Critical Contours: An Invariant Linking Image Flow with Salient Surface Organization

Benjamin S. Kunsberg
Division of Applied Mathematics
Brown University, RI, USA

Steven W. Zucker
Dept. of Computer Science
Yale University, CT, USA

We exploit a key result from visual psychophysics – that individuals perceive shape qualitatively – to develop a geometrical/topological invariant (the Morse-Smale complex) relating image structure with surface structure. Differences across individuals are minimal near certain configurations such as ridges and boundaries, and it is these configurations that are often represented in line drawings. In particular, we introduce a method for inferring qualitative 3D shape from shading patterns that link the shape-from-shading inference with shape-from-contour. For a given shape, certain shading patches become “line drawings” in a well-defined limit. Under this limit, and invariantly, these shading patterns provide a topological description of the surface. We further show that, under this model, the contours partition the surface into meaningful parts using the Morse-Smale complex. Critical contours are the (perceptually) stable parts of this complex and are invariant over a wide class of rendering models. Intuitively, our main result shows that critical contours partition smooth surfaces into bumps and valleys, in effect providing a scaffold on the image from which a full surface can be interpolated.

Keywords: shape perception, 3D shape reconstruction, orientation fields, shape from shading, shape from line drawings, non-photo-realistic rendering, Morse-Smale complex, combinatorial topology

1. Introduction

Mathematically it is well known that the problem of inferring shape from shading information – or from contour information – is ill-posed: there exist many possible surfaces that could give rise to any type of image structure. In our everyday experience, however, we unconsciously ‘solve’ this inverse problem routinely; we readily and effortlessly infer three-dimensional shape from ambiguous image information. This presents a huge conundrum for theorists: does there exist an invariant that could be used to ground inferences about surface structure on particular types of image structures and, if so, how might this invariant be used by both brains and machines?

Standard computational approaches to resolving this conundrum rely on either imposing strong priors (on light sources, reflectance models, etc.) or on imposing a form of regularization tied to a reflectance model; in either case the goal is to select a unique surface from among the different possibilities. (This material, plus that below, is reviewed in the Background section where references are provided.) The difficulties with this approach are also well-known: reflectance models are rarely, e.g., Lambertian and light sources can be complicated. It is difficult to determine where the light sources are and, moreover, there is little convincing evidence that brains extract sources before inferring shapes.

Our approach is motivated by an important (but frequently overlooked) property of human perception: different individuals (or the

\footnote{We thank Steve Cholewiak, Roland Fleming and Daniel Holtmann-Rice for useful discussions. Research supported by NSF, the Paul G. Allen Family Foundation, and the Simons Foundation.}
same individual at different times) perceive quantitatively different but qualitatively similar surfaces – not identical ones – from either shading or contour information. We take this property to be key: the goal is not to find a unique map between an image and a surface, but rather to identify an equivalence class structure; i.e., to identify which classes of images are consistent with which classes of surfaces. The identification map is then at this abstract level. As we shall show, there are parts of images that do indeed provide a kind of scaffold from which (parts of) surfaces can be reconstructed.

A second insight, intuitively understood by artists and actively studied in computer graphics, further influenced our approach. Non-photorealistic rendering seeks to identify those contours that convey rich surface information. If the surface were known, then one could calculate certain contours in the image that properly convey salient surface properties. One example is the bounding contour in the image; for standard models this is the locus of positions at which the surface normal is orthogonal to the viewing vector. Approximating this are ‘suggestive contours’ or places where the object almost occludes itself. These suggestive contours are not unlike what artists draw, but so far has only been approachable for the forward problem. That is, given the surface, then the suggestive contours can be calculated. We shall show that it is possible to define a class of image contours – what we call critical contours – that work for the inverse problem: they are computable from the image and they have surface interpretations. (Not surprisingly, they turn out to be related to, but not identical with, suggestive contours.)

The third key fact is that many useful shape descriptions are global rather than local. Grasping provides motivation: oranges are round and dogs have legs; one might grasp an orange in any pose but dogs might be grasped by the tail (but probably shouldn’t). Parts, like tails and handles, are object components with particular global properties (extended, thin, attached), but surface normals are local (the derivative of depth at a point). Returning to the qualitative point above, it is relevant that prehensile hand movements form around the pose of the shape (qualitative) and then close upon the object (quantitative).

Accomplishing the transition from local to global has been awkward computationally, although we seem to do it effortlessly. The classical computational approaches yield a numerical surface normal from a partial differential equation subject to boundary conditions, while part-based schemes, such as geons or the medial axis, presume the entire bounding contour is given. That is, the boundary conditions and the bounding contour provide global information. There is another mathematical analog to this: classically, differential geometry is local and topology is global. We exploit these mathematics, starting at the local, differential level and building up to the topological level.

Our mathematical analysis is mainly at the abstract, topological level, and this is where we have found the invariant linking image patterns to surface patterns. It is part of the Morse-Smale complex (to be defined shortly) and, in particular, it is portions of this complex that yields the framework from which the surface can be built. Stated differently, the bounding (occluding) contour can be supplemented with interior critical contours to provide a type of scaffold on which the surface can be reconstructed. The result is ‘locked down’ mainly along these contours and interpolated between them. Differences can arise in the interpolation, of course, which is what gives rise to the qualitative rather than quantitative ‘percept’ seen at the population level. Importantly, any individual ‘sees’ (i.e., reconstructs) a distinct exemplar at a given time. In effect, the generality for the image formation process that is necessary to understand natural scenes has been obtained at a cost: we do not have perfect reconstruction, but only reconstruction up to the critical contour invariants.

Understanding the local-to-global transition is key, we believe, to understanding how brains could perform 3-dimensional inferences from 2-dimensional images. Shading information is first represented in visual cortex by orientationally-selective neurons, as are contours. This is a local representation. But higher cortical areas encode shape, so somehow the brain has accomplished this transition. The Morse-Smale complex provides a framework for understanding this, by starting with flow patterns (e.g. the gradient of the image intensity) and providing global organization (the abstract manner in which these flows are organized).

Problems with building shape inferences on image orientation flows, as they could exist in visual cortex, have been identified previously. Even a small change in light source position could significantly alter the isophotes. (The shading flow is the tangent map to the isophotes.) But isophotes change more in some places than others, and the conditions in our shading limit proposition identify precisely those locations where the isophote structure remains invariant. Anchoring the shape reconstruction on the locations where
the isophote structure is consistent could explain how it is possible for brains to make robust (but qualitative) inferences about shape in 3D. Neural responses should be robust around critical contours, but not necessarily elsewhere, which implies that different positions are represented differently. (Earlier computational approaches treat all positions as equals.)

Our two principal results are as follows. The main idea is to find anchoring configurations in the image that correspond to general surface patterns. While in general huge variation is possible, boundaries and ridges are much more constrained. (We shall show in the Background section that there is little variation across subjects at these positions.) This motivates our first result, which relates the characteristic bright-dark-bright pattern normally associated with ridges on the object to the types of contours artists might sketch. Formally we show that there is a well-defined limiting process in which the ridge-like pattern gets steeper and steeper, until it approaches an ideal critical contour. These critical contours coincide with suggestive contours in certain conditions but, importantly, they are computed from the image and they relate to an analogous contour constraint on the (slant function of the) surface. Proving this relationship is our second main result, and it is done by showing that these contours form part of the Morse Smale complex. Of course, to make the result general and applicable, it holds for a wide variety of possible reflectance functions. In effect, to return to the opening paragraph, we are studying qualitative properties not quantitative properties of the ill-posed differential equations.

We close the Introduction with two final comments. The Morse-Smale complex provides a segmentation of the image (based on the gradient flow of intensities) that is coupled, directly, to the Morse-Smale complex on (the slant function of) the surface. This implies a segmentation – image parts are related to surface parts – in a way that provides a kind of natural semantics to them. This will become clear in the Results section of the paper.

Because our approach is motivated by these many different aspects of the shape inference problem, we next provide a more referenced review of the background material. This is followed by an overview, which the (impatient) reader might wish to peruse next, to get a sense of the technical developments in the paper.

2. Background

In this section we provide a review of several areas of research that are relevant to the inference of shape information from shading and contours. While work in these different areas is often concentrated in separate communities and stresses different aspects of the problem, there is a Gestalt among them. We begin with local approaches, highlighting the classical approach that each image intensity seems to give rise to a surface point (or normal), transition to the many different arguments in support of global structures, and finally broaden the review to highlight the qualitative, rather than quantitative, nature of the inference process.

Classical Approaches and Constancy

The classical view (which goes back to Ernst Mach (Mach, 1965) in 1865) relies on a Lambertian reflectance model that relates image brightness (at a point) to the surface normal at that (projected) location. This is the basic model that drove much of the computational literature (reviewed in (Horn & Brooks, 1989; Brooks & Horn, 1985; Prados & Faugeras, 2006; Zheng & Chappella, 1991)), and results in either a partial differential equation (e.g., the image irradiance equation) or an integral formulation imposing regularization and optimization. The goal is to infer the shape that gave rise to a given image by getting ‘the’ answer at every point. Because shape inference from shading is an ill-posed inverse problem, however, there are an infinity of possibilities. To reduce these possibilities, either very strict assumptions about the rendering function, lighting, material composition, and viewing geometry are required, or these effects must be regularized away. Given that the stated goal is to find a unique surface (regularization implies compromise) that gave rise to a particular image, the regularization can become quite elaborate: Among the more recent papers in computer vision, nearly a dozen different regularization terms are combined together (Barron & Malik, 2012). The result is brittle – it can be made to work for a collection of images but is brittle outside that. It is worth noting that Mach was skeptical that the differential equations approach was the correct one.

Research in visual psychophysics has emphasized a rather different point: there are occasionally conditions under which shape
constancy holds – that is, we seem to solve the inverse problem for small variations in the light source and reflectance model – and many others in which there is significant variation; see, e.g., (P. Sun & Schofield, 2012; Mooney & Anderson, 2014; Mamassian & Kersten, 1996; Mingolla & Todd, 1986; Todd, Egan, & Phillips, 2014; Egan & Todd, 2015; Christou & Koenderink, 1997; Egan & Todd, 2015; Seyama & Sato, 1998; Curran & Johnston, 1996; Khang, Koenderink, & Kappers, 2007; J. J. Koenderink, Doorn, Christou, & Lappin, 1996). Perception can differ in the details, even for special shapes such as cylinders that should be ‘easy’ (but see (Holtman-Rice, Kunsberg, & Zucker, 2017; Kunsberg, Holtman-Rice, & Zucker, 2017)). This has given rise to a number of different viewpoints, for example whether there are different operational modes (P. Sun & Schofield, 2012); or priors about e.g. where the sun is (J. Sun & Perona, 1998). The end result is that, in cognitive science (and much of computer vision), Bayesian approaches are paramount (Kersten, Mamassian, & Yuille, 2004).

The other approach current in computer vision uses neural networks to learn the map from images to depth (or normals) directly from labelled examples. Using range images and RGB color information, this is possible for given scene classes, such as university offices, but doesn’t seem to extend beyond them (Tang, Salakhutdinov, & Hinton, 2012; Eigen & Fergus, 2014; Chakrabarti, Shao, & Shakhnarovich, 2016). This is again an attempt to map intensities at every point to e.g. depth at that point. It, too, is brittle, but now as a function of training data.

