A Single Shape from Multiple Cues: How Local and Global Information Organizes Shape Inference

Benjamin Kunsberg and Steven W. Zucker

Where should I grasp the cup? Which point on the rim is closer? These questions force a dichotomy over much of shape research. Blobs, bumps, and handles underlie the manipulation of a shape, while surface normals and depth underlie the description of a shape. The former are global representations; the latter can be local. The former are qualitative; the latter are quantitative. But both can be inferred from natural images or drawings, and both are (approximately) invariantly perceived over a variety of lighting and material contexts. How does one consistently see the same qualitative parts (such as a bump) regardless of the cue type? How do we rigorously describe a "bump" in the image and integrate pointwise properties into such an intermediate part?

We are developing a theory of shape inference that formally unifies these questions. Rather than seeking the 'unique' shape from a shaded image under given conditions, which is ill-posed, we seek an abstraction over the image and over the surface that are related to one another. Starting with the observation that smooth images are represented in visual cortex as orientation (shading) flows, we study (qualitative) invariants over flows. An important one arises with the gradient flow – the Morse-Smale (MS) complex – which separates the flow into cells according to singularities (maxima, minima, saddles). The abstraction is topological, and consists of parts (2-cells) defined by curves (1-cells) that link the singularities (0-cells) together. The 2-cells provide an organization of the image (into parts) that relates one-to-one with the organization of (the slant of) the surface into parts (Fig. 1). Referring back to our title, the flow over the whole is dictated by the flow over the parts under the control of the singularities.

Image flows concentrate in certain cases, which we call the shading-contour limit. The convergence is pointwise to image contours, and identifies the common ground between shaded images and artist's drawings. It holds precisely for certain orientation patterns, such as those that arise along ridges, and is stable with respect to reasonable shading and texture rendering functions. It is this limiting process that allows us to formally relate the image structure to (certain) 1-cells in the MS complex. In the end, it is precisely these 1-cells that form the anchors on which full shapes can be based, and which constrain, we predict, the psychophysical properties of much of shape perception.

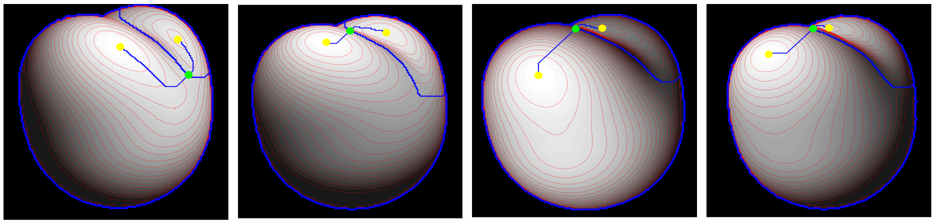

Figure 1: Note the equivalent graph structure regardless of the changes in light source (which yield drastic changes in the image.) The MS complex structure (shown with nodes in yellow and green, 1-cells in blue, isophotes in red) segments the surface into topological parts.