Current Biology, Vol. 13, R234–R235, March 18, 2003, ©2003 Elsevier Science Ltd. All rights reserved. DOI 10.1016/S0960-9822(03)00159-3

Shape Perception: Complex Contour Representation in Visual Area V4

Dispatch

Kate Gustavsen[1] and Jack L. Gallant[2]

A recent study has put forward a physiologically plausible population model that implements a parts-based shape-coding scheme for macaque visual area V4.

When perceiving a scene, the visual system must knit together responses from thirty or more interconnected brain areas, each composed of intricate networks of neurons. To understand how any of these areas represents a stimulus, it would be desirable to record from many neurons simultaneously. While *in situ* recordings from single neurons have been possible for decades, the technology necessary to achieve stable simultaneous recordings from multiple visual neurons remains elusive. Population activity can, however, be modeled mathematically using single neuron data. Pasupathy and Connor [1] have recently used this technique to demonstrate the feasibility of their theory of shape coding in visual area V4, an intermediate stage in the shape processing pathway of the visual system.

Visual analysis of shape is performed by a hierarchically organized series of brain areas (for a brief review, see [2]). The representation in each area, and the response selectivity or tuning of individual neurons, increases in complexity from the earliest to the latest stages of processing. The first stage of cortical visual processing is area V1, which represents local image characteristics such as orientation and spatial scale. Form and color information pass from V1 to area V2, and from there to V4. Neurons in these intermediate visual areas appear to be tuned to respond to shape information of moderate complexity, such as illusory contours and curvature. Shape information eventually reaches the inferior temporal cortex, where complex shapes and even entire objects appear to be encoded in the responses of single neurons.

Area V4 is a crucial link between simple form analysis and the perception of complete scenes; it must mediate the transformation between the local image analysis of earlier visual areas and the object-based representations of later stages of processing. Yet there is little consensus on the principles by which V4 neurons actually encode visual stimuli. The responses of V4 neurons are not easily predicted from those of earlier areas, nor do they directly reflect our subjective perception of a scene. An accurate model of V4 response properties is essential for understanding how the visual system represents shape.

Initial studies of shape representation in area V4 [3] reported neurons that were tuned to the orientation and spatial frequency of sinusoidal grating patterns, the same dimensions represented in area V1. These dimensions are widely viewed as particularly efficient for encoding natural scenes [4]. It is unclear, however, what advantage the visual system would gain by representing them again at a later stage of processing. An intermediate visual area should represent more elaborate aspects of shape. In fact, many V4 neurons give stronger, more selective responses to sinusoidal gratings modulated in polar and hyperbolic coordinates [5,6]. Theoretical studies have argued that these non-Cartesian coordinate transforms might be useful for visual processing [7,8]. But it is not yet clear whether these dimensions can serve as an efficient substrate for shape processing in visual areas beyond V4.

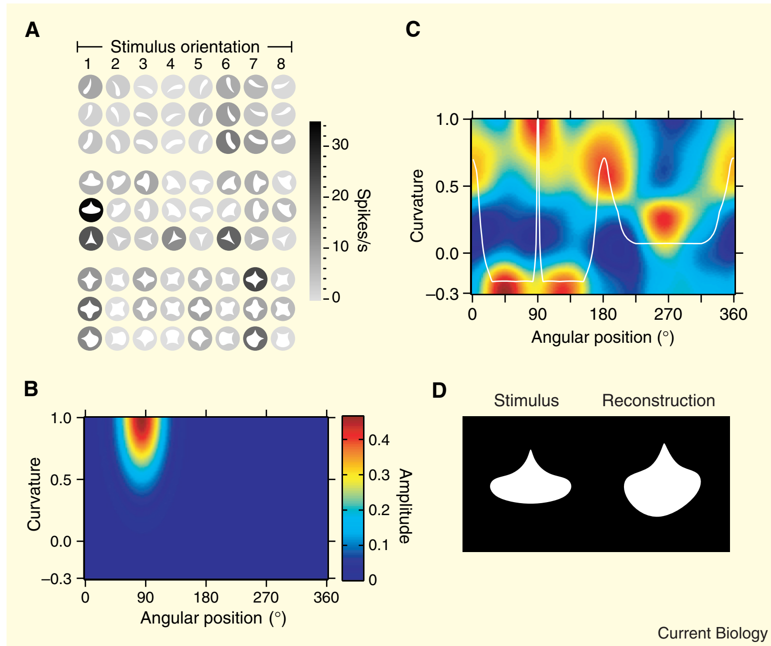
Earlier studies assumed that area V4 uses a rather abstract mathematical representation. In that scheme, later areas transform the intermediate representation into one that is more closely related to our subjective experience. An alternative hypothesis is that area V4 directly represents local features of objects. In this scenario, later areas need only combine these local features to represent entire objects. Pasupathy and Connor [9] investigated this possibility using simple stimuli that contained four to eight boundary segments, the curvature of which could be varied parametrically (Figure 1A). They found that many V4 neurons were tuned to the curvature of specific object parts. One neuron might respond, for example, to a sharp convexity at the lower right of the stimulus irrespective of the configuration of the remaining border segments. These results led Pasupathy and Connor [9] to hypothesize that area V4 might use a parts-based representation to encode object shape.

