TOPOLOGICAL FEATURES ANCHOR VISUAL SHAPE REPRESENTATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Computer vision is outstanding on fixed datasets, but lacks the effortless generality and flexibility of human perception. As a step toward understanding this flexibility, we show that there are features – critical contours – readily detectable in images that have counterparts – extremal contours – on 3D shapes. Both are part of a topological structure, Morse-Smale complex, and together they provide an invariant linking image to shape structure. An open neighborhood around these contours is related to human perception of shape features, such as bumps, which suggests that training for these features could provide a generalization benefit in algorithm development.

1 INTRODUCTION

Vision and robotics provide clues about what is where. Increasingly applications, such as object manipulation, are pushing representations toward three spatial dimensions. Handles, for example, provide grasping points, but systems capable of identifying such extended "bumps" remains a challenge ten Pas et al. (2017), even when working with 3D point clouds. We propose a topological feature ideally suited for such tasks based on the 2D image.

Specifying what should be represented in current approaches, and why, is often instructed by classical algorithms. For example, keypoint detection (classic approach: Lowe (2004); current deep network Barnes & Posner (2020)) is used to locate objects, or features on objects, and surface normals are used to describe shape (classic: Horn (1986); Horn & Brooks (1989); current: Eigen & Fergus (2014); Wang et al. (2020)). But keypoints are intuitive and heuristic, and surface normal inference is ill-posed. The latter is the classical shape-from-shading problem, for which regularizers or shape priors Kersten et al. (2004); Barron & Malik (2012); Chakrabarti et al. (2016) are invoked. These can be powerful in restricted contexts, e.g. faces Kulkarni et al. (2015) or hands Ge et al. (2019), but are brittle and, when generalized, tend to over regularize Zhou et al. (2017). Multiple views help Sridhar et al. (2019) but reconstructions are still category specific. This holds for approaches using voxel grids or polygonal meshes as well Su et al. (2015); Wang et al. (2018). We propose that a path out of these rigid, constrained corners of shape space that involves postulating an intermediate level of representation. This intermediate level is generic and based on topological constructs.

To motivate our approach, consider the occluding contour. It is special because it enjoys both salient image domain and shape domain structure (Fig. 1). Note in particular how the concentrated intensities imply the surface normal is orthogonal to the line of sight. Others have used this implicitly (e.g.Wang et al. (2018)); we go a step further by defining interior contours that also enjoy image and shape salience. Our goal is to replace empirical interior features with generic ones.

2 CRITICAL CONTOURS AND EXTREMAL CURVES

The image signature of occluding contours is distinctive; away from them, the variation of intensities with viewpoint or lighting changes is endlessly varying almost everywhere (Fig. 2(a). Some schemes exploit this variation Eslami et al. (2018) – they work when structure among the queries can be identified as in CAD models Lei et al. (2020). But this is limited. We choose the opposite – we seek those positions that are invariant across the queries (Fig. 2(a, bottom), at least for generic (Morse) smooth surfaces.

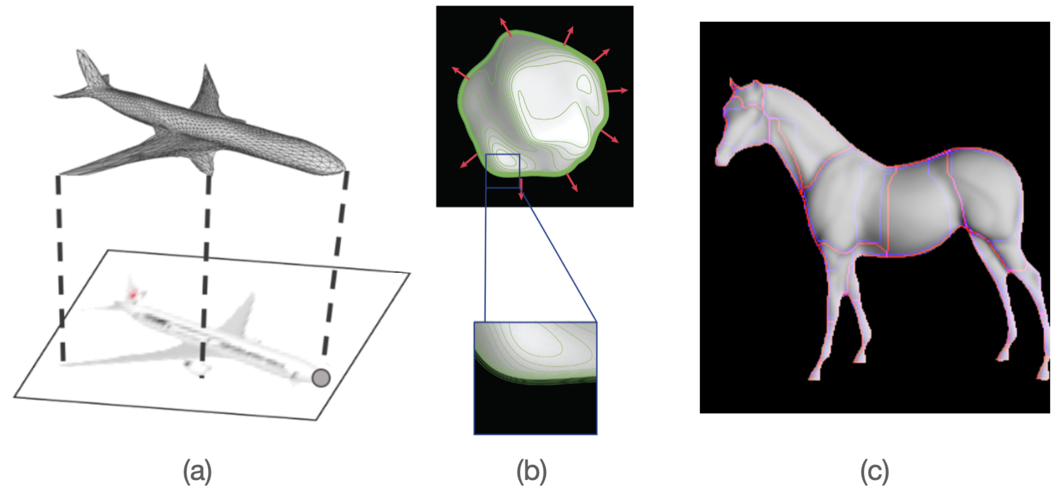

Figure 1: Anchoring shape inferences. (a) Algorithms for simple objects evolve from the occluding contour. (b) Level sets of intensity, or isophotes, concentrate in parallel along the occluding contour, while the (projected) surface normal (red arrows) is orthogonal to it. This links the two domains. (c) Our proposal: critical contours (blue) can be derived from the image; they imply the existence of extremal curves of surface slant (red). These are interior features that ground surface constraints in image structure. Fig. (a) adapted from Wang et al. (2018).

