Understanding Qualitative 3D Shape from Texture and Shading

Benjamin Kunsberg and Steven W. Zucker

What are the visual structures that encode 3D perception? How can we see reliable 3D shape in such a diverse range of cues like texture, shading, and line drawings? The usual attempt to recover 3D shape proceeds by isolating the cue, then applying inverse optics based on that cue and a cost function minimization. Although this may sometimes lead to decent reconstructions, it requires a source separation and precise knowledge of the cue and the way it renders an image from the surface. For example, in Lambertian shading, the intensity field (cue) is modeled by the dot product of the light source and the normal field. If this Lambertian shading model is wrong, as it invariably will be in natural scenes, a cost function minimization will be using inappropriate penalty terms and thus may end up with an arbitrarily wrong minimizer. However, many psychophysics studies show that qualitative shape constancy is perceived under diverse illumination conditions.

We are developing a theory that can formally deal with this conundrum while also explaining electrophysiological data such as that found in [2]. We believe the key lies in analyzing the topological properties of the orientation field of the image. We believe the Morse Smale complex (calculated via the orientations and consisting of multiple contours) provides the image contours that 'anchor' the surface perception. Last year, we proved a theorem that showed that the Morse Smale complex remains stable when illumination conditions vary. Now, we take it one step further: we show that the stable Morse Smale complex can also be computed from textured images. We do this by computing orientations via filter responses maximized over scale such as in [1] and then applying the same topological inference. Thus, textured, shaded, and line drawing images can be united under a common model.

In addition, we will demonstrate how the Morse Smale complex can be used to recover qualitative 3D shapes, such as a segmentation of the surface into regions consisting of concave (bumps), convex (valleys) and hyperbolic (slopes) as in Figure 1. Thus, we will describe a computational model with the following properties:

- Computation on biologically accurate units
- Resulting in stable (regardless of the cue) intermediate features yielding
- Qualitative 3D representations matching psychophysical data

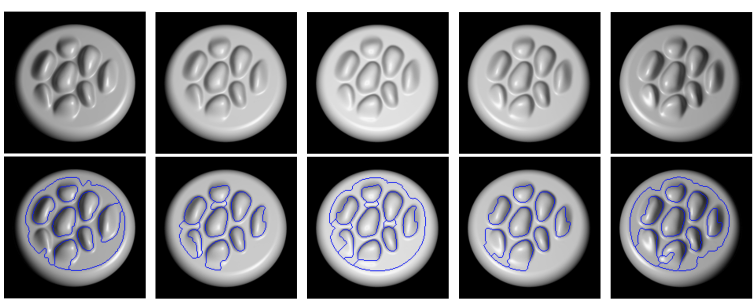

Figure 1: First Row: Images of a 'cobblestone' with different lightings and specularity (courtesy of James Todd.) Second Row: We use the MS complex of the orientation field to find convex bumps according to the above theory. Highlights may still cause some issues, but note the common segmentations. Every convex bump is captured in every image but the first.

References

[1] Elder, J. H. Are edges incomplete? *International Journal of Computer Vision 34*, 2 (Aug 1999), 97–122.

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