This brittle behavior is not a property of our perception. We can perform robustly well beyond everyday scene classes, as is illustrated in Figure 2. Here, we have generated several images of the same blob shape with different rendering functions, by mapping intensity values in [0, 1] through different non-linear (and highly non-convex) functions to other intensity values in [0, 1] via Photoshop. While these images are completely “unnatural shadings,” our visual perception of the major features remains strong. How can we explain this? We suggest, in this paper, that our perception is “anchored” by certain key features, and it is these key features that are invariant across these very different renderings. We seek not to average or regularize over them, but to find a feature that remains distinctive across them. This is, in our terms, where the constancy lies. One possibility for such features are contours, to which we now turn.

**Shape-from-Contour and Non-Photorealistic Rendering**

Contours and line drawings are another source of image structure on which three-dimensional inferences can be based, and these are also classical in computer vision (Barrow & Tenenbaum, 1981; Stevens, 1982; Malik, 1987; Huffman, 1976). Importantly, shape-from-contour is normally considered a separate problem from shape-from-shading: Contours are one-dimensional entities, while shading is a two-dimensional distribution of intensities. For this case the emphasis tends to be on junctions (Waltz, 1975) – the places where surfaces join – and constraint satisfaction, to make sure that the surfaces “fit” together properly. Huffman (Meltzer, Michie, Schank, & Colby, 1975) famously formulated “impossible objects as nonsense sentences” while considering e.g. the Penrose impossible triangle; that is, line drawings for which no continuous scene can exist.

We focus on scenes that do exist. A different type of insight into contours derives from considering not the inverse inference problem but, instead, the forward problem: constructing a line drawing from a scene model. The suggestive contours of De Carlo (DeCarlo, Finkelstein, Rusinkiewicz, & Santella, 2003) were inspiring for us. They are widely used in visualization (Lawonn & Preim, 2016) and relate to the contours artists often draw (Cole et al., 2009); see also (Judd, Durand, & Adelson, 2007; Sahner, Weber, Prohaska, & Lamecker, 2008). The idea behind suggestive contours extends from the information in occluding contours (J. J. Koenderink, 1984); namely, suggestive contours are those ‘almost occluding’ contours that would become occluding with a small change in viewpoint. They tend to occur along extended high-curvature regions of the shape (Lawlor, Holtmann-Rice, Huggins, Ben-Shahar, & Zucker, 2009), which will matter for us. However, to draw suggestive contours in the image one must know the shape: our goal is to infer something like suggestive contours directly from the image, and in a manner that could then be related to the global shape.

Folds in material are one example of when suggestive contours are useful (Jung et al., 2015), and folds have rather structured shading across them. This has been exploited in graphics (Lee, Markosian, Lee, & Hughes, 2007; Gingold & Zorin, 2008), the general forward problem, and is not unrelated to the importance of ridges in computer vision (the inverse problem). However the characterization
of ridges in computer vision remains local, and they are defined in terms of differential geometry (Hallinan, Gordon, Yuille, Giblin, & Mumford, 1999) or singularity theory (Damon, Giblin, & Haslinger, 2016). Yet somehow one must make the transition to global structures, as happens in neurobiology.

Qualitative Representations of Shape

Orientation-selective neurons in visual cortex provide a natural substrate for representing shading information as a flow pattern (J. J. Koenderink & Doorn, 1980; Breton & Zucker, 1996; Ben-Shahar & Zucker, 2004), and this can be used directly as a basis for shape-from-shading-flow computations (Kunsberg & Zucker, 2014a). Although many researchers support this basis (Fleming, Rice, & Bulthoff, 2011; Kim, Marlow, & Anderson, 2014; Holtmann-Rice, Alexander, Fleming, & Zucker, 2013; Anderson & Kim, 2009; Kim et al., 2014), in general the problem remains ill posed. While boundary effects (Ramachandran, 1988) might help, there is a matter of more detailed concern. Todd (Todd et al., 2014; Egan & Todd, 2015), among others, has questioned (forcefully and, to us, in an influential way) whether isophotes and shading flows suffice, because the isophote pattern changes significantly for different renderings and lightings of the same object; but our perception hardly varies (with regard to shape). If our percepts were based on the isophotes alone then they, too, should also change. (Illustrations of the isophotes on a simple furrow shape are shown in Figure 10 for different lightings; indeed, there are positions where they rotate almost 90 degrees.)

The first step out of this dilemma follows from our earlier analysis: there is a sense in which surface structure is more constrained around ridges (Kunsberg & Zucker, 2014b) than elsewhere, although it is still mathematically not unique. This gives a family of surfaces of more similar form, but still an infinite family. Although the variation is higher-order, it brings bas-relief to mind (P. Belhumeur & Yuille, 1999). Nevertheless, there must be a transition to global representations in higher cortical areas (Pasupathy & Connor, 2002; Yamane, Carlson, Bowman, Wang, & Connor, 2008). How, then, can the mathematical non-uniqueness be reconciled with the (apparently) stable neural representations underlying shape perception?

We believe the answer lies in a psychophysical observation: perception is qualitatively similar but not quantitatively identical across subjects. This is a different phrasing of the material discussed above concerning constancy, and virtually all of those references apply here as well (P. Sun & Schofield, 2012; Mooney & Anderson, 2014; Mamassian & Kersten, 1996; Mingolla & Todd, 1986; Todd et al., 2014; Egan & Todd, 2015; Christou & Koenderink, 1997; Egan & Todd, 2015; Seyama & Sato, 1998; Curran & Johnston, 1996; Khang et al., 2007; J. J. Koenderink et al., 1996; Wijntjes, Doerschner, Kucukoglu, & Pont, 2012), plus many others. This suggests that we should not be seeking a single surface from an image, but should be seeking that family of surfaces that are ‘locked down’ by the image. And this is consistent with the physiology: see the examples of ridge-like structure in (Yamane et al., 2008), Fig. 2.

We illustrate this qualitative behavior with the furrow shape (Figure 3), a cartooning of the data from (Nefs, Koenderink, & Kappers, 2005). This shows coarse variation across subjects (labeling of hyperbolic vs. elliptic points) based on estimated surface structure. The variation around the parabolic line speaks against a parabolic line as a descriptor (this was Klein’s famous proposal in (Hilbert & Cohn-Vossen, 1952)) but the lack of variation along the blue contour suggests that it is more likely the invariant descriptor. This blue contour is what we shall call a critical contour. In this case it is part of a suggestive contour but can be found from the image. In more formal language that we shall introduce in Section 4, we shall show that these critical contours are 1-cells of the MS complex of the shading function with high transversal intensity change. (Notice the bright-dark-bright transition across this blue contour.) Such shading patterns are similar (in a limiting sense to be developed) to line drawing contours and they relate directly to surface slant (a perceptually relevant representation of the surface (Stevens, 1983)) segmentations regardless of the rendering functions (under quite weak assumptions). In effect, then, the problem of shape-from-contour is connected to shape-from-shading; it is a kind of robust skeleton that can be used to anchor shading perception. Critical contours provide a scaffold over which the shape can be ‘draped’. Different individuals will drape it in (slightly) different ways, thus providing qualitatively similar but quantitatively different representations of a given shape. (For examples of how this draping is accomplished in graphics, see (Jung et al., 2015).)

The qualitative nature of the solution we are proposing for shape perception bears some resemblance to the categorical parts and
Figure 1: We are interested in reconstruction of 3D shape from an unknown cue. The method we introduce will be inspired by line drawing images (first column (DeCarlo et al., 2003)). However, it will also apply to shading (column 2) and textured images (not shown). It will be invariant with the light source direction to some extent. In the shaded horse example, we chose a light source in the upper right. In the shaded blob example, we chose a light source in the upper left. The third column consists of the slant function of the respective normal field. Note the commonality between the images in each row. We will develop this visual commonality using vector field topology into our theory of critical contours.

necks that can be inferred from the medial axis (or skeleton) (Kimia, Tannenbaum, & Zucker, 1995; Hung, Carlson, & Connor, 2012), but our scheme works for interior lines and shading distributions rather than bounding contours. Importantly, just as the medial axis can be used to structure grasping of objects (Przybylski, Asfour, & Dillmann, 2011), our critical contours may suffice for this as well. Qualitative descriptions can work. Other research in robotics that presumes accurate 3D point clouds has evolved pliable grippers to thwart the inaccuracies of computer vision (Spiers, Liarokapis, Calli, & Dollar, 2016). Moreover, it is relevant that, when reaching to grasp an object, we preform our hands to reflect the pose and extension of the handle (Jeannerod, Arbib, Rizzolatti, & Sakata, 1995; Paulignan, Frak, Toni, & Jeannerod, 1997) In both of these cases, qualitative properties suffice. And notice that points are no longer treated equally: those along the scaffold and the boundary matter more. Flexibility in the rendering function has been traded for a less complete output representation; the gripper’s design and dynamics then quantify it.

The transition to qualitative properties will involve considering topological as well as differential properties. We will provide an introduction to these ideas shortly. For now we simply note that it truly allows a characterization of global properties, and provides an image representation (eventually, a segmentation) that is connected to a surface representation (also a segmentation). Thus it provides a semantics for image parts related to surface parts, and allows us to fill in negative space between artists’ strokes. In effect we get global constraints from local conditions on contours, and evidence is just beginning to accumulate that segmentations such as this induce psychophysical limits (J. Koenderink, Doorn, & Wagemans, 2015).
Figure 2: First row, first two images: Different lightings of a surface, viewpoint constant. Second two: Different viewpoints of the same surface, lighting constant. Note the similar perception even though the local intensity changes can be drastic. Something must remain invariant. We propose it is the Morse-Smale 1-cells of the shading, a geometric/topological construct, that are anchoring the perception to keep it stable. Second, Third Row: Some monotonic (first and second) and non-monotonic (third and fourth) transformations of intensity, as in (Fleming, 2014). Note that shape perception is robust even across these unnatural rendering functions. By connecting the blue contour to shading, we thus inherit geometric meaning (for the surface) from the presence of a contour. In the ideal case, we will model the critical contour as a shaded region elongated along the contour with infinite shading variation across it.
Overview

We now show how the rest of the paper is organized, via the diagram in Figure 4. Starting in the upper left corner, we are given an image of a surface created via an unknown rendering function. It might be a shaded image, a textured image, a line drawing, etc. Classical shape-from-shading methods compute directly from this pixel representation and must confront an ill-posed problem. In Background we discussed some regularizations that have been proposed, so that it can produce an image (dotted right arrow). Instead, we proceed to first identify image features (that will correspond to surface features) that are invariant to the (unknown) rendering function; this is shown as “Image parts”. For shading functions, these image parts are abstracted to stylized lines in Section 5; the relationship is summarized in Lemma 1. This allows us to move to “Critical Contours,” defined in Section 7. Then, Corollary 7 allows us to interpret these Critical Contours as surface curves with important properties (1-cells of the slant Morse-Smale complex), so that we arrive at “Surface 1-cells”. These Surface 1-cells correspond to qualitative parts of the surface (bumps, valleys, ridges, etc.) and function as a kind of scaffold on which the surface can be ‘built.’ Various inpainting or diffusion algorithms could complete this scaffold of 1-cells back into a scalar field; for an example, see Figure 14. The completion is not unique, which relates directly to perception; the scaffold is the qualitative invariant on which different subjects build their quantitative percept. Thus we arrive at a surface exemplar in the upper right.

