Martinez-Conde, Macknik, Martinez, Alonso & Tse (Eds.)
Progress in Brain Research, Vol. 154
ISSN 0079-6123

CHAPTER 16

Neural basis of shape representation in the primate brain

Anitha Pasupathy*

The Picower Institute for Learning and Memory, RIKEN-MIT Neuroscience Research Center and Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Abstract: Visual shape recognition — the ability to recognize a wide variety of shapes regardless of their size, position, view, clutter and ambient lighting — is a remarkable ability essential for complex behavior. In the primate brain, this depends on information processing in a multistage pathway running from primary visual cortex (V1), where cells encode local orientation and spatial frequency information, to the inferotemporal cortex (IT), where cells respond selectively to complex shapes. A fundamental question yet to be answered is how the local orientation signals (in V1) are transformed into selectivity for complex shapes (in IT). To gain insights into the underlying mechanisms we investigated the neural basis of shape representation in area V4, an intermediate stage in this processing hierarchy.

Theoretical considerations and psychophysical evidence suggest that contour features, i.e. angles and curves along an object contour, may serve as the basis of representation at intermediate stages of shape processing. To test this hypothesis we studied the response properties of single units in area V4 of primates. We first demonstrated that V4 neurons show strong systematic tuning for the orientation and acuteness of angles and curves when presented in isolation within the cells' receptive field. Next, we found that responses to complex shapes were dictated by the curvature at a specific boundary location within the shape. Finally, using basis function decoding, we demonstrated that an ensemble of V4 neurons could successfully encode complete shapes as aggregates of boundary fragments. These findings identify curvature as a basis of shape representation in area V4 and provide insights into the neurophysiological basis for the salience of convex curves in shape perception.

Keywords: V4; Monkey; Electrophysiology; Curvature; Object-centered position; Form perception; Object recognition; Vision

Introduction

One of the primary attributes of the human visual system is its ability to perceive and recognize objects. Humans can recognize an infinite number of complex shapes rapidly. Such recognition proceeds seemingly effortlessly under a variety of viewing conditions — different viewing angles and distances, illumination levels, partial occlusion, etc. Such invariant recognition is a computationally challenging problem because, depending upon the viewing conditions and the presence of occluding objects, the image cast by the object on the retina can be dramatically different. Even so, the primate visual system segments the relevant parts of the shape from the scene and then perceives and recognizes the appropriate object with a speed and precision that is unmatched by even the most cutting-edge machine vision systems. Very little,

*Corresponding author. Tel.: +1–617-324-0132; Fax: +1–617-258-7978; E-mail: anitha@mit.edu

DOI: 10.1016/S0079-6123(06)54016-6

however, is known about the neural basis of such visual shape recognition. Discovering the mechanisms that underlie object recognition will further our understanding about the workings of the primate brain. In addition, knowledge of the underlying brain mechanisms will help construct better automated recognition systems that can perform visual tasks as well as humans.

Ventral visual pathway

The ventral stream or temporal processing pathway in the primate brain is implicated in the processing of object shape and color information (Ungerleider and Mishkin, 1982; Felleman and Van Essen, 1991). The temporal pathway runs from primary visual cortex (V1) to V2 to V4 and into various regions of temporal cortex. This pathway has been worked out in detail in the macaque monkey (Felleman and Van Essen, 1991), and physiological studies in the macaque provide most of our knowledge about shape processing in the primate brain.

Neurophysiological studies addressing the question of shape representation in the ventral visual pathway suggest a hierarchical model for shape processing: successive stages are characterized by larger receptive fields (RFs) and more nonlinear response properties. In primary visual area V1 (the first stage of cortical processing), neurons have small RFs and the responses encode visual stimuli in terms of local orientation and spatial frequency information (Hubel and Weisel, 1959, 1965, 1968; Baizer et al., 1977; Burkhalter and Van Essen, 1986; Hubel and Livingstone, 1987). V1 simple cell responses are usually modeled as a linear weighted sum of the input over space and time (with output nonlinearities) and complex cell responses as a sum of the outputs of a pool of simple cells with similar tuning properties but different positions or phases. Such models fit observed data quite well (for a review see Lennie, 2003).

In the next processing stage, area V2, cells have larger receptive fields — on average the area of a V2 RF is approximately six times that of a V1 RF (based on data from Gattass et al., 1981). V2 neurons encode information about complex stimulus characteristics, in addition to local orientation and frequency information. Many V2 neurons are sensitive to illusory or subjective contours (while cells in V1 are not), and this selectivity is thought to be achieved by pooling from several end-stopped V1 cells (von der Heydt and Peterhans, 1989). There is also some evidence that V2 responses may encode stimulus characteristics such as polarity of angles and curves (Hegde and Van Essen, 2000), texture borders and stereoscopic depth cues (von der Heydt et al., 2000). V2 lesions impair discrimination of shapes defined by higher-order cues but not those defined by luminance cues (Merigan et al., 1993). Thus, while V1 responses primarily encode contours defined by luminance, V2 responses encode contours defined by second-order cues as well (Gallant, 2000).

Several lesion and neurophysiological studies have demonstrated that area V4, the next stage along the temporal processing pathway, plays a crucial role in form perception and recognition. Bilateral V4 lesions result in severe impairment in form discrimination (Heywood and Cowey, 1987). V4 lesions in macaques affect perception of intermediate aspects of stimulus form while sparing stimulus properties explicitly represented in V1 (Schiller and Lee, 1991; Schiller, 1995; Merigan and Phan, 1998). For example, V4 lesions reduce the ability to discriminate the orientation of illusory contours (De Weerd et al., 1996) and to identify borders defined by texture discontinuities (Merigan, 1996). In a human patient with a putative ventral V4 (V4v) lesion, Gallant et al. (2000) reported impairments in the discrimination of illusory contours, glass patterns, curvatures and non-Cartesian gratings.

RFs of V4 neurons are much larger than those in V1 and V2 — on average, at a given eccentricity, a V4 RF is 16 times the area of a V1 RF. In addition, V4 neurons have been reported to have large suppressive surrounds (Desimone et al., 1985). Thus, V4 neurons have access to stimulus information over a large region of the stimulus space and this is thought to serve global perceptual mechanisms, and contribute to the selectivity of V4 neurons to stimulus shape. Some V4 neurons, like complex end-stopped cells in V1 and V2, are selective for stimulus length, width, orientation

and spatial frequency (Desimone and Schein, 1987) while others show greater nonlinearity in their responses than V2 neurons. Some cells encode edges explicitly (unlike V1 cells) as suggested by their strong responses to bars and square-wave gratings but not to sinusoidal gratings. Gallant et al. (1993, 1996) demonstrated V4 selectivity for non-Cartesian polar and hyperbolic gratings. The authors suggest that selectivity for polar gratings may mediate perception of curvature. Selectivity for hyperbolic gratings may play an important role in image segmentation, since these cells would respond well to line-crossings, a good source of information about image structure. Kobatake and Tanaka (1994) have demonstrated that many V4 neurons, unlike V2 neurons, respond stronger to complex than simpler stimulus features. Thus, while a lot of studies have demonstrated complex response selectivities in V4 neurons, no coherent, principled representational bases have been identified.

Unlike the areas that precede it, inferotemporal cortex (IT) has no discernible visuotopic organization (Desimone and Gross, 1979). IT neurons have very large RFs that almost always include the center of gaze and frequently cross the vertical meridian into the ipsilateral visual field (Gross et al., 1972). Anterior IT appears to be the last stage in the shape processing pathway, since its efferents project to areas in the temporal and frontal lobe that are not exclusively concerned with vision (Desimone et al., 1985). IT neurons are often selective for complex shapes like faces and hands (Gross et al., 1972; Perrett et al., 1982; Desimone et al., 1984; Tanaka et al., 1991). Several studies have demonstrated that IT neurons, at the single cell, columnar and population levels, encode information about features or parts of complex objects (Schwartz et al., 1983; Fujita et al., 1992; Kobatake and Tanaka, 1994; Wang et al., 1996; Booth and Rolls, 1998). Kobatake and Tanaka (1994) measured the "complexity" of the encoded features by measuring the ratio between best responses to complex and simple stimuli and found that the complexity of encoded shapes increased from V4 to posterior IT to anterior IT. These complex shapes often comprised simpler shapes with a specific spatial relationship between the component parts (Tanaka et al., 1991; Kobatake and Tanaka, 1994). Responses were sensitive to small changes or deletions of the critical attributes of the multipart objects (Desimone et al., 1984; Tanaka, 1993; Kobatake and Tanaka, 1994), but relative responses were insensitive to changes in absolute position (Sato et al., 1980; Schwartz et al., 1983; Ito et al., 1995). Studies about the dependence of IT responses on size, position and cue attributes of an object have demonstrated that single IT neurons can abstract shape properties from widely varying stimulus conditions (Sary et al., 1993).