To be flexible, an invariant must work across different rendering functions, such as Lambertian shading, specular shading, texture, etc., as well as line drawings DeCarlo et al. (2003); Judd et al. (2007). Importantly, all of these renderings involve the surface normal field to create an image: View the surface normal field $N(x, y)$ as a map from the image domain $\mathbb{R}^2$ to the unit sphere S^2. A rendering function $F \in \mathcal{F}$ is then a smooth map from the unit sphere to the real line, that is $F : S^2 \to \mathbb{R}$, and the image $I(x, y) = F(N(x, y))$. Building on Kunsberg & Zucker (2018), the *critical contour* (blue curve in Fig. 2(b) is defined

DEFINITION *Critical contours are gradient flows in the image with large transversal second derivatives.*

Critical contours are invariant to the choice of F:

THEOREM *Given a surface normal field $N(x, y)$ and any two choices of generic rendering functions $F_1, F_2 \in \mathcal{C}$, construct $I_1 = F_1(N(x, y)), I_2 = F_2(N(x, y))$. If a critical contour is present in I_1, then there is a arbitrarily close critical contour in I_2.*

We now turn to the surface. The orientation of a normal vector at a point can be specified in slant σ and tilt τ coordinates. Since there is a normal at every point, the slant coordinate defines a scalar function over the object. In these coordinates the occluding contour can be defined by maximal values of the slant along it. But in the interior of the shape the slant can no longer be critical everywhere, or it would involve total self occlusion. This motivates a relaxation:

DEFINITION *Slant extremal contours are the saddle-maxima 1-cells of the MS complex of the slant function.*

A corollary can be restated:

COROLLARY *An extremal contour must lie in the tubular neighborhood of a critical contour and have the same endpoints.*

Extremal contours lie near the critical contours, a proof of which follows since slant can be viewed formally as a rendering function $F \in \mathcal{F}$. Further, the invariance statements from Kunsberg & Zucker (2018) show that extremal contours will nearly always be salient from the image, regardless of the rendering function. The main point is this: *Extremal contours delineate the surface features while the critical contours are salient from the image. We can observe a critical contour, infer an extremal contour near it, and then use the theorem to attribute surface protrusions to image*

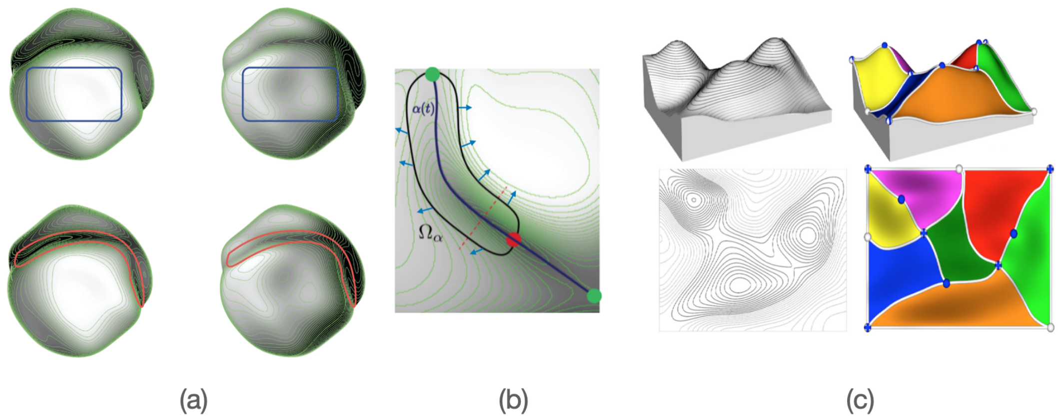

Figure 2: Distinctive features of image intensities are topological. (a) (top) The isophotes move significantly with changes in lighting, even for a lambertian shape; compare the patterns in the blue squares with those in the red sausages (bottom). In neighborhoods near ridges the isophote pattern barely changes. (b) These red sausages are open neighborhoods containing critical contours, an integral curve of the gradient flow. The arrows surrounding it point in the positive gradient direction, and the red and green dots are singularities (see Appendix). (c) The critical contours are part of the Morse-Smale complex. Note how, when shown in perspective (top), the 1-cells (contours) resemble an artist's line drawing. (c) adapted from Gyulassy (2008).

regions. Moreover, because the Corollary implies an image signature, we conjecture that this should be readily learnable. Experiments in our group are now confirming this.