The following remaining sections of the paper are:

- We introduce the Morse Smale complex – an abstract representation for a scalar field that gives a rigorous correspondence between 2D contours and 3D shape.
- We discuss the image formation and various assumptions, defining an admissable class of rendering functions.
- We interpret a contour from a line drawing as a limit of shaded images, defining a relationship between line drawings and shading.
- We rigorously define critical contours and prove our main result: that critical contours correspond to Morse Smale 1-cells that are invariant across images.
- We illustrate the result (and invariance) on several shaded surfaces.
Figure 4: Overview of our approach, starting in the upper left corner. Suppose given an image of a surface created via an unknown rendering function. Classical shape-from-shading methods attempt to infer, from this pixel representation, a unique surface. We follow the vertical path, and identify those image features that will correspond to surface features. These are the critical contours that delimit ‘image parts.’ Critical contours are invariant to the (unknown) rendering function. We show specifically how shading ‘concentrates’ into such contours; we call this the shading-contour limit. This allows us to move to “Critical Contours,” defined in Section 7, and Corollary 7 allows us to interpret these Critical Contours as surface curves with important properties. We arrive at “Surface 1-cells”, which correspond to qualitative parts of the surface (bumps, valleys, ridges, etc.); these can be completed back into a scalar field such as surface slant, which leads, finally to a surface exemplar in the upper right.

3. The Morse-Smale Complex

Our goal is to find patterns in the image, computable from orientations and invariant to a large class of rendering functions, that “anchor” the ill-posed shape-from-shading problem in a qualitative manner. This is comparable to the qualitative understanding of the phase space in nonlinear dynamics. In that setting, one seeks to understand the (simpler) problem of local neighborhoods of critical points rather than the full geometric problem. Here, we will focus on understanding how particular contours on the image constrain the global qualitative shape.

To do this, we use the Morse-Smale (MS) complex. It is the framework that allows us to convert sets of 2D contours into a 3D shape in a mathematically rigorous way. It uses the sources and sinks of the gradient flow to assign different regions of the domain to critical points. For our purposes, it will treat the 2D contours as partitions of a shape into monotonic regions (called 2-cells), which we think of as the “parts” of the shape. We now provide introductory (but necessary) background on Morse functions and the MS complex.
This introduction is necessarily brief. More complete treatments can be found in (J. Milnor, 2016; Forman, 1998, 2002; Matsumoto, 2002; Biasotti et al., 2008) and, for motivation, see (J. W. Milnor, 1997).

Given an $n$-manifold $M$, consider a smooth scalar function $f : M \rightarrow \mathbb{R}$. (In our case, we will consider the specific case where the image is a scalar function on $M = \mathbb{R}^2$ and align our domain with an $\{x, y\}$ axis.) The gradient

$$\nabla f = \left( \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right)$$

exists at every point. A point $p \in \mathbb{R}^2$ is called a critical point when $\nabla f(p) = 0$. The function $f$ is a Morse function if all its critical points are non-degenerate (meaning the Hessian, or matrix of second derivatives, at those points is non-singular) and if no two critical points have the same function value.

The gradient field gives a direction at every point (except for the critical points, a set of measure zero) in the image. Starting at any point, we can move infinitesimally in the gradient direction to a new point and repeat. This process will trace out an integral line. Precisely, an integral line is a maximal path on the image whose tangent vectors agree with $\nabla f$ at every point of the path. These integral lines naturally must end at critical points (where the gradient direction becomes undefined.) Thus, one can define an origin and destination for each integral line. Further, for each critical point, one defines its ascending manifold as the union of integral lines having that critical point as a common origin. Similarly, one defines its descending manifold as the union of integral lines with that critical point as a common destination.

Each critical point is defined by its index: the number of negative eigenvalues of the Hessian at that point. For scalar functions on $\mathbb{R}^2$, there are only three types of critical points: a maximum (with index 2), a minimum (with index 0) and a saddle point (with index 1). For Morse-Smale functions, which are dense in the set of continuous functions, the integral lines only connect critical points of differing index. The ascending manifold associated with a critical point of index $k$ is of dimension $n - k$. Similarly, the descending manifold for an index $k$ critical point is dimension $k$.

For two critical points $p$ and $q$, with the index of $p$ one greater than the index of $q$, consider the intersection of the descending manifold of $p$ with the ascending manifold of $q$. This intersection will be either a 1D manifold (a curve called a 1-cell) or the empty set. For two critical points $r$ and $s$, with the index of $r$ two greater than the index of $s$, the intersection of the descending manifold of $r$ with the ascending manifold of $s$ will either be a 2D manifold (a region called a 2-cell) or the empty set. Thus, the intersection of all ascending manifolds with all descending manifolds partition the manifold $M$ into 2D regions surrounded by 1D curves with intersections at the critical points.

Consider Figure 5. In the left column, a surface and its level curves are shown in greyscale from two different views. The right column shows the MS complex from two different views. The critical points are marked: the small blue circles are maxima, the small white circles are minima, and the crosses are saddle points. The 1-cells connecting critical points whose indices differ by exactly 1 are shown in white. Note that they surround the 2-cells shown as differently colored regions. Each 2-cell is combinatorially a quadrilateral. Its boundary consists of four 1-cells connected at vertices in the following cyclic order: min, saddle, max, saddle. In addition, the function value along each integral line (not shown) in a 2-cell must be monotonically increasing. Thus, values of the function on the 2-cells are highly constrained given the 1-cells and critical points.

Why use the MS complex as an scalar field representation? Suppose that one only had knowledge of the scalar function at the critical points and 1-cells. One could reconstruct the 2-cells (and thus the entire function) relatively accurately. For some insight, see (Allemand-Giorgis, Bonneau, & Hahmann, 2014; Weinkauf, Gingold, & Sorkine, 2010). The position, heights and boundaries of all the bumps, dimples and ridges are already known and the only choices left are how steep to make the transitions in between. Thus, there is a natural connection between the scalar function restricted to the 2D curves (the salient 1-cells) and the scalar function on the entire domain (the unknown 2-cells). We believe this natural connection can model the visual system’s perceptual ability to “see” 3D shape from a charcoal sketch, for example.
Figure 5: Illustration of the Morse-Smale complex for a scalar function in two dimension. (left column) A ‘mountain range’ seen in perspective and from above. Contours are level sets in height. If the scalar function were image intensity, the level curves would be isophotes. Colored regions represent 2-cells of the MS complex. White curves represent 1-cells (contours) of the MS complex. Maxima, saddles, and minima are represented by solid blue points, crosses and solid white points, respectively. This figure shows how 2D contours can represent the 3D surface, up to monotonic transformations on each part. Figure from (Gyulassy, 2008).

Our main theorem will show that particular 1-cells of the image will be nearly invariant under changes in the rendering function. In the next section, we define the image formation process so we can make this statement precise.

4. Image Formation and Assumptions

We now describe the image formation process. First, we define our 3D coordinate axis \((x, y, z)\) so that \((x, y)\) parametrize the image plane while \(z\) is the view direction. We consider a orthogonal projection model. Let \(e_1, e_2, e_3\) represent the standard basis, as unit vectors in these cardinal directions. We think of the image as being created from a “cue”, or more precisely, by applying a rendering function on the unit sphere \(F : S^2 \to \mathbb{R}\) to the normal field \(N(x, y) = (n_1, n_2, n_3)\) of a smooth surface \(S(x, y)\). That is,

\[
I(x, y) = F(N(x, y))
\]  

Many familiar cues have this structure. For example, Lambertian shading is equivalent to \(F_L(N(x, y)) = \sum_i L_i \cdot N(x, y)\) for diffuse light sources \(L_i\). The spatial frequency cue of isotropic texture (once ideally blurred) is monotonically related to \(F_T(N(x, y)) = [0, 0, 1] \cdot N(x, y)\). Specular shading is equivalent to \(F_S(N(x, y)) = \sum_i [(L_i - 2(N \cdot L_i)N) \cdot e_3]^p\) for specular light sources \(L_i\) and some constant \(p\). We seek image contours that are always present independent of the choice of \(F\).

For the theoretical analysis, we consider differentiable rendering functions \(F\) so that we can consider Lipschitz continuous gradient fields. Note that in practice, as in the textured image, we do not need the rendering function to be differentiable but instead need the ability to calculate this consistent field of orientations. Practically, this can be done on a non-differentiable rendering function using the Structure Tensors or a set of orientation filter responses (see e.g. (Ben-Shahar & Zucker, 2003; Franken & Duits, 2009)).
apply the below theory to textured images in a future paper.)

We take the differential of the above rendering equation to get:

\[ DI(x,y) = DFN(x,y) \circ DN(x,y) \]  

\[ = DF^T DN(x,y) \]  

\[ = \begin{bmatrix} F_x & F_y & F_z \end{bmatrix} \begin{bmatrix} n_{1,x} & n_{1,y} \\ n_{2,x} & n_{2,y} \\ n_{3,x} & n_{3,y} \end{bmatrix} \]  

To get 4 from 3, we just note that function composition in this case is just matrix multiplication. To simplify notation, we also drop the point of application of \( DF \). (A comment on notation: we will use bold lettering to denote tensors and matrices while vectors will be in standard lettering. For an introduction to this tensor notation, see the appendix in (Holtman-Rice et al., 2017).)

Here, we see that \( DI \) is the 1-form corresponding to dot product with the gradient \( \nabla I \), \( DF \) is equivalent to a \( 1 \times 3 \) vector and \( DN \) can be seen as a \( 3 \times 2 \) matrix. We see that the image gradient orientation, that is the angle of the vector \( \nabla I \), is generally dependent on both the surface through the operator \( DN \) and the material/cue through the operator \( DF \). In the most general scenario, where \( DF \) has no constraints, we cannot constrain \( DN \) from the data \( DI \). Thus, we will now put some weak constraints on \( DF \), \( DN \), and \( D^2N \).

### Rendering Function Assumptions

**Definition 1.** We define the admissible cue class as the set of differentiable rendering functions \( F \) satisfying the following two criteria:

1. **Bounded variation:** there exists a \( C_1 \) such that for all rendering functions \( F \) and \( N_0 \in S^2, ||\nabla F|| < C_1 \).

2. **\( F \) is concave and \( D^2F \) is bounded:** There exists a constant \( C_2 \) such that \( C_2 < D^2F_{N_0}(u,u) \leq 0 \) for all \( N_0 \in S^2, u \in T_{N_0}S^2 \).

We elaborate on the conditions. The bounded variation condition ensures that arbitrarily large changes in the image cannot be due to the rendering function alone, but must also require some change in the normal field. Without a constraint on the rendering function such as this one, we could not decipher between gradients due to material changes (such as a painting) and gradients due to natural shading changes. There is perceptual evidence supporting this condition. If an image feature is to be seen as “shading”, it must have generally low contrast. Very high contrast features are often seen as material changes (Holtmann-Rice et al., 2013). The concave condition ensures that, if we imaged the unit sphere with our rendering function, we would see only one highlight (point of maximum brightness).