Investigating shape representation in area V4

The results discussed above suggest that, in the earliest stage of shape processing, responses faithfully encode local stimulus characteristics. At successive levels, neural responses are dictated by larger regions of the visual field and become increasingly nonlinear functions of the stimuli. As a result, neural responses reflect the abstraction of stimulus form, with local characteristics becoming increasing irrelevant in determining the neural response. At the highest stages, responses encode the global shape and are largely invariant to local cue information, position, size, etc. How is this achieved, i.e. how are local orientation and spatial frequency signals in V1 transformed into selectivity for complex shapes in IT? To gain insights into the mechanisms that underlie this transformation of stimulus representation, we need to decipher the neural basis of representation at intermediate stages between V1 and IT. At successive stages of processing we need to discover the relevant stimulus dimensions and the quantitative relationship between those dimensions and the neural response. The relationship between the bases of representation at successive stages will reveal the computations required and will thereby provide insights into the underlying mechanisms. This will also provide clues about how these representational schemes might underlie perception and recognition.

To this end, we sought to discover the bases of shape representation in area V4, an intermediate stage in the ventral visual pathway. As described

above, V4 neurons show selectivity for a wide variety of complex stimuli but the specific shape dimensions represented are as yet unknown. We therefore studied V4 neurons with the specific aim of discovering the underlying tuning dimensions. The ideal, most desirable, approach to investigating the basis of shape representation in any neuron would be to study the responses to a large set of complex natural stimuli created by uniformly sampling all of "shape space". Then, analytical approaches with minimal assumptions, such as spike-triggered covariance, could be used to extract the shape dimensions along which the responses vary maximally. However, even a coarse, uniform sampling of shape space would require thousands of stimuli, since object shape varies along a very large number of dimensions. For example, all possible combinations from a 10×10 pixel patch, whose pixels take one of two values (background or preferred color of cell), will result in 2^{100} stimuli. Such an exhaustive search approach would be impractical, since a primary experimental consideration is the length of time, on average 1–2 hours, a single cell can be studied in an awake behaving animal. A pragmatic alternative would be to explore shape space in a directed fashion for principled investigations of a specific hypothesis about shape representation. Depending on the hypothesis, an answer could be arrived at by targeted sampling of the relevant subregion of the extremely large shape space consisting of all possible two-dimensional (2D) shapes. We followed this approach to decipher the bases of representation in area V4. On the basis of theoretical, psychological and physiological studies we identified candidate dimensions for representation in area V4 (see below) and then studied the response characteristics of neurons as a function of the shape dimensions in question.

Contour characteristics as basis of shape representation

What might be a good candidate dimension of shape representation in area V4? To answer this question we turned to theoretical and psychophysical studies. Many modern shape theories and computational models propose a scheme of recognition by hierarchical feature extraction, i.e. objects are first decomposed into simple parts that are pooled at subsequent stages to form higher-order parts. This approach has been motivated in part by physiological findings of selectivity for simpler features in earlier stages (see above). It has also been motivated by the human ability to discern the various parts of an object and their relative positions in addition to recognizing the object as a whole. There is also some psychological evidence for the hypothesis that recognition proceeds by parsing shapes into component parts and then identifying the parts and the spatial relationships between them (Biederman, 1987; Biederman and Cooper, 1991). Finally, from a practical and informational standpoint, such basis representations are advantageous — they provide data compression and are therefore compact and good for further computations.

Feature-based models for visual shape representation vary widely in the number of hierarchical levels and the nature of features or shape primitives extracted at each stage. Beyond local edge orientation, which most models invoke as the first-level feature, shape primitives extracted at intermediate levels are related to the object boundaries (Attneave, 1954; Milner, 1974; Ullman, 1989; Poggio and Edelman, 1990; Dickenson et al., 1992) or to its volume (Biederman, 1987; Pentland, 1989). While boundary-related features (i.e. contour features) include 2D angles and curves and three-dimensional (3D) corners, curved surface patches and indentations, solid or volumetric primitives (generalized cones or geons) are simple 3D shapes such as cylinders, spheres, etc. defined by the orientation of their medial axes and several cross-sectional attributes.

Several theoretical studies have asserted the importance of contour features as a basis of representation: they are high in information content and lead to an economical representation (Attneave, 1954), they are largely invariant to deformation (Besl and Jain, 1985) and can be constructed from local edge orientation and curvature signals (Milner, 1974; Zucker et al., 1989). Further, they form natural parts for constructing more complex representations. Psychological findings imply specialized

mechanisms for the perception of contour features: angle perception acuity is higher than that predicted by component line orientation acuity (Chen and Levi, 1996; Heeley and Buchanan-Smith, 1996; Regan et al., 1996), the detection thresholds for curvilinear glass patterns is much lower than that for radial glass patterns (Andrews et al., 1973; Wilson et al., 1997) and detection of curved targets among straight distractors is faster than that for straight targets among curved distractors (Triesman and Gormican, 1988; Wolfe et al., 1992).

Thus, based on theory and psychophysics, contour features seem to be a good choice for the basis of shape representation beyond local orientation. But does the primate visual system actually extract contour features as intermediate-level shape primitives beyond oriented edges, in the process of recognizing visual shape? In this paper we describe experiments that explicitly test this hypothesis. If contour features, i.e. convex projections and concave depressions along the object contour, are indeed extracted in area V4 then one would expect:

a) At least some cells to be strongly and systematically tuned to contour features.
b) Responses to complex shapes to be dictated primarily by specific contour features along the object contour.
c) Population representation of complex shapes, in terms of their component contour features, to be complete and accurate.

Below, we demonstrate that each of these statements is true for area V4. In our experiments, roughly 33% of V4 neurons showed strong systematic tuning for contour features, i.e. angles and curves, presented in isolation within their RFs. (Pasupathy and Connor, 1999). Responses of many V4 neurons to moderately complex shapes are dictated by the curvature of the contour at a specific location relative to object center (Pasupathy and Connor, 2001). Finally, the estimate of V4 population representation, derived from single cell responses and curvature tuning properties, provides a complete and accurate representation of complex 2D contours (Pasupathy and Connor, 2002). These results suggest that contour characteristics (parameterized in terms of contour curvature) at a specific spatial location relative to object center are a basis for shape representation in area V4. Certainly there are other important shape dimensions encoded in area V4 (see the section ''Discussion''). Our results described below demonstrate that contour features are an important dimension.

Tuning for isolated contour features in area V4

If contour features are a basis of representation then we would expect single cells to respond selectively to angles and curves presented in isolation within the cell's RF. This is the first question we addressed with our experiments.

We studied the responses of 152 V4 neurons to a large parametric set of single contour features. Many V4 neurons showed strong systematic tuning to these contour feature stimuli responding preferentially to angles and/or curves oriented in a specific direction (Pasupathy and Connor, 1999). Fig. 1A shows the responses of a V4 neuron to the angles and curves that we used in this study. Each icon shows a stimulus with the background gray level representing the average response rate based on five stimulus repetitions. The stimuli were angles, curves or straight edges, presented in isolation within the center of the RF, in the optimum color for the cell under study (shown here in white) as determined by preliminary tests. Stimulus luminance was constant within the estimated RF boundary and gradually faded into the background gray over a distance equal to estimated RF radius.

The stimulus set included convex, concave and outline angles and curves that pointed in one of eight directions at 45° intervals (orientation 45°, 90°, etc.) and subtended one of three angular extents (acuteness: 45°, 90°, 135°). Four edges and lines were also included for a total of 156 stimuli. Stimuli were flashed for 500 ms each and were separated by an interstimulus interval of 250 ms. A sequence of five stimuli comprised a trial. The entire stimulus set was presented in random order without replacement five times.