3 Bumps: closed extremal contours

A special case, closed extremal contours, illustrates the power of the Theorem. Recall that we began with the task of finding bumps (or protrusions for grasping)[1]. Consider, for example, a sigmoidal bump and ask yourself where it begins. Although in robotics one seeks exact quantitative locations, we show (Fig. 3(a)) that humans are quite variable in their estimates of 3D shape; we say they are qualitative rather than quantitative Nartker et al. (2017). This is typical of perception Sun & Schofield (2012); 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); Koenderink et al. (1996) show similar results.

In other words, we all agree there are bumps when viewing a shape, but the precise location of where they begin is problematic. We cartoon this variability with the yellow 'sausage' in 3(c); a technical version is shown in (d), which surrounds a (piece of) critical contour. Note how the gradient points away from the sausage, the entry and the exit sides contain extrema, and nearby is an extremal contour (dashed). Importantly, critical contours and extremal contours are integral curves through the gradient flow; i.e. are orthogonal to the level sets. In the limit as the 'sides' of the bump get steeper the contours converge toward one another, just as an artist's drawing would convey shape. It is this construction that underlies the proof that critical and extremal contours are 1-cells in the Morse-Smale complex (see Appendix for introduction to Morse-Smale).

Examples of the critical contour in the image and the extremal contour on the surface are shown in blue for a sigmoidal bump. Because of the global nature of extremal contours, examples were computed using persistence simplification Reininghaus & Hotz (2011). In this artifical example they coincide, but for the more natural 'horse' example (Fig. 1(c)) they lie near each other (within the sausage neighborhood). In addition we show the attached portions of the Morse-Smale complex. Note how the extremal contour passes through the max in slant (and contains the minimum inside);

[1]Due to space limitations, we do not consider whether the bump is convex or concave in this abstract; all statements are correct in both cases. In addition, there are constraints between bumps and dents, that we also do not discuss here.

this signature is the relaxation of the occluding contour to the interior of the shape (see Appendix for development of the Morse-Smale cell for a bump.)