A given rendering function \( F \) creates a image \( I(x,y) \) of a given imaged surface \( S \). Suppose we now choose a new rendering function \( \tilde{F} \) (e.g. by changing the light source) to get a new image \( \tilde{I}(x,y) \) of the same surface. Our main theorem describes an important commonality between these images. A registration correspondence exists between any two images, \( I(x,y) \) and \( \tilde{I}(\tilde{x}, \tilde{y}) \) but we will not focus on describing or calculating it here. Instead, we will think of the second image as a new scalar field on the same coordinate system: we will prove things regarding \( \tilde{I}(x,y) \). This simplifies the notation.

We must not only restrict our rendering function, but also restrict our surface slightly. We will restrict to generic interactions between the rendering function and imaged surface, which we define below.

### Generic Surface Assumptions

An image of a surface rarely contains a full description of the surface. As shape reconstruction is generally an ill-posed problem, surfaces can collude with rendering functions to create images that hide surface features. We attempt to remove these rare cases via assumptions here. For example, in Lambertian shading, the image is a projection of the surface normal field onto an unknown direction (given by the light source). Thus, the normal field can vary in the directions perpendicular to the light source with no effect on the
image. For example, a Lambertian cylinder could be oriented with respect to the light source in such a way as to create a zero image. In this case, $|DN|$ can be arbitrarily large while $|DI|$ is 0. We want to avoid cases like these. However, a slight change of light source or viewpoint would make the curvature of the cylinder visible and would therefore lead to large changes in the image. Thus, we use the term “generic” to represent stability in the image with respect to slight changes in the rendering function. There are two forms of generic that we will assume for our setup.

1. Given a curve $\alpha(t)$ on a surface $S$, we may assume that the three column vectors of the unfolded tensor $D^2_\alpha(t)N$,

$$\{D^2N(u,u), D^2N(u,w), D^2N(w,w)\} \in \mathbb{R}^3$$

contain at least two that are linearly independent.

2. Let $\theta(v_1, v_2)$ represent the acute angle between two vectors in $\mathbb{R}^3$. Given a curve $\alpha(t)$ on a surface $S$ and rendering function $F$, there exists an $\epsilon$ such that for all $t$:

$$\theta(DF,D_{\alpha(t)}(\cdot)) > \epsilon$$ (6)

$$\theta(DF,D^2_{\alpha(t)}(u,u)) > \epsilon$$ (7)

$$\theta(DF,D^2_{\alpha(t)}(w,w)) > \epsilon$$ (8)

$$\|DF\| > \epsilon$$ (9)

That is, the rendering function’s differential does not happen to align along certain differential properties of the surface normal field. Many of these properties are arbitrarily small measure in the space of continuous configurations. (This removes the Lambertian cylinder example above.) Experimentally, violations of these conditions are rare.

5. The Contour Interpreted as a Shading Limit

We are seeking a visual pattern that is present across multiple views and renderings. We are inspired by artist sketches, where one sees a collection of thin black strokes on a white piece of paper. We would like to think of these strokes as a robust skeleton for describing part boundaries of a shaded image. What is the physical meaning (constraint on the viewed surface) inherent in each stroke? Our main theorem will relate these “ideal” strokes to a Morse Smale complex on the surface normal field. To be precise, we need to investigate differential properties related to one of these contours.

We consider a line drawing image as a collection of these 1D contours. We focus on one contour $\alpha(t)$ and assume it has bounded image (planar) curvature. For each value $\alpha(t)$ on each contour, we have a scalar intensity value $I(\alpha(t))$ and require this value to be 0 at the endpoints. Without loss of generality, let $\alpha(t)$ be arclength parametrized.

**Definition 2.** Then, a ideal 1D contour $\alpha(t) \in \{\alpha_i\}$, $0 \leq t \leq 1$ (a 1D curve) can be expressed as a scalar field $I_\alpha \subset \mathbb{R}^2$ in the following way:

$$I_\alpha(x,y) > 0 \text{ if } (x,y) \in \alpha(t), t \in (0,1)$$

$$I_\alpha(x,y) = 0 \text{ if } (x,y) \in \alpha(t), t \in \{0,1\}$$

$$I_\alpha(x,y) = 0 \text{ if } (x,y) \notin \alpha(t)$$ (10)

We would like to understand the behavior of the image derivatives of $I_\alpha$. However, as $I_\alpha$ is discontinuous, the derivatives do not exist. However, we can approximate these derivatives by considering $I_\alpha$ as the limit of a sequence of shaded images $\{I^\sigma_\alpha\}$, $\sigma \to 0$ on
Figure 6: We inherit the meaning of a contour by considering it as a limit of shaded images. First row: From left to right, we start with the lower left contour from Figure 1, drawn in blue. Next, we represent an element of the shading sequence $I_\sigma^\alpha$ of the contour (a 2D image). Then, we consider another element with larger $\sigma$ and draw its isophotes in green. Second row, left: we show true isophotes from a Lambertian shaded image of the surface. Note the similarity of the two isophote patterns near the contour: on either side of the contour the direction of the level curves rotate to be nearly tangential to the contour. Also, note how the contour (blue) is nearly a gradient flow of the shaded image. The dotted red line represents a transversal direction, with plotted pixel values in Figure e). Note the steep local shading minimum across the contour.
We define each shaded image $I^\sigma_\alpha$ as a convolution with Gaussian functions $G(\sigma)$ of $I_\alpha$ with successively smaller standard deviation $\sigma$ in the following manner.

For every point $p$ on $\alpha(t)$, we can parametrize the local neighborhood $U_\alpha(t)$ with two directions. For convenience, write $u(t) = \alpha'(t)$. Define $w(t)$ as the transversal direction at the point $p$ so that $w(t) \cdot u(t) = 0$. At the point $p$, $\{u, w\}$ is an orthonormal basis and let $\xi, \eta$ be the corresponding coordinate functions. As we are only interested in the limiting behavior and as $\alpha(t)$ has bounded curvature, we can realign our frame so that $p$ is at the origin and the following holds:

$$I_\alpha(p + \xi u + \eta w) = \lim_{\sigma \to 0} I_\alpha(p + \xi u) e^{-\frac{\sigma^2}{2\sigma^2}}$$

(11)

$I_\alpha(p + \xi u) e^{-\frac{\sigma^2}{2\sigma^2}}$ for $\sigma > 0$ approximates the image $I_\alpha$ (which is only nonzero on a 1D manifold) in a local neighborhood by blurring it in the transverse direction. However, this approximation still results in a discontinuous image at the endpoints $\alpha(0)$ and $\alpha(1)$. Thus, we require one more blurring operation in the tangent direction to get a continuous shading function on $\Omega_\alpha$.

Recall the Gaussian kernel:

$$G(\sigma) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\xi)^2}{2\sigma^2}}$$

(12)

We convolve $I_\alpha(p + \xi u + \eta w)$ with $G(\sigma)$ to get:

$$I_\alpha(p + \xi u + \eta w) = \lim_{\sigma \to 0} \frac{1}{\sqrt{2\pi\sigma}} \int I_\alpha(p + xu) e^{-\frac{\sigma^2}{2\sigma^2}} e^{-\frac{(x-\xi)^2}{2\sigma^2}} dx$$

(13)

For any nonzero $\sigma > 0$, we can define

$$I^\sigma_\alpha(\xi, \eta) = \frac{1}{\sqrt{2\pi\sigma}} \int I_\alpha(p + xu) e^{-\frac{\sigma^2}{2\sigma^2}} e^{-\frac{(x-\xi)^2}{2\sigma^2}} dx$$

(14)

which defines a sequence of shaded images so that $\lim_{\sigma \to 0} I^\sigma_\alpha = I_\alpha$. (Note that we cannot simply convolve $I_\alpha$ with a standard 2D Gaussian kernel as it is only positive on a set of zero measure. The difference would be an extra $\sigma$ in the denominator of the above expression.) We will now compute image derivatives, up to second order, along our contour for these shaded images, $I^\sigma_\alpha$. By taking the limit as $\sigma \to 0$, we will inherit derivatives for $I_\alpha$.

We differentiate $I^\sigma_\alpha(\xi, \eta)$ in the tangent direction:

$$\frac{\partial}{\partial \xi} I^\sigma_\alpha(\xi, \eta) \bigg|_p = \int \frac{\partial}{\partial \xi} \left( I_\alpha(p + xu) e^{-\frac{\sigma^2}{2\sigma^2}} e^{-\frac{(x-\xi)^2}{2\sigma^2}} \right) dx$$

(15)

$$= \frac{\partial}{\partial \xi} I_\alpha \bigg|_p e^{-\frac{\sigma^2}{2\sigma^2}}$$

(16)

For each point on $\alpha(t)$, $t \in (0, 1)$, the limit as $\sigma \to 0$ is $\frac{\partial}{\partial \xi} I_\alpha$ as expected. (That is, the image derivative along the contour $\alpha$ is the limit of the derivatives along the shading approximations to the contour.)

We repeat the same process for the remainder of the derivatives and get:

$$\lim_{\sigma \to 0} \frac{\partial}{\partial \eta} I^\sigma_\alpha(\xi, \eta) \bigg|_p = 0$$

(17)

$$\lim_{\sigma \to 0} \frac{\partial^2}{\partial \xi^2} I^\sigma_\alpha(\xi, \eta) \bigg|_p = I''_\alpha(t)$$

(18)

$$\lim_{\sigma \to 0} \frac{\partial^2}{\partial \eta^2} I^\sigma_\alpha(\xi, \eta) \bigg|_p = -\infty$$

(19)

$$\lim_{\sigma \to 0} \frac{\partial}{\partial \xi} \frac{\partial}{\partial \eta} I^\sigma_\alpha(\xi, \eta) \bigg|_p = 0$$

(20)
We also would like the images derivatives at the endpoints of $\alpha(t)$. To calculate these approximations, we apply a Heaviside $H_0(x)$ step function to the image $I_\alpha$. To make the integral feasible, we Taylor approximate the intensity on the contour $I_\alpha(p + xu)$ up to second order:

$$I_\alpha(p + xu) = c_0 + c_1x + c_2x^2$$ (21)

for some constants $\{c_0, c_1, c_2\}$. This “approximation” becomes exact as $\sigma \to 0$. If we are at an endpoint $\alpha(0)$ and we move in the positive tangent direction, the image intensity is defined by this Taylor expansion. If we move in the negative tangent direction, the image intensity is zero. For example,

$$\frac{\partial}{\partial \xi} I_\alpha^\sigma(\xi, \eta) \Big|_{\alpha(0)} = \int \frac{\partial}{\partial \xi} \left(I_\alpha(p + xu)e^{-\frac{x^2}{2\sigma^2}}e^{-\frac{(x-\xi)^2}{2\sigma^2} H_0(x)}\right) dx$$ (22)

$$= \int \frac{\partial}{\partial \xi} \left((c_0 + c_1x + c_2x^2)e^{-\frac{x^2}{2\sigma^2}}e^{-\frac{(x-\xi)^2}{2\sigma^2} H_0(x)}\right) dx$$ (23)

$$= \frac{c_1}{2} + \frac{c_0 + 2c_2\sigma^2}{\sqrt{2\pi\sigma}}$$ (24)

As we require the contour intensity to be 0 at the endpoint, we can set $c_0 = 0$ and calculate the limit of $\frac{\partial}{\partial \xi} I_\alpha^\sigma(\xi, \eta) \Big|_{\alpha(0)}$ as $\epsilon \to 0$ to get $\frac{c_1}{2}$.