For the cell in Fig. 1A, response rates range from 0 (light gray) to 42 ± 2.3 spikes/s (black) as indicated by the scale bar. Responses to acute and right angle convex stimuli pointing in the 135–180° range were strong while the same stimuli drawn as

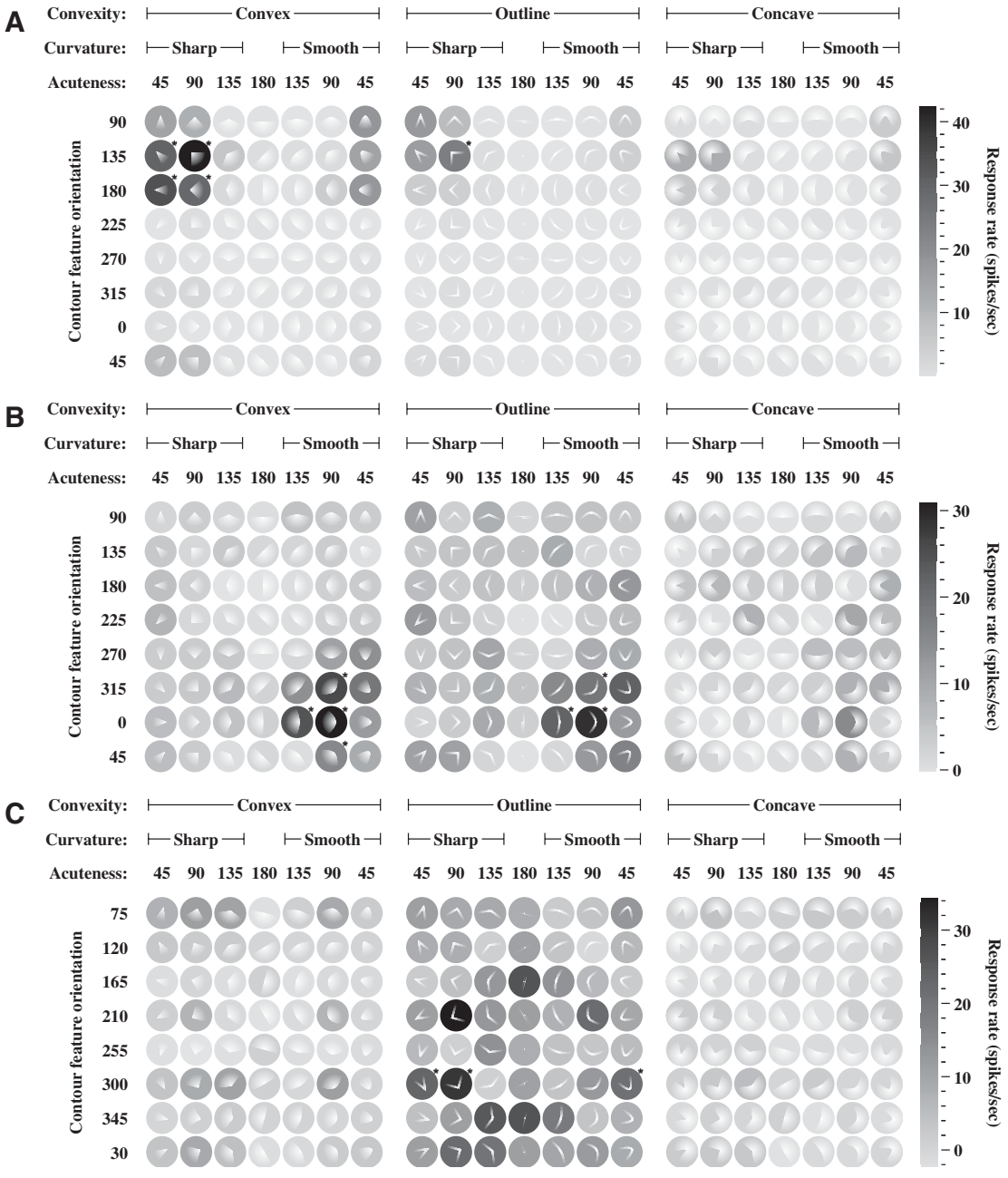

Fig. 1. Examples of V4 responses to contour feature stimuli. Each icon represents a stimulus drawn against a dark background. Circular boundaries were not a part of the actual stimulus display. Each stimulus consists of a straight edge, angle, or curve drawn as a convex projection, outline or concave indentation, centered within the RF. Boundaries are sharp within the RF and gradually fade into the background. Entire extent of fading is not shown here. (A) Example response pattern showing contour feature tuning. Responses of a cell showing strong tuning for sharp angles pointed in the 135–180° range. Background gray levels indicate average responses over five repetitions as per scale bar on the right. (B) Another example of contour feature tuning. Responses show tuning for curves pointed in the 315–360° range. (C) Example response pattern showing orientation tuning. Cell responded strongly to many stimuli that contain a 75° oriented edge or line. Figure originally published in Pasupathy and Connor (1999).

outlines elicited moderate responses. Responses were weaker to smooth (vs. sharp), obtuse (vs. acute) and concave (vs. convex) features. The responses of this cell contribute to a clear, strong, unimodal peak in the orientation × acuteness contour feature space (for detailed analysis see Pasupathy and Connor, 1999). These results cannot be explained in terms of orientation tuning for individual edges, since many of the least effective stimuli (including the straight edges) contain the same edge orientations as the most effective stimuli. This result is strikingly different from the recently reported angle selectivity in area V2 (Ito and Komatsu, 2004). Unlike in V4, angle-selective V2 units responded in comparable amounts to the preferred angle and to one or both of its component end-stopped lines.

A second example of contour feature tuning is shown in Fig. 1B. Convex and outline curves oriented at 315°–360° elicited strong responses from this cell. Responses were strongest to 90° curves but the cell responded well to some obtuse curves as well. Unlike the previous example, this cell did not respond to sharp angle stimuli. Here again, stimuli that drive the cell are clustered together contributing to a single strong peak of high responses in the stimulus space.

Unlike the previous examples, the cell in Fig. 1C did not show clustering of strong responses in the stimulus space. This cell responded well to a variety of angle and curve outline stimuli containing edges oriented near 75° as part of the stimulus. In the preliminary bar orientation test, this cell demonstrated strong orientation tuning with a peak at 75°. The response pattern exhibited by this cell reflects this orientation tuning, since the responses to contour features are primarily dictated by the component orientations of the contour feature stimuli.

Contour feature tuning described above (Figs. 1A, B) cannot be explained in terms of tuning along simpler stimulus dimensions such as edge orientation, contrast polarity or spatial frequency. Component edges of the optimal stimuli also appeared in other stimuli that failed to evoke strong responses from the cell. Secondly, tuning for the orientation of a contour feature was consistent across the three (45°, 90°, 135°) acuteness values despite differences in component edge orientation and, finally, responses to contour feature stimuli were almost always greater than to single lines or edges. Tuning for spatial frequency or contrast polarity fails to explain contour feature tuning because both optimal and nonoptimal stimuli were composed of similar spatial frequencies and contrast polarities. Further, contour feature tuning was consistent across convex, concave and outline stimuli despite substantial differences in contrast polarity and spatial frequency content. Consistency of tuning across convex, concave and outline stimuli also argue against explanations in terms of area of stimulation or differential surround stimulation hypotheses. Finally, contour feature tuning cannot be explained in terms of hotspots in the RF because tuning profile (rank-order of responses) was similar at different positions in the RF (Pasupathy and Connor, 1999).

Roughly one-third of the entire sample of 152 V4 cells showed strong unimodal peaks (as in Figs. 1A, B) in this orientation × acuteness contour feature space. Most cells responded best to either convex or outline stimuli. This bias towards convex features parallels findings of perceptual significance of convex features reported by several psychophysical studies (see the section "Discussion"). A majority of cells also responded preferentially to 45° (more acute) stimuli but this acuteness bias was weaker than that for convexity. Many cells responded equally well to both angle and curve counterparts while other cells were clearly biased toward either sharp or smooth stimuli. In the case of curves, the angle subtended by the contour feature is analogous to contour curvature — more acute angles correspond to higher curvature. Thus, tuning along the acuteness dimension suggests that curvature may be an important shape dimension encoded by V4 neurons. We explored this further in the next set of experiments.

The results above demonstrate that V4 neurons carry explicit information about the orientation and acuteness of contour features when presented in isolation. This suggests that angles and curves along the object contour could serve as intermediate-level primitives, and that V4 neurons could encode 2D contours in terms of their component angles and curves. To ascertain if they do, in our

second set of experiments, we tested how V4 neurons respond when contour features are presented in the context of a closed 2D contour. Do V4 neurons tuned for contour features represent complex shapes in terms of their component contour features, i.e. are the responses to complex shapes dictated by convex and concave features along the shape contour? To answer this question we studied the responses of single V4 neurons to a set of complex shapes created by a systematic combination of convex and concave contour features and the results are discussed below.