No neurophysiological study of V4 has systematically investigated tuning for more than a few shape attributes at a time. Still, all of the above studies found that most V4 neurons were selective along one or more distinct stimulus dimensions. It is likely that single V4 neurons are in fact tuned in many dimensions, from Cartesian grating orientation to local boundary curvature. What are we to make of this multimodal tuning? Can neurons respond to several aspects of shape simultaneously? Or is there a single ‘correct’ stimulus representation for area V4?

This is a fundamental problem, but an extremely difficult one to address experimentally. The demonstration of tuning is necessary but not sufficient evidence that a specific dimension can serve as a good foundation for shape representation. Pasupathy and Connor [1] reasoned that a modeling study might show how the the local contour responses of individual V4 neurons could be combined to represent a complex shape. They incorporated single neuron data into a population model of boundary curvature coding, then assessed the model’s ability to represent complete shape boundaries and predict responses to novel stimuli.

[1]Department of Molecular and Cell Biology, [2]Department of Psychology, University of California at Berkeley, Berkeley, California 94720-1650, USA.
E-mail: gallant@socrates.berkeley.edu

Figure 1.
(A) A subset of the stimuli used by Pasupathy and Connor [1] to assess shape tuning in V4. The rows represent various combinations of two (top group), three (middle group), or four (bottom group) convex segments; the columns are the eight orientations at which each shape was presented. The background gray level of each stimulus indicates the response elicited from a single, representative V4 neuron. (B) Gaussian tuning function in a coordinate system defined by angular position (horizontal axis) and curvature (vertical axis), for the neuron whose responses are illustrated in A. The color scale indicates response strength. The peak at a single curvature and angular position indicates that this neuron was tightly tuned for sharp convexities at 90 degrees. This is consistent with the responses shown in panel A for the same neuron. (C) Population response to the 'squashed raindrop' shape, the fifth row in panel A. The white line maps the boundary of this shape into a space whose axes are curvature and angular position. Color peaks indicate the curvature extrema of the stimulus as represented by the simulated V4 population. (D) The stimulus used to create the population response in C (left), and the reconstruction of this stimulus based on the peaks in that response (right).



To construct their V4 population model, Pasupathy and Connor [1] used data from 109 neurons that showed sensitivity to complex curvature. For each one they constructed a two-dimensional tuning function summarizing response strength as a function of boundary curvature and angular position. These tuning functions were fit with two-dimensional Gaussian surfaces to provide interpolation between the measured stimulus values (Figure 1B). Each tuning function was included in the model population eight times, with the angular position rotated in 45 degree steps. (This is equivalent to assuming that tuning for any particular curvature value occurs at all angular positions with equal likelihood.) The response of the entire model population could then be estimated by summing the tuning functions, weighting each by the strength of its predicted response. The resulting population response surface indicated the most likely boundary contour for the stimulus (Figure 1C). The fidelity of the population responses to the original shapes was demonstrated by reconstructing stimuli from their corresponding responses (Figure 1D).

The population model captured the overall configuration of the shape boundaries, though it tended to underestimate sharp convexities and overestimate broad convexities. This reflects limitations of the original neuronal sample, which contained few cells tuned to straight edges. It remains to be seen whether this was an artifact of sampling or an intrinsic bias within V4 [6]. The current model does not incorporate some additional shape attributes that may be represented in V4 [10,11], but the generality of the modeling framework may allow these attributes to be incorporated in future work.

The modeling study of Pasupathy and Connor [1] significantly advances our ability to assess how objects are represented by the visual system. It provides the first good evidence that a specific aspect of shape could in principle be extracted directly from the population activity of area V4. Most prior neurophysiological studies that investigated tuning did not address whether neural responses could support a robust stimulus representation. This modeling strategy is likely to be replicated and extended in future studies to test the feasibility of stimulus representations at other levels of visual processing.

References

1. Pasupathy, A. and Connor, C.E. (2002). Population coding of shape in area V4. Nat. Neurosci. *12*, 1332-1338.
3. Desimone, R., Schein, S.J., Moran, J. and Ungerleider, L.G. (1985). Contour, color and shape analysis beyond the striate cortex. Vis. Res. *25*, 441-452.
2. Van Essen, D.C., and Gallant, J.L. (1994). Neural mechanisms of form and motion processing in the primate visual system. Neuron *13*, 1-10.
4. Simoncelli, E.P. and Olshausen, B.A. (2001). Statistical properties of natural images. Annu. Rev. Neurosci. *25*, 1193–1216.
5. Gallant, J.L., Braun, J. and Van Essen, D.C. (1993). Selectivity for polar, hyperbolic, and Cartesian gratings in macaque visual cortex. Science, 259, 100-103.
6. Gallant, J.L., Connor, C.E., Rakshit, S., Lewis, J. W. and Van Essen, D.C. (1996). Neural responses to polar, hyperbolic, and Cartesian gratings in area V4 of the macaque monkey. J. Neurophysiol. *76*, 2718-2739.
7. Caelli, T.M. (1976). The prediction of interactions between visual forms by products of Lie operators. Math. Biosci. *30*, 191-204.
8. Dodwell, P.C. (1983). The Lie transformation group model of visual perception. Percept. Psychophys. *34*, 1-16.
9. Pasupathy, A. and Connor, C.E. (1999). Responses to contour features in macaque area V4. J. Neurophysiol. *82*, 2490-2502.
10. Hinkle, D. A. and Connor, C.E. (2002). Three-dimensional orientation tuning in macaque area V4. Nat. Neurosci. *5*, 665-670.
11. Kobatake, E. and Tanaka, K. (1994). Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. J. Neurophysiol. *71*, 856-867.