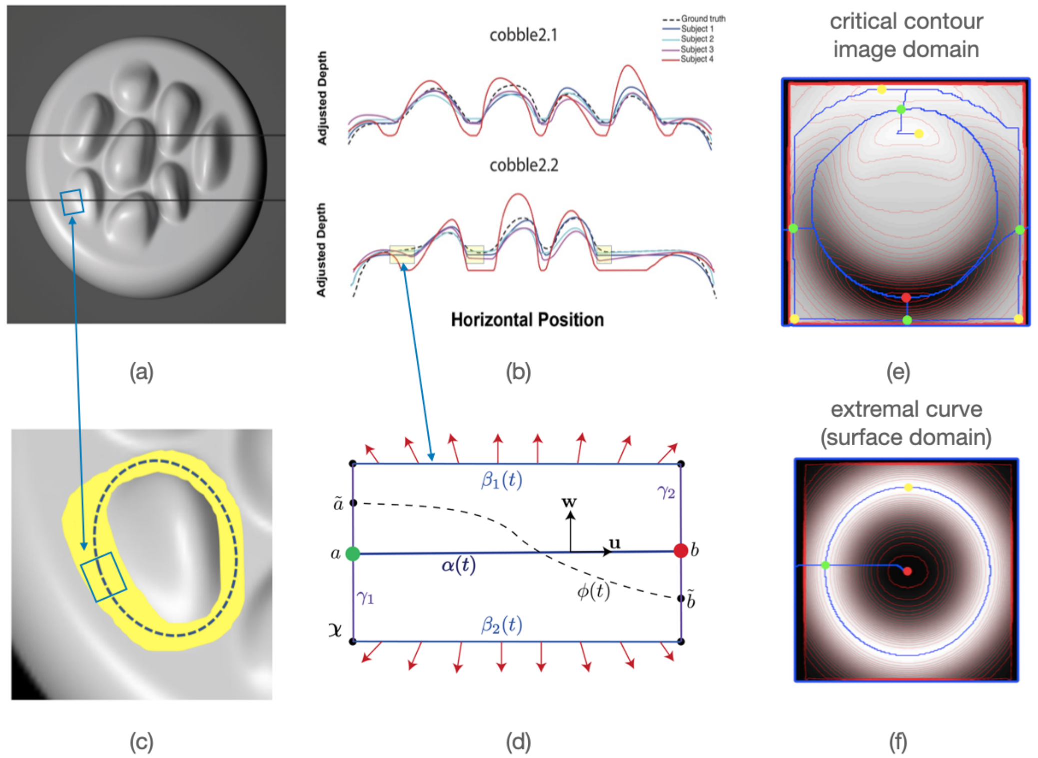

Figure 3: Overview of argument. (a) Human perception of shape is qualitative. (b) Different subjects see different depths when viewing a series of shaded bumps. (c) We cartoon these differences by placing a neigborhood over the loci of positions where a bump could begin. The theorem (d) states that for any critical contour (blue) in the sausage neighborhood there will be an extremal contour (dashed) nearby. (e) The Morse-Smale complex for the image and (f) slant function for a bump. The crtical and (closed) extremal contours are shown in blue. Fig. (a) from Nartker et al. (2017)).

4 Summary and Conclusions

The problems associated with shape inference are currently a barrier to flexible vision algorithms. Physically-crafted priors, smoothing regularizers, and task-limiting training datasets, have led to algorithms that are brittle and difficult to generalize. We argue that there is something basic about the qualitative nature of human shape perception, and that this variability is naturally expressed in topological terms. Critical contours and extremal curves of slant are a specific realization of these topological terms, and provide a generic invariant linking the image domain to the shape domain. Incorporating these topological constraints into recognition algorithms could extend their flexibility and generalizability. It is perhaps no accident that our hands were designed as qualitative grippers; only after our fingers tighten are the exact, geometric parameters known.

Acknowledgments

Supported by the Simons Collaboration on the Global Brain, the Allen Distinguished Investigator Program through The Paul G. Allen Frontiers Group, and the National Science Foundation Collaborative Research in Computational Neuroscience Program.

REFERENCES

L. Allemand-Giorgis, G-P. Bonneau, and S. Hahmann. Piecewise Polynomial Reconstruction of Functions from Simplified Morse-Smale complex. *IEEE Visualization Conference*, 2014.

Dan Barnes and Ingmar Posner. Under the radar: Learning to predict robust keypoints for odometry estimation and metric localisation in radar. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 9484–9490. IEEE, 2020.

Jonathan T Barron and Jitendra Malik. 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, 2012.

S. Biasotti, L. De Floriani, B. Falcidieno, P. Frosini, D. Giorgi, C. Landi, L. Papaleo, and M. Spagnuolo. Describing shapes by geometrical-topological properties of real functions. *ACM Comput. Surv.*, 40(4):12:1–12:87, October 2008. ISSN 0360-0300. URL `http://doi.acm.org/10.1145/1391729.1391731`.

Ayan Chakrabarti, Jingyu Shao, and Gregory Shakhnarovich. Depth from a single image by harmonizing overcomplete local network predictions. *CoRR*, 2016. URL `http://arxiv.org/abs/1605.07081`.

Chris G. Christou and Jan J. Koenderink. Light source dependence in shape from shading. *Vision Research*, 37(11):1441 – 1449, 1997. ISSN 0042-6989.

William Curran and Alan Johnston. The effect of illuminant position on perceived curvature. *Vision Research*, 36(10):1399 – 1410, 1996. ISSN 0042-6989.

Doug DeCarlo, Adam Finkelstein, Szymon Rusinkiewicz, and Anthony Santella. Suggestive contours for conveying shape. *ACM Trans. Graph.*, 22(3):848–855, 2003.

Eric J. L. Egan and James T. Todd. 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, 2015.

David Eigen and Rob Fergus. Predicting depth, surface normals and semantic labels with a common multi-scale convolutional architecture. *CoRR*, 2014. URL `http://arxiv.org/abs/1411.4734`.

SM Ali Eslami, Danilo Jimenez Rezende, Frederic Besse, Fabio Viola, Ari S Morcos, Marta Garnelo, Avraham Ruderman, Andrei A Rusu, Ivo Danihelka, Karol Gregor, et al. Neural scene representation and rendering. *Science*, 360(6394):1204–1210, 2018.