Conformation of 2D contours dictate responses of V4 neurons

If contour features are a basis for shape representation in area V4, then responses of V4 neurons to complex-shape stimuli will carry reliable information about the component contour features of the shape in question. To test this question, in the second set of experiments, we studied the responses to complex-shape stimuli of 109 V4 neurons that showed preferential responses to contour features in preliminary tests (Pasupathy and Connor, 2001). A large parametric set of complex shapes, created by systematic variation of curvature along the object contour, was used to study the responses of V4 neurons. An example is shown in Fig. 2.

Each stimulus (shown in white) consisted of a closed shape, presented in the optimal color for the cell, in the center of the RF (represented by the surrounding circle). All parts of all stimuli lay within the estimated RF for the cell under study. The stimulus set consisted of 366 complex shapes with varying combinations of convexities and concavities along the contour. Stimuli had two, three or four convex projections that were separated by 90°, 135° or 180°. Convex projections were sharp points (0° location stimulus *2*) or medium convexities (135° location stimulus *2*). The intervening contours between the convex projections were rendered as arcs of circles whose curvature and convexity were defined by the angular separation and curvature of the adjoining convex projections. Smooth transitions between contour segments were achieved by rendering the stimuli as piecewise cubic b-splines through the control points associated with the contour segments. Stimuli were presented in eight orientations (*horizontal* within each block) at 45° intervals except for stimuli that were rotationally symmetric. The stimulus set also included two circles with diameters 3/16 and 3/4 of estimated RF diameter. Stimulus presentation protocol was as in experiment above.

For the example in Fig. 2, response rates ranged from -6.3 ± 0.0 (light gray) to 38.1 ± 7.0 spikes/s (black). This cell responded strongly to a wide variety of shapes that had a sharp convexity in the lower left corner of the object (angular position 225°) with an adjoining concavity in the counterclockwise direction. Stimuli with a medium convexity at 225° elicited a moderate response from the cell. For example, compare the responses to stimuli *1* and *3* or stimuli *2* and *4*, which differ in their contour only at angular position 225°. In both cases the stimulus with the sharp convexity (stimuli *1* and *2*) elicited the stronger response. In contrast, stimuli with a broad convexity or a concavity at 225° did not drive the cell. Curvature of the contour from 0° to 90° varied widely across stimuli that evoked strong responses from the cell. The results described so far suggest that the contour curvature at a specific angular position strongly dictates the responses to complex shapes.

To quantify the functional relationship between the contour characteristics and the neural response, we represented each stimulus in a multidimensional shape space in terms of its contour characteristics and then modeled the relationship between the stimulus dimensions and the neuronal responses. Each stimulus in our set can be approximated by 4–8 constant curvature segments that are connected by short segments with variable curvature entirely determined by the adjoining segments. Therefore, each shape can be uniquely represented using a discretized and simplified representation of just the constant curvature segments. These constant curvature segments that comprise the stimuli could be uniquely described just in terms of the angular position relative to center (or tangential orientation) and contour curvature. For a general unconstrained 2D closed contour unique representation would require additional dimensions such as position along the

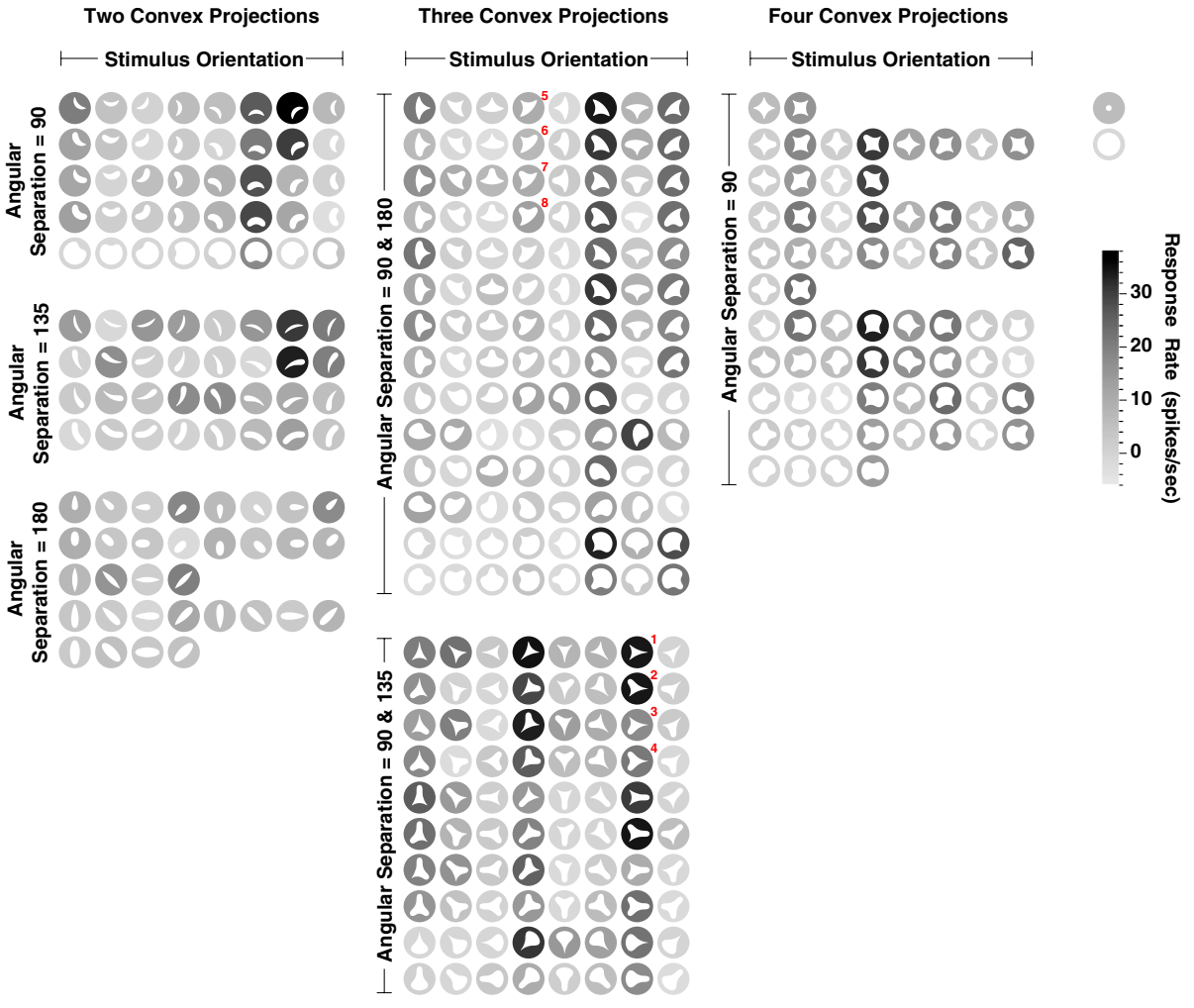

Fig. 2. Example cell from the complex shape test. Within each dark circle, which represents the RF, a complex shape stimulus is drawn in white. The circular boundaries were not an actual part of the visual display. Stimuli were constructed by systematic variation of the convexities and concavities along the contour. Stimuli are divided into groups on the basis of the number of convex projections and the angular separation between convex projections. Background gray level represents average response, to the corresponding stimulus, as per scale bar on the right. Shapes with a sharp convexity pointing to the lower left and an adjoining concavity in the counterclockwise direction elicit strong responses from this cell. Numbers to the upper right of stimuli correspond to references in text. Original figure published in Pasupathy and Connor (2001).

contour, radial position relative to the center of the stimulus and local tangential orientation. Here, since several of these dimensions co-vary, representation in terms of just curvature and angular position is unique.