We can also calculate the other image derivatives at the endpoint $\alpha(0)$. (Note that the other endpoint, $\alpha(1)$, is just the mirror version; we use $-H_0(x)$ instead of $H_0(x)$.)

$$\lim_{\sigma \to 0} \frac{\partial}{\partial \eta} I_\alpha^\sigma(\xi, \eta) \Big|_{\alpha(0)} = 0$$ (25)

$$\lim_{\sigma \to 0} \frac{\partial^2}{\partial \xi^2} I_\alpha^\sigma(\xi, \eta) \Big|_{\alpha(0)} = \infty$$ (26)

$$\lim_{\sigma \to 0} \frac{\partial^2}{\partial \eta^2} I_\alpha^\sigma(\xi, \eta) \Big|_{\alpha(0)} = -\infty$$ (27)

$$\lim_{\sigma \to 0} \frac{\partial}{\partial \xi} \frac{\partial}{\partial \eta} I_\alpha^\sigma(\xi, \eta) \Big|_{\alpha(0)} = 0$$ (28)

We now consolidate these calculations on this sequence of shaded images.

**Lemma 1.** Let $\alpha(t), \Omega_\alpha, I_\alpha, u, w$ be defined as above. The sequence of shaded images $\{I_\alpha^\sigma\}$ converge pointwise to the original line drawing $I_\alpha$ and have the following properties on the derivatives:

1. As $\sigma \to 0$, $I_{uw}(\alpha(t)) \to -\infty$ for every $t \in [0, 1]$.
2. As $\sigma \to 0$, $I_{uu}(\alpha(t)) \to \infty$ for $t = \{0, 1\}$.
3. As $\sigma \to 0$, there exists a constant $M$ such that $|I_w|, |I_u|, |I_{uw}|, < M$ for every $t \in [0, 1]$.

See Figure 6. As we’ve shown, the first condition implies that the contour sides “pinch in” as $\sigma \to 0$. Note that the gradient on one side of the endpoint is continuous, whereas the gradient on the other side is discontinuous; this is the intuition (proof as above) for the second condition.

We do all this to show rigorously that line drawings can be seen as pointwise close to shading patterns with large derivatives as above. Equivalently, an artist’s stroke may be seen as infinitesimally thin shading pattern. This leads us to define critical contours in the next section, which are nearly invariant shading patterns that mimic these artist’s strokes. Importantly, these critical contours will relate to the Morse Smale complex in an simple way.
Figure 7: In a surface region with anisotropic curvature (the two principal curvatures differ vastly), the image gradient flows are robust as we change the light source. Each curve represents a MS 1-cell (highlight line) on the image corresponding to the light source with the same color. As we move the light source from A to B, the integral path shifts a small amount.

6. Main Thereom

We now define a critical contour, the (nearly) invariant visual pattern across renderings. This critical contour will have image derivatives similar to those calculated in Lemma 1 for the ideal contour.

Definition 3. A $K$-Critical Contour $(\alpha(t), I(x, y), M, K)$ is a curve $\alpha(t)$ on an image $I(x, y)$ such that the following conditions hold for all $t$:

1. $|I_{ww}(\alpha(t))| > K$ for every $t \in [0, 1]$.
2. $|I_{uu}(\alpha(t))| > K$ for $t = \{0, 1\}$.
3. $|DI|, |I_{uw}| < M$ for every $t \in [0, 1]$.

As $K \to \infty$, K-critical contours converge pointwise to the ideal contour we defined in 10. In Theorem 1, we show these K-critical contours are also 1-cells of the Morse Smale complex in any image obtained from most rendering functions in our admissible class.

Theorem 1. Let $F, \tilde{F}$ be any two rendering functions in the admissible cue class. Applying these rendering functions to a generic surface $S$, we obtain two corresponding images $I(x, y), \tilde{I}(x, y)$. For any $\epsilon > 0$, there exists a $K \in \mathbb{N}$ such that the surface region corresponding to an $\epsilon$–neighborhood of a $K$-critical contour in $I$ contains a Morse Smale 1-cell for image $\tilde{I}$.

We explain the intuition for the proof of Theorem 1. Consider the surface $f(x, y) = ax^2 + by^2 + c$ and note that $(0, 0)$ is a critical point of $f$. We think of $f$ as a height function above some plane with normal vector $(0, 0, 1)$. Define $g$ to be a height function from the surface defined by $f$ to a different plane with normal vector $(g_1, g_2, g_3)$. In general, $(0, 0)$ is not a critical point for $g$. However, if $|a|, |b|$ are large enough, then $g$ will have a critical point arbitrarily close to $(0, 0)$. Thus, $f$ and $g$ almost “share” critical points.

From the above example, it is plausible to believe that if the local curvature of a surface $M$ is large enough, scalar fields on $M$ resulting from different projections of the normal field may share critical points. We will take this one step further and show that they also share MS 1-cells, which are one dimensional analogues to critical points. See Fig. 7. We will need to show that the presence of an critical contour implies sufficient curvature across the contour to apply the above intuition.
Figure 8: We have drawn $\Omega$ with $\alpha(t)$ a straight line for simplicity. We show there exists a path consisting of 1-cells $\phi(t)$ of the new image $\tilde{I}$ given the conditions on $\alpha(t)$. The previous lemmas show the existence of $\tilde{a}, \tilde{b}$ and $\beta_1(t), \beta_2(t)$ with the properties on $D\tilde{I}$ as shown.

The proof will follow in 3 steps.

1. Consider Figure 8. We would like to show that the presence of a $K-$critical contour in $\alpha(t) \subset I(x,y)$, implies the following “$\epsilon-$box structure” $\Omega$ on the unknown image function $\tilde{I}(x,y)$. First, we would like to show that there exists critical points $\tilde{a}, \tilde{b} \in \tilde{I}(x,y)$, that are $\epsilon-$close to the endpoints $a, b$ of $\alpha(t)$. (Lemma 3).

2. We show that there exists two curves $\beta_1(t), \beta_2(t)$ in the $\delta-$tubular neighborhood where the gradient $\nabla \tilde{I}$ points away from $\alpha(t)$. (Lemma 4).

3. We then show the “box structure” in Figure 8 will contain a $\tilde{I}$ 1-cell, shown as $\phi(t)$. This is proven by first showing, WLOG, that all integral paths flow from left to right. Then, either $\tilde{a}$ or $\tilde{b}$ must be a saddle and there is an integral path that traverses $\Omega$ connecting to at least one of them. See Figure 8.

To show step 1, we prove Lemma 2 and Lemma 3. To show step 2, we prove Lemma 4. To show step 3, we prove Lemma 5.

**Lemma 2.** **We use notation, generic, and rendering function assumptions as in the previous Sections. If** $|DI_{\alpha(t)}| < M$, then $||DN_{N(\alpha(t))}(\cdot)||$ **is bounded. Similarly, $|DI_{\alpha(t)}| < M$ and $|I_{uw}(\alpha(t))| < M$ imply $||D^2N_{N(\alpha(t))}(u,w)||$ **is bounded.**

**Proof.**

\[ |DI_{p}(\cdot)| < M \quad (29) \]
\[ |DF^TDN_{N(p)}(\cdot)| < M \quad (30) \]

By genericity property 2, we must have $||DN_{N(p)}(\cdot)||$ bounded in Frobenius norm $\forall p \in \alpha(t)$. (We see here that this prevents an infinitesimal change in the rendering function (that is, $DF$) resulting in an unbounded change in the image gradient $\nabla I$.)

We repeat the same argument for $|I_{uw}| < M$, taking one further derivative and leaving off a subscript for clarity:

\[ |I_{uw}| < M \quad (31) \]
\[ |D^2F(DN(u),DN(w)) + DF^TD^2N(u,w)| < M \quad (32) \]
\[ |DF^TD^2N(u,w)| < M + |D^2F(DN(u),DN(w))| \quad (33) \]
Lemma 3. Let \( a \in \Omega_\alpha \subset \mathbb{R}^2 \) be an endpoint of the \( K \) critical contour with the conditions from Definition 3. Given a new rendering function \( \tilde{F} \), resulting image \( \tilde{I} \) and any \( \delta > 0 \), there exists a \( K_0 \) such that if \( K > K_0 \), the following holds: \( \exists \tilde{a} \in \Omega_\alpha \), such that \( ||a - \tilde{a}|| < \delta \) and \( \tilde{a} \) is a critical point of \( \tilde{I} \).

Proof. We will consider the image of the normal field on the Gauss sphere in the neighborhood of \( N(a) \). The main idea is, if a differentiable function has a large enough gradient at a point \( a \), then it has a zero inside a neighborhood of \( a \). We take two derivatives of the equation \( I(x, y) = F(N(x, y)) \) to get the following equation of matrices:

\[
D^2 I_{(x,y)}(u_1, u_2) = D^2 F(DN_{(x,y)}(u_1), DN_{(x,y)}(u_2)) + D\tilde{F}^T D^2 N_{(x,y)}(u_1, u_2)
\]  

(34)

To calculate \( I_{uw}(a) \), we replace \( u_1, u_2 \) both with \( w \) and let \( (x, y) = a \), where \( a \) is any point on \( \alpha(t) \). By the concave rendering function assumption:

\[
0 > D^2 F(DN_a(w), DN_a(w))
\]  

(35)

If \( I_{uw} > K \), then

\[
K < D\tilde{F}^T D^2 N_a(w, w)
\]  

(36)

As \( ||\nabla F|| < C_1 \) by the rendering function assumptions,

\[
\frac{K}{C_1} < ||D^2 N_a(w, w)||
\]  

(37)

Similarly, for \( I_{uw} \), we get

\[
\frac{K}{C_1} < ||D^2 N_a(u, u)||
\]  

(38)

Recall that \( D\tilde{F} \) is the differential of the second rendering function. We expand the operator \( D\tilde{F}^T DN_a(\cdot) \) in a first order Taylor expansion around \( a \).

\[
D\tilde{F}^T DN_{a+qu+rw}(\cdot) = D\tilde{F}^T DN_a(\cdot) + q D\tilde{F}^T D^2 N_a(u, \cdot) + r D\tilde{F}^T D^2 N_a(w, \cdot)
\]  

(39)

We know that \( ||D^2 N_{N_a}(u, u)|| \) and \( ||D^2 N_{N_a}(w, w)|| \) are sufficiently large; we want to find a \( (q_0, r_0) \) such that \( D\tilde{F}^T DN_{a+q_0u+r_0w}(\cdot) \) is precisely 0.

We use the first generic property to assume that the span of three vectors \( \{D^2 N_a(u, u), D^2 N_a(u, w), D^2 N_a(w, w)\} \) contains at least two linearly independent ones. This implies that \( D\tilde{F}^T D^2 N_a(w, \cdot) \) and \( D\tilde{F}^T D^2 N_a(u, \cdot) \) are not parallel vectors and thus they span a plane \( P \). Generically, \( P \) contains an intersection point with the unknown vector \( -D\tilde{F}^T DN_a(\cdot) \). That intersection point defines a \( (q_0, r_0) \) satisfying \( D\tilde{F}^T DN_{a+q_0u+r_0w}(\cdot) = 0 \).