W. Freeman. The generic viewpoint assumption in a framework for visual perception. *Nature*, pp. 542 – 545, 1994.

Liuhao Ge, Zhou Ren, Yuncheng Li, Zehao Xue, Yingying Wang, Jianfei Cai, and Junsong Yuan. 3d hand shape and pose estimation from a single rgb image. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 10833–10842, 2019.

Attila Gabor Gyulassy. *Combinatorial construction of Morse-Smale complexes for data analysis and visualization*. PhD thesis, University of California, Davis, 2008.

Daniel N Holtmann-Rice, Benjamin S Kunsberg, and Steven W Zucker. Tensors, differential geometry and statistical shading analysis. *Journal of Mathematical Imaging and Vision*, 60(6):968–992, 2018.

Berthold Horn. *Robot Vision*. The MIT Press, Cambridge, Massachusetts, 1986.

B.K.P. Horn and M.J. Brooks. *Shape from Shading*. The MIT Press, Cambridge, Massachusetts, 1989.

Tilke Judd, Fredo Durand, and Edward H. Adelson. Apparent ridges for line drawing. *ACM Trans. Graph.*, 26(3):19, 2007.

Daniel Kersten, Pascal Mamassian, and Alan Yuille. Object perception as bayesian inference. *Annu. Rev. Psychol.*, 55:271–304, 2004.

Byung-Geun Khang, Jan J Koenderink, and Astrid M L Kappers. Shape from shading from images rendered with various surface types and light fields. *Perception*, 36(8):1191–1213, 2007. URL `http://dx.doi.org/10.1068/p5807`.

Jan J Koenderink, Andrea J van Doorn, Chris Christou, and Joseph S Lappin. Perturbation study of shading in pictures. *Perception*, 25(9):1009–1026, 1996. URL `http://dx.doi.org/10.1068/p251009`.

Tejas D Kulkarni, William F Whitney, Pushmeet Kohli, and Josh Tenenbaum. Deep convolutional inverse graphics network. In *Advances in neural information processing systems*, pp. 2539–2547, 2015.

Benjamin Kunsberg and Steven W Zucker. Critical contours: An invariant linking image flow with salient surface organization. *SIAM Journal on Imaging Sciences*, 2018.

Jiahui Lei, Srinath Sridhar, Paul Guerrero, Minhyuk Sung, Niloy Mitra, and Leonidas J Guibas. Pix2surf: Learning parametric 3d surface models of objects from images. In *European Conference on Computer Vision*, pp. 121–138. Springer, 2020.

David G Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004.

Pascal Mamassian and Daniel Kersten. Illumination, shading and the perception of local orientation. *Vision Research*, 36(15):2351 – 2367, 1996.

Yukio Matsumoto. *An introduction to Morse theory*, volume 208. American Mathematical Soc., 2002.

John Milnor. *Morse Theory*. Princeton University Press, 1963.

E. Mingolla and J. T. Todd. Perception of solid shape from shading. *Biological Cybernetics*, 53(3): 137–151, 1986.

M. Nartker, J. Todd, and A. Petrov. Distortions of apparent 3d shape from shading caused by changes in the direction of illumination. *Journal of Vision*, 17(324), 2017.

Jan Reininghaus and Ingrid Hotz. Combinatorial 2d vector field topology extraction and simplification. *IEEE Trans Vis Comput Graph.*, pp. 1433–1443, 2011.

Jun'ichiro Seyama and Takao Sato. Shape from shading: estimation of reflectance map. *Vision Research*, 38(23):3805 – 3815, 1998.

Srinath Sridhar, Davis Rempe, Julien Valentin, Sofien Bouaziz, and Leonidas J Guibas. Multiview aggregation for learning category-specific shape reconstruction. *arXiv preprint arXiv:1907.01085*, 2019.

Hang Su, Subhransu Maji, Evangelos Kalogerakis, and Erik Learned-Miller. Multi-view convolutional neural networks for 3d shape recognition. In *Proceedings of the IEEE international conference on computer vision*, pp. 945–953, 2015.

Peng Sun and Andrew J. Schofield. 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, 2012.

Andreas ten Pas, Marcus Gualtieri, Kate Saenko, and Robert Platt Jr. Grasp pose detection in point clouds. *CoRR*, abs/1706.09911, 2017. URL `http://arxiv.org/abs/1706.09911`.

James T. Todd, Eric J. L. Egan, and Flip Phillips. Is the perception of 3d shape from shading based on assumed reflectance and illumination? *i-Perception*, 5(6):497–514, 2014. URL `http://dx.doi.org/10.1068/i0645`.