In this scheme, each stimulus is represented by four to eight points in the curvature by angular position space. For example, stimulus *1* is represented by six points: one for each of the convex points and the intervening concavities. To quantify the dependence of the neural response on angular position and contour curvature, we modeled the neural response as a 2D Gaussian function (product of two 1D Gaussians with no correlation terms) of contour curvature and angular position. Here, response of a neuron to a stimulus was the maximum of the responses predicted by its component contour segments. The parameters for these

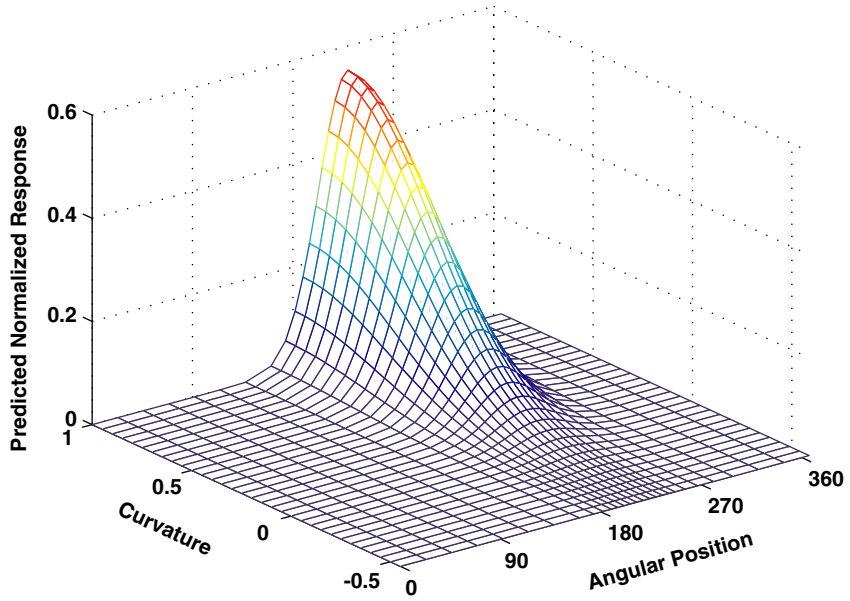

Fig. 3. Best-fitting 2D model for the responses illustrated in Fig. 2. Horizontal axes represent angular position and contour curvature. Vertical axis and surface color represents normalized response predicted by the 2D angular position-curvature model. The Gaussian model had a peak at 229.6° (angular position) and 1.0 (curvature) predicting strong responses for stimuli with sharp convexities pointing to the lower left. Original figure published in Pasupathy and Connor (2001).

curvature × position tuning functions (peak position and standard deviation (SD) for each of the two dimensions and the amplitude of the Gaussian) were estimated by minimizing the sum of squared errors between predicted and observed responses (for further details, see Pasupathy and Connor, 2001).

The tuning surface predicted by the best fitting curvature × position model, for the example in Fig. 2, is shown in Fig. 3. The horizontal axes represent angular position (0°–360°) and contour curvature (−0.31–1.0; negative values represent concavities and positive represent convexities). The *z*-axis and surface color represent the predicted normalized response. The peak of the Gaussian surface is at contour curvature equal to 1.0 and angular position equal to 229.6°, predicting strongest responses to stimuli with a sharp convex projection at or near the lower left corner of the object. Tuning along the angular position dimension was narrow (SD = 26.7°), implying that a small range of positions (centered at 229.6°) of the sharp convexity evoked strong responses. Predicted responses were close to zero for stimuli with broad convexities and concavities at the lower left. The goodness of fit of the model, in terms of correlation between observed and predicted values, was 0.70.

The predictions of the 2D model do not accurately reflect the responses of the cell. The model predicted high responses to all stimuli with a sharp convexity at the lower left of the object. However, several stimuli with sharp convex projections at the lower left elicited weak responses from the cell (see stimuli *5–8*, Fig. 2). Among the stimuli with a sharp convexity at the lower left, those with a concavity in the counterclockwise direction (~270°) evoked stronger responses than those with a broad convexity in that same location (stimuli *5–8*). Thus, an adjoining concavity is essential for a strong response.

To include the influence of the adjoining contour segments, we built a four-dimensional (4D) curvature × position model. In addition to angular position and the corresponding contour curvature, we included curvatures of the adjoining contour segments as independent variables. The 4D model was a product of four 1D Gaussians with no second-order correlation terms. This model had nine parameters: peak position and SD associated with

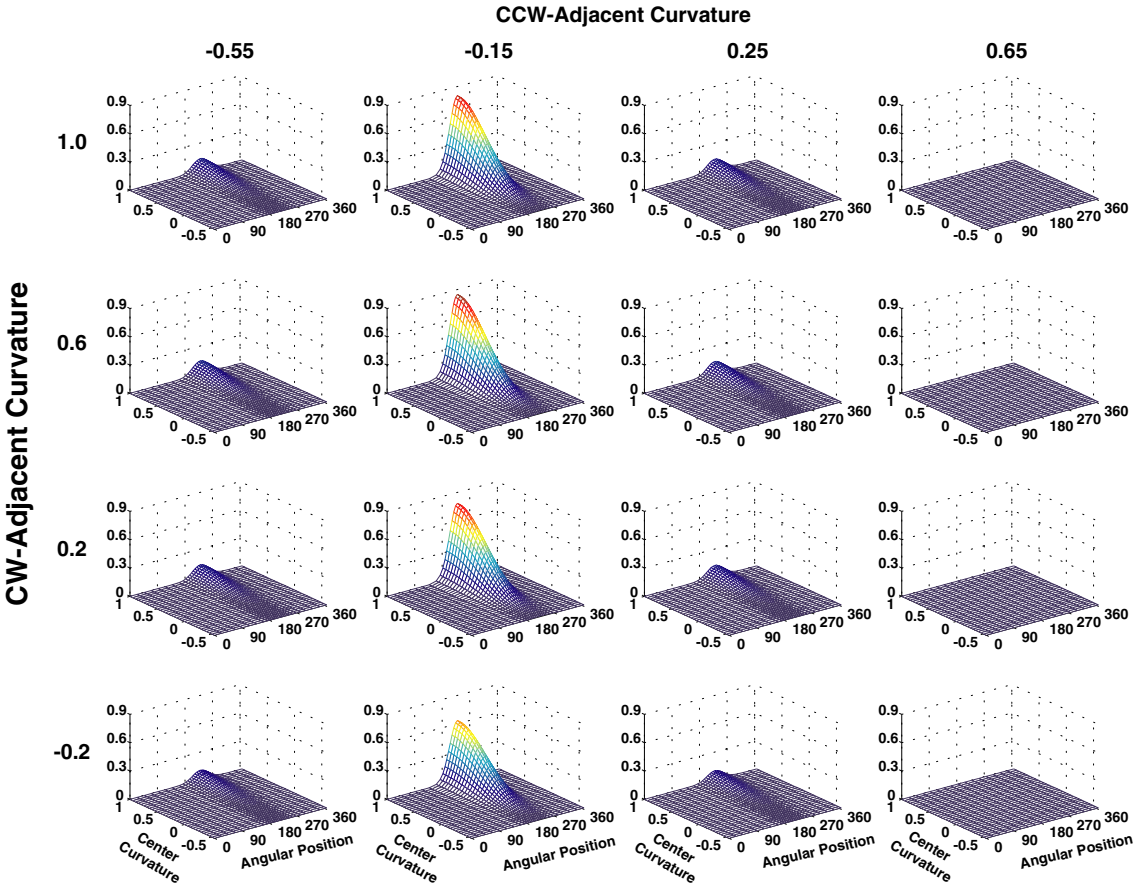

Fig. 4. Results of the 4D model for the responses illustrated in Fig. 2. Each plot represents a slice through the multi-dimensional Gaussian surface. The slices are perpendicular to the CCW and CW curvature dimensions. In the 4×4 grid of plots, horizontal position of the plot represents CCW curvature and vertical position represents CW curvature, as labeled along the top and left, respectively. The horizontal axes of each plot represent angular position (0–360°) and curvature of the central contour segment. The peak of the Gaussian model was located at 230.04° (angular position), 1.0 (curvature of central contour segment), −0.15 (curvature of CCW segment), 0.6 (curvature of CW segment). Model predicts strong responses for shapes with a sharp convexity pointing to the lower left when the adjoining CCW segment is concave. Influence from the CW segment is minimal. CCW: counterclockwise, CW: clockwise. Figure from Pasupathy and Connor (2001).