Define \( \tilde{a} = a + q_0u + r_0w \). It remains to show \( ||a - \tilde{a}|| < \epsilon(c_1, c_2) \). Recall that \( D\tilde{F}^T DN_a(\cdot) \) is a bounded vector by Lemma 2. From 39:

\[
-D\tilde{F}^T DN_a(\cdot) = q_0 D\tilde{F}^T D^2 N_a(u, \cdot) + r_0 D\tilde{F}^T D^2 N_a(w, \cdot)
\]  

(40)
The above matrix equation represents two equations:

\[ -DF^T DN_a(u) - q_0 D_{\tilde{F}}^T D^2 N_a(u, u) = r_0 D_{\tilde{F}}^T D^2 N_a(w, u) \tag{41} \]
\[ -DF^T DN_a(w) - r_0 D_{\tilde{F}}^T D^2 N_a(w, w) = q_0 D_{\tilde{F}}^T D^2 N_a(w, w) \tag{42} \]

We note that \( q_0 D_{\tilde{F}}^T D^2 N_a(u, w) \) and \( r_0 D_{\tilde{F}}^T D^2 N_a(w, u) \) are bounded by Lemma 2. Consider Equation 41. There exists some \( \Gamma \in \mathbb{R} \) such that:

\[ |DF^T DN_a(u)| + |r_0 D_{\tilde{F}}^T D^2 N_a(w, u)| < \Gamma \tag{43} \]
\[ |q_0 D_{\tilde{F}}^T D^2 N_a(u, u)| < \Gamma \tag{44} \]
\[ |q_0| |D_{\tilde{F}}| |D^2 N_a(u, u)|| |\cos(\theta)| < \Gamma \tag{45} \]
\[ |q_0| < \frac{\Gamma}{\cos(\theta) |D^2 N_a(u, u)|||DF|} \tag{46} \]
\[ |q_0| < \frac{\Gamma C_1}{\epsilon^2 K} \tag{47} \]

where \( \theta \) is the angle between the vectors \( D_{\tilde{F}} \) and \( D^2 N_a(u, u) \). By generic assumptions (Equation 8), \( |D_{\tilde{F}}|, |\cos(\theta)| \) are bounded below by \( \epsilon \). To go from 46 to 47, we also substituted in from 38. Similarly, from Equation 42, we get \( |r_0| < \frac{\Gamma C_1}{\epsilon^2 K} \).

Now, \( r_0 \) and \( q_0 \) are the displacements from point \( a \) to point \( \tilde{a} \):

\[ ||a - \tilde{a}|| = \sqrt{r_0^2 + q_0^2} \tag{48} \]
\[ < \sqrt{2} \frac{\Gamma C_1}{\epsilon^2 K} \tag{49} \]

Define \( \delta(K) = \sqrt{2} \frac{\Gamma C_1}{\epsilon^2 K} \) to complete the proof.

\[ \square \]

**Lemma 4.** Let \( \alpha(t) \) be a \( K \)-critical contour with the conditions from Definition 3. Recall that \( w(t) \) is the transversal direction in \( \mathbb{R}^2 \) to \( \alpha \) for each point \( \alpha(t) \). Given a new rendering function \( \tilde{F} \), resulting image \( \tilde{I} \) and \( \delta > 0 \), there exists a \( K_0 \) such that if \( K > K_0 \), the following holds:

Define two curves \( \beta_1(t) = \alpha(t) + \delta w \) and \( \beta_2(t) = \alpha(t) - \delta w \). On \( \beta_1(t) \), \( D\tilde{I}_{\beta_1(t)}(w(t)) > 0 \) and on \( \beta_2(t) \), \( D\tilde{I}_{\beta_1(t)}(w(t)) < 0 \).

**Proof.**

\[ K < I_{ww} \tag{50} \]
\[ < DF^T D^2 N(w, w) \tag{51} \]

by applying equations 34 and 35. By the rendering function assumption \( ||DF|| \) is bounded by \( C_1 \):

\[ \frac{K}{C_1} < ||D^2 N(w, w)|| \tag{52} \]

By the generic assumptions Equation 8:

\[ |D_{\tilde{F}}^T D^2 N(w, w)| > |DF| |D^2 N(w, w)|| \cos(\theta) \tag{53} \]
\[ > \frac{K \epsilon^2}{C_1} \tag{54} \]
Expand the scalar function $DF^TDN(w)$ to first order in the $w$ direction around any point on the contour $\alpha(t)$:

$$\nabla \tilde{I}(\alpha(t) + sw) \cdot w = DF^T D_{\alpha(t) + sw}(w)$$

$$\approx DF^T D_{\alpha(t)}(w) + s DF^T D^2_{\alpha(t)}(w, w)$$

(55)

(56)

Fix $\delta$. It suffices to show that this quantity is positive for $s = \delta$ and negative for $s = -\delta$. If so, then we have shown e.g. $D\tilde{I}_{\beta_1(t)}(w) > 0$ for $\beta_1(t) = \alpha(t) + \delta w(t)$. We now note that the first term on the RHS of equation 56 is upper bounded for all $t$ on the critical contour $\alpha(t)$ by Lemma 2. And note that the second term is lower bounded proportional to $K$ from 54. Thus, for the given $\delta$ and for all $t$, we can choose a $K$ such that

$$|DF^T D_{\alpha(t)}(w)| < |\delta DF^T D^2_{\alpha(t)}(w, w)|$$

(57)

We see that in equation 56, the first term is dominated by the second term, whose sign is dependent on $s$. Thus, we get the necessary gradient conditions on $D\tilde{I}$.

Lemma 5. Let $\alpha(t)$ be a $K-$ critical contour with the conditions from Definition 3. Given a new rendering function $\tilde{F}$, resulting image $\tilde{I}$ and $\delta > 0$, apply the previous two lemmas. We can find critical points $\tilde{a}, \tilde{b}$ of $\tilde{I}$ and two curves $\beta_1(t), \beta_2(t)$ arbitrarily close to $\alpha(t)$. (See Figure 8.) Parametrize two line segments $\gamma_1(s), \gamma_2(s)$ with the following properties:

$$\gamma_1(0) = \beta_1(0), \gamma_1(1) = \beta_2(0), \gamma_1(0.5) = \tilde{a}$$

$$\gamma_2(0) = \beta_1(1), \gamma_2(1) = \beta_2(1), \gamma_2(0.5) = \tilde{b}$$

Define the region $\Omega$ bounded by the curves $\{\gamma_1, \beta_1, \gamma_2, \beta_2\}$. Without loss of generality, every integral path that intersects $\Omega$ enters from a point on $\gamma_1$ and leaves on a point on $\gamma_2$.

Proof. First, we assume that there are no critical points in the interior of $\Omega$. If there are, bisect $\Omega$ into $\Omega_1$ and $\Omega_2$ and repeat the following argument.

Let $S_1$ be the set of all points in $\Omega$ on integral curves entering from points on $\gamma_1$. We say that an integral path $P$ enters from $\gamma_1$ when there exists a $r, s$ such that $\gamma_1(r) = P(s)$ and $P'(s) \cdot \alpha'(0) > 0$. Let $S_2$ be the set of all points in $\Omega$ on integral curves entering from points on $\gamma_2$.

Clearly, $\Omega = S_1 \cup S_2$. It suffices to show that one of the $S_1$ is empty. Suppose not; suppose $S_1 \neq \emptyset, S_2 \neq \emptyset$. Being a tubular neighborhood of a curve $\alpha(t)$, $\Omega$ is a topologically connected space in $I$. Thus, $\tilde{S}_1$ and $\tilde{S}_2$ must not be disjoint. There exists a point $p \in \tilde{S}_1 \cap \tilde{S}_2$. As there are no critical points in $\Omega$, $\nabla I_2(p) \neq 0$. For any $\epsilon$, there exists a $\epsilon$ neighborhood of $p$ containing both an integral path $\psi_1 \subset S_1$ and an integral path $\psi_2 \subset S_2$. However, an integral path is the solution to a differential equation $\psi(t) = \nabla I_2(\psi(t))$ with initial condition $\gamma(0) = q$. For a Lipschitz continuous gradient field, there is continuous dependence of solutions on the initial conditions $\gamma(0)$ (?, ?). Thus, $\psi_1$ and $\psi_2$ must be arbitrarily close together, which yields a contradiction, as they go through points on opposite sides of $\Omega$.

We now have all the pieces to prove the main theorem. All that is necessary is to show that, given the conditions in the above lemma, there is a Morse Smale 1-cell of $I_2$ contained in $\Omega$.

Proof of Theorem. From Lemma 5, we see that all integral lines flow from the right side of $\Omega (\gamma_1)$ to the left side of $\Omega (\gamma_2)$ or vice versa. $\tilde{a}$ is a critical point on $\gamma_1$ and $\tilde{b}$ is a critical point on $\gamma_2$. As the flow direction on $\beta_1, \beta_2$ points outwards for all $t$, the critical index of $\tilde{a}$ and $\tilde{b}$ can only differ by at most 1. Without loss of generality, $\tilde{b}$ has an incoming integral path $\phi$ starting from the other side of $\Omega$. Thus $\tilde{b}$ must be a saddle point and $\phi$ must be a Morse Smale 1-cell traversing $\Omega$. □
Corollary 6. As $K \to \infty$, as in the case of our ideal contour, a $K-$critical contour $\alpha(t)$ in any admissible image represents a Morse Smale 1-cell in any other admissible image.

Proof. As $K \to \infty$, the tubular neighborhood $\Omega$ of $\alpha(t)$ shrinks to zero width. The integral path $\phi(t)$ must traverse $\Omega$ and thus must eventually lie on $\alpha(t)$. 

This means that an ideal contour represents a visual commonality among all images of the surface $S$. As the normal slant function is a member of our admissible rendering functions, an ideal contour also represents a surface property: a MS 1-cell of the slant function. Thus, we can now interpret an ideal contour as a surface property that “shines through” in every image created by any of the rendering functions.

Corollary 7. For $K$ sufficiently large, a $K$-critical contour $\alpha(t)$ in image $I$ of surface $S$, aligns with a Morse Smale 1-cell of the slant of the surface normal field of $S$.

Proof. Define $\tilde{F}(N) = \langle e_3, N \rangle$ as the rendering function corresponding to a Lambertian surface with light source in the view direction $e_3$. Define $\tilde{I}$ as the image associated with the surface $S$ using this rendering function. As the slant of the normal field is a monotonic function of the image $\tilde{I}$, it shares the same Morse Smale complex as $\tilde{I}$. Apply the Theorem to show that $\alpha(t)$ aligns with a Morse Smale 1-cell of $\tilde{I}$.

7. The Morse-Smale Complex on Shading and Slant

We now apply the above theory to a number of different shapes, to illustrate the Morse-Smale complex and critical contours computed from the image related to the Morse-Smale complex on the slant function of the surface. The results are ordered in complexity; see Figs. 9, 10, and 11.

A note on methodology: A 3D mesh was generated for each figure, which was then rendered under different conditions to produce each image. We use (Reininghaus & Hotz, 2011) to calculate the MS complex and consider persistence simplifications with few critical points from these images. We experimentally verified that MS 1-cells with large $I_{ww}$ remain positionally stable across these images, as predicted by our theorem. We observe that, because the computations are run directly on quantized pixel values, there are certain numerical issues. We do not believe that these would be problems for biological applications, in which the MS complex were computed directly from the gradient flow rather than the pixel values; algorithms for accomplishing this will be reported in future work.

