual cortex. *Journal of Neuroscience* 20, 6594–6611.
- Zipser, Karl, Lamme, Victor A.F., Schiller, Peter H., 1996. Contextual modulation in primary visual cortex. *Journal of Neuroscience* 16 (22), 7376–7389.
- Zucker, S.W., Hummel, R.A., 1979. Toward a low-level description of dot clusters: labelling edge interior and noise points. *Computer Graphics and Image Processing* 9, 213–233.