each of the four dimensions and a parameter for the amplitude of the Gaussian. Fig. 4 illustrates the tuning profile predicted by the 4D model for the example in Fig. 2. This figure shows 16 surface plots, each of which is a slice through the multi-dimensional Gaussian model at different values for the adjoining curvature segments. Horizontal position of each plot represents curvature of the counterclockwise (CCW) contour segment, and vertical position represents curvature of the clockwise (CW) curvature segment. For example, the surface plot in the lower left corner represents the curvature × position tuning profile for the cell when the adjoining contour segments are both concave (curvatures: CCW = –0.55, CW = –0.2). The horizontal axes for each surface plot represent contour curvature and angular position of the central contour segment, and the z-axis plots predicted normalized response. The peak of the 4D model was at 230.04° and 1.0 along the angular position and central curvature dimensions, similar to the results of the 2D model. In addition, the curvature of the CCW segment had an influence on the predicted response as illustrated by the

scaling of the Gaussian peak with CCW curvature (across each row). The tuning profile along the CCW curvature dimension had a peak at broad concavity (curvature = –0.15, second column), and the predicted response dwindled with more convex or concave curvature values. Thus, for stimuli with a sharp convexity at the lower left of the object, the model predicted strong responses when it was flanked by a concavity in the CCW direction and weaker responses when it was flanked by a broad convexity (such as stimuli *5–8*, Fig. 2). In contrast, the curvature of the CW segment exerted a weak influence on the response profile as indicated by the almost identical tuning profiles down each column. Tuning along the angular position dimension was narrow (SD = 25.8°), implying that the preferred pattern of contour curvature evoked a strong response if it occupied a small range of angular positions in the lower left of the object.

The SD parameters for the CCW, CW and central curvature dimensions were 0.21, 1.08 and 0.5, respectively. High SD (compared to 1.31, which is the range of curvatures sampled) implies a flat, broad-tuning profile with a small maximum–minimum response difference over the range of curvatures sampled, and therefore a weak dependence of the response on the corresponding dimension. A low SD implies narrow tuning with a large maximum–minimum response difference, and therefore a strong influence on the response. Just as illustrated by the surface plots in Fig. 4, the SD parameters for the 4D model imply a strong dependence of the predicted response on the central and CCW contour segments. The goodness of fit represented by the correlation coefficient was 0.82 for this 4D model.

The coefficient of correlation between observed and predicted responses (r) was used to assess the goodness of fit of the 2D and 4D curvature × position models. For 101/109 cells, the 2D model produced a significant fit (F test, $p<0.01$) and the median of this distribution was 0.46. Inclusion of the two adjoining curvature segments in the 4D model significantly improved the fit in 93/109 cases (partial F test, $p<0.01$) and the median of the r distribution increased to 0.57. The model's validity was confirmed by predicting responses to stimuli that were not used to derive the tuning functions.

Alternate hypotheses

Can the response patterns described above be explained in terms of previously described tuning properties of V4 neurons such as edge orientation, bar orientation or contrast polarity? We tested these alternate hypotheses explicitly and found that none explained the results presented here. For each cell, we modeled the responses to complex shapes as a function of edge orientation and contrast polarity and assessed the goodness of fit. We also modeled responses as a function of bar orientation, length and width. Third, we tested whether responses to complex shapes were a function of the mass-based shape of the object by representing each shape in terms of its principal axis orientation and the aspect ratio of the approximating ellipse. For most cells, the predictive power of these models was poor (median $r<0.3$) and the 2D and 4D curvature × position models provided superior fits to the observed data. The 2D model had as many or fewer parameters than these alternate models. Therefore, the quality of fits was not simply a function of the number of parameters in the model but a true reflection of the cell's tuning along the corresponding stimulus dimensions (such as curvature, edge orientation, etc.).

Another potential explanation is that the apparent shape tuning is due to selective activation of hotspots in the RF. In other words, rather than the shape of the bounding contour, the strength of the response may simply depend on whether certain subregions of the RF are stimulated by the visual stimulus. If this were the case then shape tuning would be dependent on the position of the stimulus within the RF. We investigated position dependence of shape tuning in roughly one-third of the neurons. For each cell, we presented a subset of stimuli at multiple positions within the RF. The subset included at least one stimulus that had the preferred contour segment (as determined by its responses to complex shapes and the angular position-curvature model) as part of its contour and another stimulus that did not. In all we found that responses to all stimuli were modulated as a function of the position of the stimulus in the RF. However, the responses to the preferred stimulus were greater than those to the nonpreferred for all

stimulus locations, i.e. while responses decreased in magnitude the preferred features were consistent at all locations within the RF (Pasupathy and Connor, 2001).

The curvature × position models discussed so far test whether complex shape responses were dictated by a specific sector along the object contour. Another possibility is that responses were dictated by two (or more) noncontiguous regions of the contour. We investigated this hypothesis by modeling the neuronal responses to complex shapes as a function of two noncontiguous sections of the contour separated by 180°. We restricted the angular separation to 180°, since curvatures of contour segments closer than 180° were not entirely independent of each other. This 3D model, consisting of one angular position and the curvatures of two contour segments as independent variables, had seven parameters. The performance of the model (median $r = 0.47$) was similar to the 2D curvature × position model, i.e. the second (noncontiguous) contour segment made a negligible contribution to the predictive power of the model. This suggests that the curvature × position models with contiguous contour segments provide a better description of the responses of V4 neurons. A sum of Gaussian's model, where each Gaussian was 4D, further confirmed the negligible influence of other regions of the contour.

As alluded to earlier, in this stimulus set object-centered angular position and tangential orientation for the curvature segments comprising each shape were completely dependent on each other, since the curvature segments were pointed directly away from or towards the object center. With a secondary test, we sought to determine whether the true dependence of the neural response was on object-centered angular position or on orientation of the corresponding curvature segment. We tested cells with a stimulus set (Fig. 5), in which these two factors are partially segregated, since the convex extremities are offset (orthogonal to the main axis) in the CW and CCW directions. If curvature orientation (i.e. the pointing direction of the convex extremities) were the only significant dimension, then tuning patterns should be similar (though scaled in amplitude) across the three blocks of rows in Fig. 5. In fact, the tuning patterns are very different. In the top block, responses are strongest for the CCW orthogonal offset; in the middle block responses are strongest for the stimuli with no orthogonal offset; in the bottom block, responses are strongest for the CW orthogonal offset. This reversal pattern reflects tuning for angular position: at each orientation, the most effective stimuli are those with convex extremities near 45° relative to the object center. A similar reversal pattern was observed for 23/29 cells. For each of the 29 cells we determined the orthogonal offset position that excited the cell most for each of the three stimulus orientations tested. A linear regression between the orthogonal offset position and stimulus orientation had a negative slope in 23 cases indicating a reversal pattern in the responses similar to the pattern shown here. Also, interaction between orthogonal position and stimulus orientation was significant in all 23 cases that showed selectivity for polar position of the feature relative to the object center. These findings suggest that the position of boundary pattern relative to object center does dictate the neural response in V4.

The experiments described so far demonstrate that the neurons in area V4 are tuned to contour curvature at an object-centered position. This tuning cannot be explained in terms of other previously described properties of V4 neurons. The responses cannot be described in terms of hotspots in the RF. Thus, these results suggest that boundary conformation, which can be parameterized in terms of contour curvature, at a specific position relative to object center, may be a basis of shape representation in area V4. If this were true then we would expect a complete and accurate representation of 2D shape contours in terms of contour curvature in the V4 population response. We tested this hypothesis explicitly and the results are described below.