The first example (Fig. 9) consists of a large bump which, as the light source moves, illustrates the common critiques of flow based approaches: large movement in the isophotes (and in the location of the maximum in intensity). Notice, however, that two of the Morse-Smale 1-cells (blue curves) form a circle and remain fixed surrounding the bump: these have large $K$; i.e. these have large bright-dark-bright transitions. These are the critical contours. The movement of critical points around the boundary of the image is a numerical issue; for the reasons above these are irrelevant to the shape representation.

The second example is the furrow shape (Fig. 10) shown from two views and with drastically different lightings. Notice how the isophotes move; how the maxima move, but how the critical contours remain fixed. Interestingly, the central saddle point moves with the light source by ’sliding’ along the 1-cell. It is interesting to compare these different 1-cells with the cartooned perceptual results (Fig. 3): the movement of the 1-cells that do not correspond to critical contours live within the perceptual region of variation across subjects. While in this case it is due to a light-source movement, in general it could correspond to the different implicit “sources” (priors) underlying the subjects’ judgements.

The next example is a random ‘blob’ shape, illuminated from very different viewpoints, plus the slant function on this shape. Notice the stability of the critical contours; how these agree across lightings and for the slant function, and how these stable 1-cells correspond to the suggestive contours that were computed from the true 3D shape.
Figure 9: Top: a slightly perturbed sigmoid rotated around the z axis. The color at each point indicates the absolute value of the Gaussian curvature. An arrow points to a turquoise band which is centered along a contour of near zero Gaussian curvature. (This is a 1-cell of the MS complex of the slant function). Under a wide class of rendering functions (as described above), the resulting shaded image will contain a 1-cell along this band. First Row: From left to right, the first 2 images are Lambertian shaded renderings of the above rotated sigmoid with different light sources. The third image is a specular rendering. The fourth image is the slant function. Second Row: Corresponding MS complexes to the images above along with isophotes in red. Blue arcs correspond to the 1-cells. Yellow, green and red points correspond to maximum, saddle and minimum critical points. Notice the blue common circular contour (which consist of unions of 1-cells).
Figure 10: A second example that shows the commonality between 1-cells over both large light source changes and large view changes. First Row: A “furrow shape” lit from three directions and true slant in fourth column. Second Row: The MS 1-cells with critical points (maxima in red, saddles in green) corresponding to the images in the first row, except for the last one which shows the critical contour sketched in red. Third Row: A second view of the furrow shape lit from three directions and true slant in fourth column. Fourth Row: The MS 1-cells corresponding to the images in the third row, plus critical contour also sketched.
Figure 11: First row: Images of a blob under different rendering functions and their corresponding MS complexes. First image: Lambertian rendering with light source at view. Second, third images: Lambertian rendering with alternate light sources (> 45 degrees off view). Fourth image: Surface rendered according to slant function. Second row: Corresponding MS complexes for above images, except fourth image which shows sketched critical contours. For contrast, note the suggestive contours (in perspective projection) (DeCarlo et al., 2003) for the same surface in the third row. The extra suggestive contour in upper right is not seen in orthographic projection.

In our next example, we experimentally verify Corollary 7. In Fig. 12, we overlay the MS complex for the horse image with the MS complex of the slant field. Note the correspondence between the red segmentation, blue segmentation, and suggestive contours, as predicted by our theory. On those curves where the two MS complexes are not in exact alignment, the value of $K$ is not sufficiently large. This indicates where the qualitative structure of the slant (of the normal field) can be immediately and robustly inferred from a shaded image via the MS complex.
Figure 12: Ideal line drawings, as modeled by suggestive contours (DeCarlo et al., 2003), relate to the Morse Smale complex of 1-cells of both the image and slant scalar functions. In particular, lines are drawn at the 1-dimensional intersections of these two MS complexes. Left: a line drawing. Center: the persistence simplified 1-cells of the image field (red) and persistence simplified 1-cells of the slant function (blue). Right: a zoomed image of the back hip. In cases where the two MS complexes don’t exactly align, the value of $K$ is not large enough.

An important consequence of the global nature of the MS complex is that it provides enough of the qualitative solution to segment the surface into salient parts as in Fig. 13. There is a maximum on each of the four primary lobes, plus several others. The part regions surrounding these are delimited by a Morse-Smale 2-cells, as are the interior (less reliable) 2-cells. (It is these interior 2-cells that will 'move around' with the light sources.

Figure 13: This figure illustrates how the surface is segmented by the full MS complex for the shading into salient parts. Notice the four major lobes, pointing outward like the ends of an ‘X’, plus some interior parts. Crisp maxima in intensity signal the four dominant lobes. The middle maximum signals an interior part.

The remaining question is how to quantitatively reconstruct a scalar field from only knowledge of its critical contours or 1-cells of its MS complex. This question has been considered, for example in (Allemand-Giorgis et al., 2014; Weinkauf et al., 2010). Here, we show a simple example for the furrow object that the segmentation induced by 1-cells of the MS complex is nearly sufficient. (See Figure 14.) In other words, solving for the 1-cells of the slant function is a qualitative representation that may be used to approximately
recover the entire function. In (Gerber, Bremer, Pascucci, & Whitaker, 2010), one can see several applications how visualizing the graph structure of the MS complex can capture the essential phenomena of real world data. Biologically plausible versions of this interpolation will be considered in subsequent work.

Figure 14: True slant function (upper left). Remaining 3 figures: Sample reconstructions (using a linear inpainting algorithm (D’Errico, 2012)) with just the red critical contour shown in Figure 10. Note the strong similarity between the slant reconstructions even though the original images (shown also in Figure 10) are pointwise very different.

8. Discussion

Many different communities have considered the shape inference problem, and many different perspectives have evolved regarding it. Crudely (‘to first order’) in computer vision it is viewed as an inverse problem, which suggests the task is to find the proper regularizers or priors that, when imposed, yield a “best” solution. In visual psychophysics, it is viewed as an experimental task, to find the similarities and differences across conditions and between observers. In computer graphics, it is viewed as a means for information display. Each of these communities has contributed key insights, each of which has been a huge influence on our work. The brittleness of priors necessitated an approach that will work for a wide range of rendering functions; the psychophysics revealed the qualitative rather than quantative nature of the process; and non-photorealistic rendering stimulated our thinking regarding curves to represent continuous shaded surfaces. This latter point is, of course, well known through artists’ sketches. In the end, for us, it takes the form of the shading-to-contour limit, which is formulated as curves through intensity distributions with ‘steep walls.’ This formally leads to a kind of semantics for the image’s critical contours, a semantics that unifies the shape-from-shading and the shape-from-contour problems. In effect, contours can be seen as infinitesimally thin shading patterns and thus inherit surface constraints.

The puzzle of how we effortlessly solve the inference problem is explained by moving to an abstraction, the Morse-Smale complex, and it is at this level that we obtain our second result: the critical contours are part of a cellular decomposition of the image that relates, formally, to the cellular decomposition of the (slant function of the) surface. In other words, the critical contours form (part of) a scaffolding on which the full surface can be reconstructed. Importantly, this captures the qualitative aspect of our percepts from psychophysics, it holds for a wide range of rendering functions, and it provides an image-based model for (the interior curves in) artists’ sketches.
By seeking an invariant between image flows and surface properties we were able to solve another puzzle: if isophotes carry all of the image information, but they change with (even small) changes in lighting or reflectance, how is it possible that our percepts remain (somewhat) constant? The answer is that, while the isophotes change in many places, they are relatively stable in others. It is these stable loci that define the critical contours. It is the topological machinery that permits the local to global transformation, and it has a deep consequence: All points are not treated equally, as in the differential equations approach.

One technical point needs polishing. In the introductory sections of the paper we refer to the critical contours as the invariant linking image structure to surface structure. In the mathematical sections, however, two technical parameters were introduced. First, the shading contour limit needed a $\sigma$ (or blur) parameter, and, second, the $K$-critical contours require ‘wall’ heights of magnitude $K$. These two parameters are describing the same phenomena – that is, describing how close a shading pattern is to a line. Either one is sufficient to define critical contours, so we use just $K$. Formally, the invariant exists in the limit as $K \to \infty$ (which implies $\sigma \to 0$).

Restating our main theorem intuitively: on any smoothly rendered image, almost every $K$-critical contour partitions the slant field into bumps and valleys. Equivalently, the $K$-critical contours are part of a meaningful segmentation of the surface shared by almost all the renderings. (As $K \to \infty$, ‘almost’ becomes ‘all.’)

Much remains to be explored further. First, what are the implications of focusing on the image neighborhood around critical contours for psychophysics and for neurophysiology? Second, the full MS complex induces a notion of image parts and their relationship to surface parts. This has implications for categorical descriptions, for the propagation (or diffusion) of image properties, and for recognition. Thirdly, we only addressed reconstruction of the slant function from the MS complex in this paper; tilt also must be considered to recover a complete surface representation. There are, of course, very similar mathematical constraints for the tilt at critical contours, so we expect an identical theory. This will be explored in a separate paper. Finally, preliminary experiments show that the same ideas apply to shape-from-texture, but there are additional issues regarding scale and the orientation flow. These, plus the above issues, will be addressed in subsequent papers.
Allemand-Giorgis, L., Bonneau, G.-P., & Hahmann, S. (2014). Piecewise Polynomial Reconstruction of Functions from Simplified Morse-Smale complex. *IEEE Visualization Conference*.

Anderson, B. L., & Kim, J. (2009). Image statistics do not explain the perception of gloss and lightness. *Journal of Vision*, 9(11), 10. Available from [http://dx.doi.org/10.1167/9.11.10](http://dx.doi.org/10.1167/9.11.10)

Barron, J. T., & Malik, J. (2012). Shape, albedo, and illumination from a single image of an unknown object. In *Ieee conference on computer vision and pattern recognition* (pp. 334–341). IEEE.

Barrow, H. G., & Tenenbaum, J. M. (1981, August). Interpreting line drawings as three-dimensional surfaces. *Artif. Intell.*, 17(1-3), 75–116. Available from [http://dx.doi.org/10.1016/0004-3702(81)90021-7](http://dx.doi.org/10.1016/0004-3702(81)90021-7)

Ben-Shahar, O., & Zucker, S. (2004). Geometrical Computations Explain Projection Patterns of Long-Range Horizontal Connections in Visual Cortex. *Neural Computation*, 16, 445 - 476.

Ben-Shahar, O., & Zucker, S. (2003). The Perceptual Organization of Texture Flow: A Contextual Inference Approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(4).

Biasotti, S., De Floriani, L., Falcidieno, B., Frosini, P., Giorgi, D., Landi, C., et al. (2008, October). Describing shapes by geometrical-topological properties of real functions. *ACM Comput. Surv.*, 40(4), 12:1–12:87. Available from [http://doi.acm.org/10.1145/1391729.1391731](http://doi.acm.org/10.1145/1391729.1391731), doi:10.1145/1391729.1391731

Breton, P., & Zucker, S. (1996, June). Shadows and shading flow fields. *Proc. IEEE Conf. on Computer Vision and Pattern Recognition - CVPR ’96*, 782 - 789.

Brooks, M., & Horn, B. (1985). Shape and Source from Shading. *Proceedings of International Joint Conference on Artificial Intelligence*, 932-936.