Nanyang Wang, Yinda Zhang, Zhuwen Li, Yanwei Fu, Wei Liu, and Yu-Gang Jiang. Pixel2mesh: Generating 3d mesh models from single rgb images. In *Proceedings of the European Conference on Computer Vision (ECCV)*, September 2018.

Rui Wang, David Geraghty, Kevin Matzen, Richard Szeliski, and Jan-Michael Frahm. Vplnet: Deep single view normal estimation with vanishing points and lines. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 689–698, 2020.

T. Weinkauf, Y. Gingold, and O. Sorkine. Topology-based Smoothing of 2D Scalar Fields with C1-Continuity. *IEEE Symposium on Visualization*, 2010.

Tinghui Zhou, Matthew Brown, Noah Snavely, and David G. Lowe. Unsupervised learning of depth and ego-motion from video. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.

A APPENDIX

A.1 INTRODUCTION TO THE MORSE SMALE COMPLEX

The Morse Smale complex is a qualitative representation emphasizing the different stable and unstable regions of a smooth scalar function. In this work, we choose the function to be the slant function of the image surface $\sigma(x, y) : \mathbb{R}^2 \to \mathbb{R}$. We will assume σ is a *Morse function*: all its critical points are non-degenerate (meaning the Hessian at those points is non-singular) and no two critical points have the same function value.

For a smooth surface, the *gradient* $\nabla\sigma = (\partial f/\partial x, \partial f/\partial y)$ exists at every point. A point $p \in \mathbb{R}^2$ is called a *critical point* when $\nabla\sigma(p) = 0$. This gradient field gives a direction at every point in the image, except for the critical points, a set of measure zero. Following the vector field will trace out an *integral line*. These integral lines must end at critical points, where the gradient direction is undefined. Thus, one can define an *origin* and *destination* critical point for each integral line.

The type of 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: a maximum (with index 2), a minimum (with index 0) and a saddle point (with index 1).

There are two types of integral lines, depending on the difference in index of the critical points it connects. If the difference is one, we call the integral line a *1-cell*. It naturally must connect a saddle with either a maximum or a minimum. For example, a *saddle-maxima 1-cell* connects a saddle and a maximum. The set of 1-cells will naturally segment the scalar field into different regions, called *2-cells*. In addition, the scalar values on the 1-cells govern the values on the 2-cells. See Fig 2(c) for insight.

Further, for each critical point, its *ascending manifold* is defined as the union of integral lines having that critical point as a common origin. Similarly, its *descending manifold* is the union of integral lines with that critical point as a common destination.

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 watershed) 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 $\mathbb{M}$ into 2D regions surrounded by 1D curves with intersections at the critical points.

The Morse Smale complex is the combinatorial structure (and the corresponding attaching maps) defined by the critical points, 1-cells and 2-cells. It is a structure that relates a set of contours (the 1-cells) to a qualitative function representation. With knowledge only of the slant 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 et al. (2014); Weinkauf et al. (2010). In this work, we wish to show how the slant saddle-maxima 1-cell can be used as a ‘bump boundaries’ to model 3D shape perception.

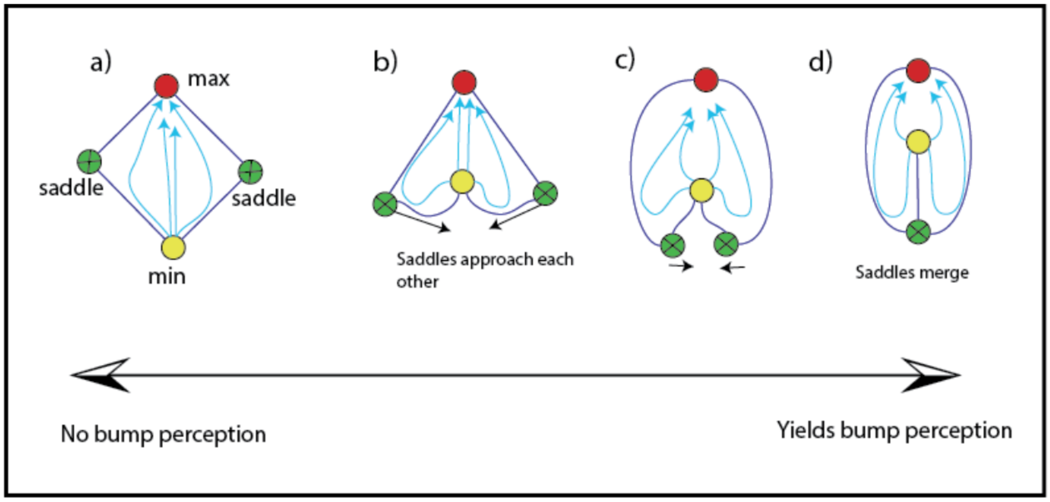

Figure 4: Development of the MS-complex for bumps. (a) Starting with the standard MS cycle, (b) smoothly deform the domain so the saddles begin to (c) approach one another until (d) they merge. Notice how the integral curves (cyan) deform along with the saddles.