Ensemble representation of 2D contours in area V4

How complete and accurate is the representation of 2D contours at the population level in area V4? We tested this question by deriving the population representation for each shape in our stimulus set using basis function decoding. As a first approach to population analysis, we focused on the simpler curvature × position 2D domain. All 109 cells that

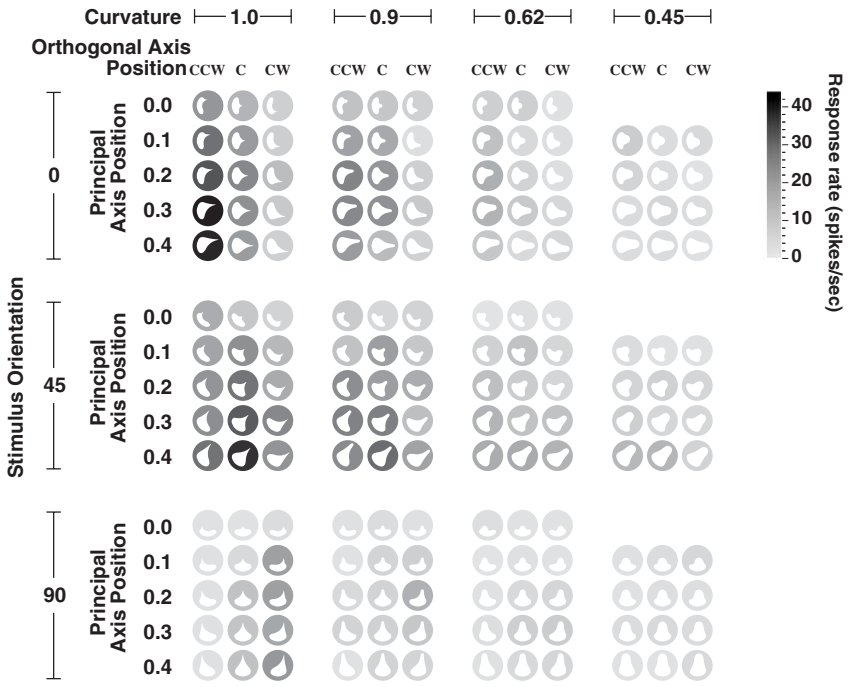

Fig. 5. Example result from the secondary shape test. Stimulus set consists of distorted versions of a teardrop-shaped stimulus. Distortions were created by varying position of the convex point along the principal and orthogonal axes and by varying the curvature of the convex point and the overall stimulus orientation. CCW: counterclockwise; C: center; CW: clockwise. As above, circular disks were not a part of the stimulus display. Background gray levels represent average response rates as per gray scale bar. Responses are strongest to sharp convexities (leftmost vertical block) that are farthest along the principal axis (last row in each horizontal block). Tuning for orthogonal position (horizontal within each block) and stimulus orientation (vertical axis) are interdependent such that responses are strongest for stimuli with the sharp convex point at the upper right corner of the object. For instance, at 0° stimulus orientation, CCW orthogonal position elicits the strongest response, at 45° center elicits the strongest response, and at 90° CW elicits the strongest response. Figure from Pasupathy and Connor (2001).

we recorded from were included in the population analysis.

We used the curvature × position tuning functions (e.g. Fig. 3) to estimate the V4 population response to each shape in our stimulus set. Fig. 6 shows the results of this analysis for an example stimulus, the "squashed raindrop" shape shown at the center. The curvature × position functional representation of this shape is drawn as a white line in the surrounding polar plot. This function has peaks and troughs corresponding to the major features of the shape: a medium convex peak at 0° (right), a concave trough at 45° (upper right), a sharp convex peak at 90° (top), etc. Our analysis was designed to determine whether similar curvature/position information could be decoded from the neural responses.

To derive the population response, we scaled each cell's tuning peak by its response to the shape in question. Thus, each cell "voted" for its preferred boundary fragment with strength proportional to its response rate. The entire set of scaled tuning peaks defined a surface representing coding strength for all combinations of curvature and position. The (Gaussian) smoothed resulting surface is shown as a color image in Fig. 6. Here, color represents the strength of the population representation for the corresponding curvature/position combination. Red represents strong population representation for the corresponding feature, and blue represents weak population representation. The actual curvature function is re-plotted in white for comparison. The population surface contains peaks (red) corresponding to the major boundary features of the

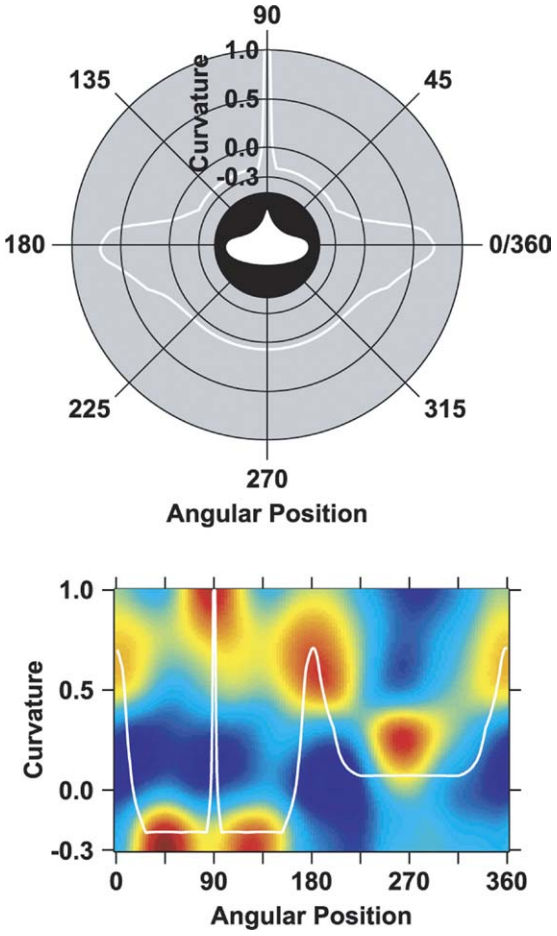

Fig. 6. Estimated population response for an example shape. *Top*: The shape in question is shown in the center. The surrounding white line plots boundary curvature (radial dimension) as a function of angular position (angular dimension) in polar coordinates to highlight the correspondence with boundary features. *Bottom*: Estimated population response across the curvature × position domain (colored surface) with the true curvature function superimposed (white line). *X*-axis represents angular position; *Y*-axis represents curvature. Color scale runs from 0.0 (blue) to 1.0 (red). The peaks and troughs in the curvature function are associated with peaks (red) in the population representation. All the prominent boundary features of the shape are strongly encoded by the V4 population representation. Figure from Pasupathy and Connor (2002).

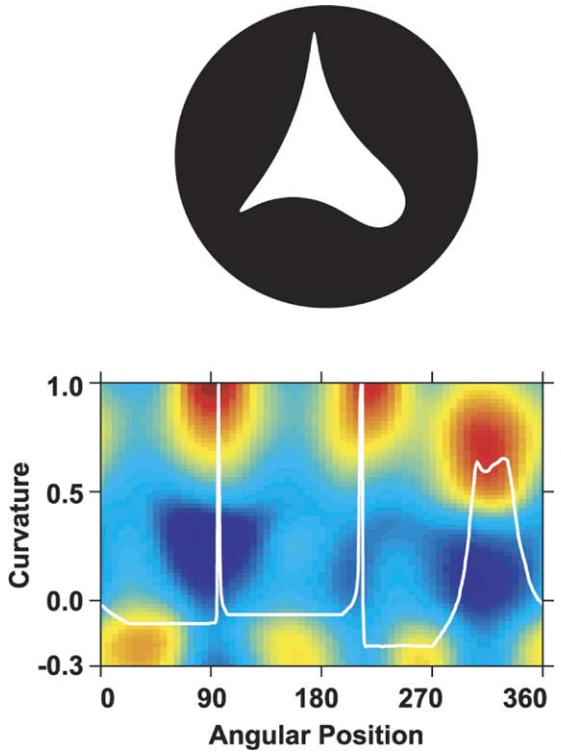

Fig. 7. Estimated population response for a second example shape. *Top*: The shape. *Bottom*: Estimated V4 population response is shown with the curvature function superimposed in white. Axes and details as in Fig. 6. The three strong peaks (red) at convex curvatures and the somewhat weaker peaks (yellow) at concave curvatures correspond well, in terms of curvature and angular position, with the six prominent boundary features of the shape in question.

stimulus: the sharp convexity at 90°, the medium convexities at 0° and 180°, the broad convexity at 270°, and the concavities at 45° and 135°.

A second example is shown in Fig. 7. Again, the population representation contains peaks associated with all of the major boundary features of the stimulus: sharp convexity at 90° and 225°, medium convexity at 315°, and the three intervening concavities. We obtained similar results for all 49 shapes in our stimulus set — the V4 population representation had peaks associated with all of the major boundary features for each shape. Thus, the V4 population signal provides a complete representation of 2D shape contours in terms of contour curvature and angular position. To assess the accuracy of the population representation, for each shape we computed the mean absolute difference between the population peaks and the nearest points on the true curvature × position function. For all shapes the errors in representation were small in both dimensions: median difference was 4.04° along the angular

position dimension and 0.0704 along the curvature dimension. Thus, the peaks in the population surface corresponded closely to the boundary components in the original shape and therefore provided a complete and accurate representation of the stimuli used here. Thus, the V4 population signal carries all the information needed to represent, perceive and recognize the 2D shape contours used here.