Chakrabarti, A., Shao, J., & Shakhnarovich, G. (2016). Depth from a single image by harmonizing overcomplete local network predictions. *CoRR, abs/1605.07081*. Available from [http://arxiv.org/abs/1605.07081](http://arxiv.org/abs/1605.07081)

Christou, C. G., & Koenderink, J. J. (1997). Light source dependence in shape from shading. *Vision Research*, 37(11), 1441 - 1449. Available from [http://www.sciencedirect.com/science/article/pii/S0042698996002829](http://www.sciencedirect.com/science/article/pii/S0042698996002829), doi:http://dx.doi.org/10.1016/S0042-6989(96)00282-9

Cole, F., Sanik, K., DeCarlo, D., Finkelstein, A., Funkhouser, T., Rusinkiewicz, S., et al. (2009). How well do line drawings depict shape? *ACM Trans. Graph.*, 28.

Curran, W., & Johnston, A. (1996). The effect of illuminant position on perceived curvature. *Vision Research*, 36(10), 1399 - 1410. Available from [http://www.sciencedirect.com/science/article/pii/0042698995002138](http://www.sciencedirect.com/science/article/pii/0042698995002138), doi:http://dx.doi.org/10.1016/0042-6989(95)00213-8

Damon, J., Griblin, P. J., & Haslinger, G. (2016). *Local features in natural images via singularity theory* (Vol. 2165). Springer.

DeCarlo, D., Finkelstein, A., Rusinkiewicz, S., & Santella, A. (2003). Suggestive contours for conveying shape. *ACM Trans. Graph.*, 22(3), 848–855.

D’Errico, J. (2012). Inpaint nans. *MATLAB Central File Exchange*.

Egan, E. J. L., & Todd, J. T. (2015). The effects of smooth occlusions and directions of illumination on the visual perception of 3-d shape from shading. *Journal of Vision*, 15(2), 24. Available from [http://dx.doi.org/10.1167/15.2.24](http://dx.doi.org/10.1167/15.2.24), doi:10.1167/15.2.24

Eigen, D., & Fergus, R. (2014). Predicting depth, surface normals and semantic labels with a common multi-scale convolutional architecture. *CoRR, abs/1411.4734*. Available from [http://arxiv.org/abs/1411.4734](http://arxiv.org/abs/1411.4734)

Fleming, R. (2014). Visual Perception of materials and their properties. *Vision Research*, 94, 62 - 75.

Fleming, R., Rice, D.-H., & Bulthoff, H. (2011). Estimation of 3d shape from image orientations. *Proceedings of the National Academy
Koenderink, J. J. (1984). What does the occluding contour tell us about solid shape? *Perception, 13*(3), 321-330. Available from [http://dx.doi.org/10.1068/p130321](http://dx.doi.org/10.1068/p130321) (PMID: 6514517), doi:10.1068/p130321

Koenderink, J. J., & Doorn, A. J. van. (1980). Photometric invariants related to solid shape. *Journal of Modern Optics, 27*(7), 981–996.

Koenderink, J. J., Doorn, A. J. van, Christou, C., & Lappin, J. S. (1996). Perturbation study of shading in pictures. *Perception, 25*(9), 1009-1026. Available from [http://dx.doi.org/10.1068/p251009](http://dx.doi.org/10.1068/p251009) (PMID: 8983042), doi:10.1068/p251009

Kunsberg, B., Holtman-Rice, D., & Zucker, S. W. (2017). What’s In A Patch, II: Visualizing generic surfaces. *ArXiv e-prints.*

Kunsberg, B., & Zucker, S. W. (2014a, April). How Shading Constrains Surface Patches without Knowledge of Light Sources. *SIAM Journal on Imaging Sciences, 7*(2), 641–668.

Kunsberg, B., & Zucker, S. W. (2014b). Why shading matters along contours. In *Neuromathematics of vision* (pp. 107–129). Springer.

Lawlor, M., Holtmann-Rice, D., Huggins, P., Ben-Shahar, O., & Zucker, S. W. (2009). Boundaries, shading, and border ownership: A cusp at their interaction. *Journal of Physiology-Paris, 103*(12), 18 - 36. Available from [http://www.sciencedirect.com/science/article/pii/S0928425709000254](http://www.sciencedirect.com/science/article/pii/S0928425709000254) (Neuromathematics of Vision), doi:https://doi.org/10.1016/j.jphysparis.2009.05.005

Lee, Y., Markosian, L., Lee, S., & Hughes, J. F. (2007, July). Line drawings via abstracted shading. *ACM Trans. Graph.*, 26(3). Available from [http://doi.acm.org/10.1145/1276377.1276400](http://doi.acm.org/10.1145/1276377.1276400), doi:10.1145/1276377.1276400

Mach, E. (1965). On the physiological effect of spatially distributed light stimuli. *Translation in Ratliff, F., Mach Bands: Quantitative studies on neural networks in the retina.*

Malik, J. (1987). Interpreting line drawings of curved objects. *International Journal of Computer Vision, 1*(1), 73–103. Available from [http://dx.doi.org/10.1007/BF00128527](http://dx.doi.org/10.1007/BF00128527), doi:10.1007/BF00128527

Mamassian, P., & Kersten, D. (1996). Illumination, shading and the perception of local orientation. *Vision Research, 36*(15), 2351 - 2367. Available from [http://www.sciencedirect.com/science/article/pii/0042698995002863](http://www.sciencedirect.com/science/article/pii/0042698995002863), doi:https://doi.org/10.1016/0042-6989(95)00286-3

Matsumoto, Y. (2002). *An introduction to morse theory* (Vol. 208). American Mathematical Soc.

Meltzer, B., Michie, D., Schank, R., & Colby, K. (1975). Machine intelligence 7.

Milnor, J. (2016). *Morse theory.* (am-51) (Vol. 51). Princeton university press.

Milnor, J. W. (1997). *Topology from the differentiable viewpoint.* Princeton University Press.

Mingolla, E., & Todd, J. T. (1986). Perception of solid shape from shading. *Biological Cybernetics, 53*(3), 137–151. Available from [http://dx.doi.org/10.1007/BF00342882](http://dx.doi.org/10.1007/BF00342882), doi:10.1007/BF00342882

Mooney, S., & Anderson, B. (2014). Specular image structure modulates the perception of three-dimensional shape. *Current Biology, 24*(22), 2737 - 2742. Available from [http://www.sciencedirect.com/science/article/pii/S0960982214012706](http://www.sciencedirect.com/science/article/pii/S0960982214012706), doi:https://doi.org/10.1016/j.cub.2014.09.074

Nefs, H. T., Koenderink, J. J., & Kappers, A. M. L. (2005). The influence of illumination direction on the pictorial reliefs of lambertian surfaces. *Perception, 34*(3), 275-287. Available from [http://dx.doi.org/10.1068/p5179](http://dx.doi.org/10.1068/p5179) (PMID: 15895627), doi:10.1068/p5179

Pasupathy, A., & Connor, C. E. (2002). Population Coding of Shape in Area V4. *Nature Neuroscience: Published Online.*

Paulignan, Y., Frak, V., Toni, I., & Jeannerod, M. (1997). Influence of object position and size on human prehension movements. *Experimental brain research, 114*(2), 226–234.

P. Belhumeur, D. K., & Yuille, A. (1999). The Bas-Relief Ambiguity. *International Journal of Computer Vision, 35*, 33–44.
Prados, E., & Faugeras, O. (2006). Shape from shading. Published in N. Paragios, Y. Chen, O. Faugeras (eds), Handbook of Mathematical Models in Computer Vision, 375 - 388.

Przybylski, M., Asfour, T., & Dillmann, R. (2011). Planning grasps for robotic hands using a novel object representation based on the medial axis transform. In Intelligent robots and systems (iros), 2011 ieee/rsj international conference on (pp. 1781–1788).

Ramachandran, V. S. (1988). Perception of shape from shading. Nature, 331, 163–166.

Reininghaus, J., & Hotz, I. (2011). Combinatorial 2d vector field topology extraction and simplification. In V. Pascucci, X. Tricoche, H. Hagen, & J. Tierny (Eds.), Topological methods in data analysis and visualization: Theory, algorithms, and applications (pp. 103–114). Berlin, Heidelberg: Springer Berlin Heidelberg. Available from http://dx.doi.org/10.1007/978-3-642-15014-2_9

Sahner, J., Weber, B., Prohaska, S., & Lamecker, H. (2008). Extraction of Feature Lines on Surface Meshes based on Discrete Morse Theory. IEEE Symposium on Visualization, 27(3).

Seyama, J., & Sato, T. (1998). Shape from shading: estimation of reflectance map. Vision Research, 38(23), 3805 - 3815. Available from http://www.sciencedirect.com/science/article/pii/S0042698997004355 , doi:https://doi.org/10.1016/S0042-6989(97)00435-5

Spier, A. J., Liarokapis, M. V., Calli, B., & Dollar, A. M. (2016, April). Single-grasp object classification and feature extraction with simple robot hands and tactile sensors. IEEE Transactions on Haptics, 9(2), 207-220. , doi:10.1109/TOH.2016.2521378

Stevens, K. A. (1982). Implementation of a Theory for Inferring Surface Shape from Contours. MIT Memo No. 676.

Stevens, K. A. (1983, March). Slant-tilt: The visual encoding of surface orientation. Biol. Cybern., 46(3), 183–195. Available from http://dx.doi.org/10.1007/BF00336800 , doi:10.1007/BF00336800

Sun, J., & Perona, P. (1998). Where is the sun? Nature neuroscience, 1(3), 183–184.

Sun, P., & Schofield, A. J. (2012). Two operational modes in the perception of shape from shading revealed by the effects of edge information in slant settings. Journal of Vision, 12(1), 12. Available from +http://dx.doi.org/10.1167/12.1.12 , doi:10.1167/12.1.12

Tang, Y., Salakhutdinov, R., & Hinton, G. (2012, June). Deep Lambertian Networks. ArXiv e-prints.

Todd, J. T., Egan, E. J. L., & Phillips, F. (2014). Is the perception of 3d shape from shading based on assumed reflectance and illumination? i-Perception, 5(6), 497-514. Available from http://dx.doi.org/10.1088/1068/10645 , doi:10.1068/i0645

Waltz, D. (1975). Understanding line drawings of scenes with shadows. In The psychology of computer vision (p. pages). McGraw-Hill.

Weintraub, T., Ginolld, Y., & Sorkine, O. (2010). Topology-based Smoothing of 2D Scalar Fields with C1-Continuity. IEEE Symposium on Visualization.

Wijntjes, M. W. A., Doerschner, K., Kucukoglu, G., & Pont, S. C. (2012). Relative flattening between velvet and matte 3d shapes: Evidence for similar shape-from-shading computations. Journal of Vision, 12(1), 2. Available from +http://dx.doi.org/10.1167/12.1.12 , doi:10.1167/12.1.2

Yamane, Y., Carlson, E. T., Bowman, K. C., Wang, Z., & Connor, C. E. (2008). A Neural Code for Three-Dimensional Object Shape in Macaque Inferotemporal Cortex. Nature Neuroscience: Published Online.

Zheng, Q., & Chappella, R. (1991). Estimation of illuminant direction, albedo, and shape from shading. IEEE Transactions on Pattern Analysis and Machine Intelligence, 13, 680–702.