For additional information, see Milnor (1963); Gyulassy (2008); Biasotti et al. (2008); Matsumoto (2002).

A.2 Extremal Rings and Occluding Contours

We first show how the standard 2-cell of the Morse-Smale complex specializes in the case of bumps; see Fig. 4. It is basically a situation in which two saddles merge into one. Notice how this further "pulls" the gradient flow along with the merging; this helps to further define the image signature.

We next show that extremal contours represent boundaries of bumps and valleys of the surface. The argument will be based on the conclusion that the surface normal field should generically point uniformly to the interior (or exterior) of the region bounded by the extremal contours.

Consider Fig 5, and focus on the enlarged portion of the extremal ring. This was defined to cut across the level sets of the slant function but, by definition, it does not inform the tilt function. Here we show that, for generic surfaces, the tilt must be constrained along the extremal ring to prevent self occlusion and image instability. There are two basic possibilities: in the first case, the surface behaves like a bump boundary, where the surface normal on the extremal ring points consistently; the second case leads to rather wild surfaces with 'crazy' curvatures. While this second case is a mathematical possibility (developed next), it violates our requirement that the surface be generic. That is, a small change in the lighting would lead to a drastically different image (Fig. 6). For the wild surfaces, the normal points in many different directions, covering a very large portion of the Gauss map.

More technically, let $\alpha(t)$ denote an extremal ring that bounds a region R, and let $\{x, y\}$ be image coordinates under orthographic projection. Let $\sigma(x, y)$ represent the slant in a neighborhood of a point (x_0, y_0) on $\alpha(t)$. Rotate the frame so that the slant gradient (tangent direction to $\alpha(t)$) points locally in the y direction. We compare two possible solutions for the local surface depth $S(x, y)$ by defining two pairs of Taylor expansions for the slant and tilt functions. Let $\sigma_1(x, y) = \sigma_2(x, y) = c_1 + c_2 y + \sigma_{xx} x^2 + \sigma_{xy} xy + \sigma_{yy} y^2$. (There is no linear x term here due to the slant gradient pointing along the y axis.) Let $\tau_1 = x$ and $\tau_2 = y$. Now, $\{\sigma_1, \tau_1\}$ defines a surface S_1 and $\{\sigma_2, \tau_2\}$ defines an alternative surface S_2. Which of these is more likely?

Both S_1 and S_2 have the same magnitude tilt gradient. However, S_1 (generic surface) has the tilt change along the contour while S_2 (twisted surface) has a tilt change perpendicular to the extremal contour.

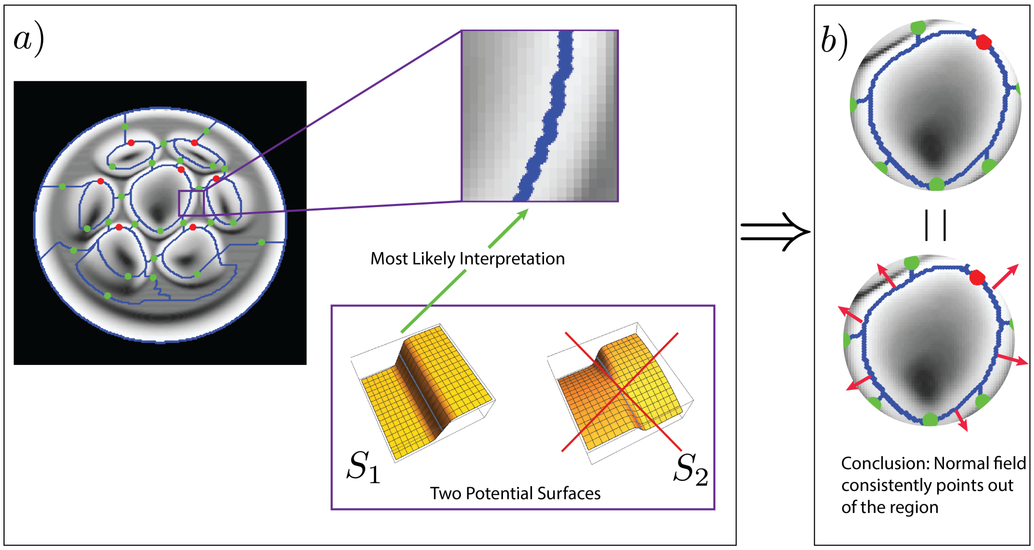

Figure 5: Extremal curves and surface normals. (a): Consider a portion of a maximal curve of slant. Two distinct types of local surfaces could have caused this local slant function, one with no twist and another with substantial twist. The first surface, without twist, is much more likely than the second surface. Note the normal field for the first surface consistently points to the same side (left) of the extremal curve. (b) Taking the most likely interpretation for each portion of the extremal curve, the normal field must point uniformly outside (or inside) the entire contour.

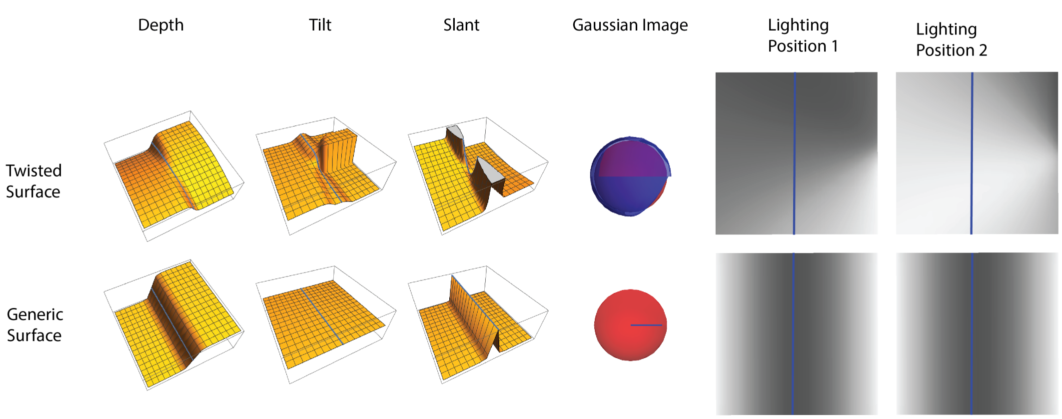

Figure 6: Comparing possible surface explanations for a maximal slant curve. Each row depicts different properties of the surfaces shown in the first column. Although both are technically solutions, the twisted surface covers most of the Gauss map and, when rendered under slightly different lightings, gives rise to drastically different images. The generic surface solution occupies a small portion of the Gauss map and renders almost the same image under different light sources. Both arguments show that the twisted surface (top row) is much less likely.

Let $N_1(x, y), N_2(x, y)$ represent the normal fields for each of these two solutions. We compare the relative probability of each of these surfaces by considering the term $T_i = \int_\Omega \det(DN_i DN_i^T)$ for each surface. This is essentially the Gaussian curvature of each solution integrated over the patch.[2] The solution with higher Gaussian curvature will be the solution that is less smooth, more dependent on lighting direction Freeman (1994) and with a higher chance of occlusion. We compare the relative likelihood of the two solutions $\frac{L_1}{L_2}$ by considering the inverse of the ratio $\frac{T_2}{T_1}$. A simple calculation shows:

$$\frac{L_1}{L_2} \propto \frac{T_2}{T_1} \propto \frac{\sigma_{xx}^2}{\sigma_y^2} \tag{1}$$

where σ_{xx} is the transversal second derivative of the slant across $\alpha(t)$ while σ_y is the gradient along $\alpha(t)$. Since the slant on α is extremal, its gradient will necessarily be small. In addition, since the slant is changing rapidly across α, σ_{xx} will be large. Thus, for a slant patch as shown in Fig 5(a), the ratio $\frac{\sigma_{xx}^2}{\sigma_y^2}$ will be large. This statement is illustrated by the comparison in Fig 6. We conclude that the image patch has a surface normal field that is not twisted. In other words, it is most likely that the normal points to a single side of the curve in the entire Taylor expansion.

Applying the above argument completely around the extremal ring $\alpha(t)$ shows that the most probable interpretation of the surface normal field along $\alpha(t)$ is that it does not have a twist and therefore must point uniformly outside (or inside) the region bounded by $\alpha(t)$. It then follows that the region R is 'higher' (or 'lower') than the surrounding area. More precisely, one can see that the R must be an ascending or descending manifold of depth; in other words the region is a bump or a valley.

REMARK *The surface normal along an extremal contour points consistently to the interior or exterior.*

[2]For an explanation of these terms, see Holtmann-Rice et al. (2018).