Discussion

Previous studies suggest that area V4 encodes shape in terms of edge orientation, spatial frequency, and bar orientation, length and width (Desimone and Schien, 1987). The results presented here add another pair of dimension to this list — contour curvature and object-centered position. Individual V4 neurons carry information about the boundary conformation at a specific location along the shape contour and the V4 population as a whole encodes a complete and accurate representation of the entire 2D contour in terms of its boundary configuration. With these results we can now successfully describe the responses of many V4 neurons to a more general class of curves; such a description using edge and bar orientation models alone is poor. Thus, these findings extend our understanding of V4 responses to a larger class of 2D shapes.

Our results are consistent with studies demonstrating selectivity for complex shapes with curved parts in area V4 (Gallant et al., 1993; Kobatake and Tanaka, 1994; Gallant et al., 1996; Wilkinson et al., 2000) and IT (Schwartz et al., 1983; Tanaka et al., 1991; Janssen et al., 1999). They also reinforce previous evidence for the extraction of coarse curvature information at earlier levels by end-stopped cells (Heggelund and Hohmann, 1975; Dobbins et al., 1987; Versavel et al., 1990). Recently, Ito and Komatsu (2004) suggested that the process of angle extraction may start but not attain completion in area V2 because angle-selective V2 units respond in comparable amounts to the component end-stopped lines as well. Perhaps then, this process of angle and curve encoding which starts in V2 (or even V1) culminates in the representation of 2D boundaries in terms of their contour features as demonstrated here.

The dimensions used to quantify the encoded boundary features need not have been curvature and angular position — other parameterization schemes that describe contour characteristics might be as or more effective. However, as discussed above, results cannot be explained in terms of mass-based orientation parameters, tuning for oriented bars or lower-level factors such as edge orientation, contrast polarity and spatial frequency. Also, response patterns exhibited by these cells cannot be explained in terms of area of stimulation or by differential surround stimulation. Thus, many V4 neurons carry information about 2D object boundaries in terms of component shape and position. At the population level, this representation is complete and accurate. These results suggest that V4 neurons represent complex 2D contours in a piecewise fashion, encoding information about curvature and other characteristics of a section of the shape boundary.

Implications for perception

Several psychophysical studies have reported high sensitivity in humans for the detection of curved elements in visual displays (Andrews et al., 1973; Triesman and Gormican, 1988; Wolfe et al., 1992; Wilson et al., 1997). Our results demonstrating explicit representation of curvature in area V4 provides a physiological basis for this increased sensitivity to curves. In the experiments presented above, when contour features were presented in isolation and in the context of complex shapes, selectivity for sharp convexities was overrepresented while that for concavities was rare. The distribution of Gaussian tuning peaks along the curvature axis was independent of RF eccentricity and showed a consistent bias toward sharp convexities at all eccentricities. This bias was evident even when curvature was finely sampled (in the secondary fine-scale tuning test). These findings suggest a biased representation of shape in terms of sharp convexities. In the complex-shape study, we sampled curvatures in the range of -0.31 (medium concavity) to 1.0 (sharp convexity). It is possible that sharp concavities are also strongly represented in the cortex and the true bias is in

favor of sharp curvatures rather than sharp convexities. However, results from the first experiment with isolated contour features suggested otherwise. In those experiments, the same ranges of concave and convex curvatures were sampled and the results showed a strong bias in favor of convexities. Taken together, these results suggest a bias toward shape representation in terms of sharp convexities.

The overrepresentation of neurons that encode shape in terms of convex projections makes sense from a functional point of view — encoding object shape in terms of sharp convexities leads to a highly efficient and economical shape description, since high curvature regions appear to be rich sources of shape information (Attneave, 1954). The neuronal bias toward convex features may also underlie their perceptual salience. Researchers have found that convex projections, rather than concave indentations, provide the basis for figure/ground interpretations (Kaniza and Gerbino, 1976) and shape similarity judgements (Subirana-Vilanova and Richards, 1996) in human observers. The greater representation of neurons that encode shape in terms of convex projections may underlie the perceptual significance of convex features. Our results provide neurophysiological evidence in support of the "curvature minima" rule (Hoffman and Richards, 1984), which hypothesizes that object segmentation, for the purpose of shape recognition, should occur along boundaries of maximum concavity, producing convex parts. Psychophysical results (Braunstein et al., 1989) in humans support this rule — observers are more likely to recognize parts from a previously viewed object if the parts are convex. Our results suggest that shape representation in the ventral visual pathway proceeds in accordance with the curvature minima rule.

Transformation from local oriented signals to complex shape selectivity

How might tuning for angles and curves in area V4 arise from tuning for local orientation and spatial frequency information available at earlier visual areas? Theorists have proposed that tuning for sharp angles could be achieved by an appropriate combination of end-stopped orientation signals (Milner, 1974; Hummel and Biederman, 1992). For example, the response pattern in Fig. 1A could be achieved by combining the signals from units tuned to a 0°-oriented edge end-stopped at the top and a 90° edge end-stopped at the right, with a preference for a specific contrast direction (brighter to the right). The end result would be a signal related to the presence of a sharp corner pointing to the upper left. The tuning width of the pooled units may then produce the tuning for contour feature orientation and acuteness, and integrating signals from multiple positions could yield position invariance. Transformation from oriented signals to angle representation possibly requires multiple stages of processing, since both in V2 (Ito and Komatsu, 2004) and in V4 there are intermediate units that respond strongly to the preferred angle and its component end-stopped lines.

Cells selective for angles constructed in the above fashion would also respond to curves that contain the appropriate component orientations. Cells that respond preferentially to curves and are tuned to contour curvature may also be derived by integrating an appropriate pattern of local orientation and curvature signals from end-stopped cells in preceding visual areas. Recently Cadieu et al. (2005) have demonstrated that V4-like curvature tuning can be achieved by combining inputs from position-invariant orientation-selective Gabor filters. This is equivalent to pooling appropriate orientation selective units from V1 or V2, followed by nonlinear processing, to derive the necessary curvature-tuned responses. Alternatively, selective responses to curved contours in area V4 may be achieved by integration of V2 signals modeled as linear-nonlinear-linear filters (Wilson, 1999). A third hypothesis suggests that position-independent tuning for corners and curves may be achieved by the processing of local orientation end-stopped signals in the dendrites of V4 neurons (Zucker et al., 1989). Further experiments are required to decide between these possible mechanisms.

At a more global level, the progression from local orientation signals to selectivity for contour conformation supports the proposal of many feature-based shape recognition schemes that suggest contour features as intermediate level shape

primitives. Our results imply a coding scheme based on structural description of the shape, consistent with the notion of representation by parts or components. As outlined earlier, in these models shapes are described in terms of the conformations and relative positions (and/or connectivity) of their simpler components (Marr and Nishihara, 1978; Hoffmann and Richards, 1984; Biederman, 1987; Dickenson et al., 1992; Riesenhuber and Poggio, 1999; Edelman and Intrator, 2000). Our findings are consistent with several recent studies that support this idea of a parts-based representation in IT (Tanaka et al., 1991; Wang et al., 2000; Op de Beeck et al., 2001; Tsunoda et al., 2001; Sigala and Logothetis, 2002).

A parts-based representation is advantageous because of its insensitivity to variations in the retinal image of an object and its alphabet-like power to encode an infinite variety of shapes. Most parts-based theories envision a hierarchical progression of parts complexity through a sequence of processing stages. Boundary fragments constitute an appropriate level of parts complexity for an intermediate processing stage like V4. More complex parts are encoded in IT, the next stage in the ventral pathway (Desimone et al., 1984; Tanaka et al., 1991; Tsunoda et al., 2001). In a recent study Brincat and Connor (2004) demonstrated that the responses of posterior IT neurons reflect (linear and nonlinear) integration of information about the characteristics and relative positions of 2–4 contour segments of complex shapes. Further integration in anterior IT could lead to selectivity for more complex features and sparser responses (Tanaka, 1996; Tsunoda et al., 2001; Edelman and Intrator, 2003) or it may culminate in holistic coding for global object shape (Logothetis and Sheinberg, 1996; Booth and Rolls, 1998; Ullman, 1998; Riesenhuber and Poggio, 1999; Baker et al., 2002).

Areas of future investigation

The results presented in this review further our understanding about the basis of 2D contour representation in area V4. Due to practical considerations, our stimuli were restricted to be a small subset of 2D contours — silhouettes with no internal structure — and so our results pertain to the representation of the boundaries of 2D contours. To gain a more complete understanding of 2D contour representation and to extend our results to more naturalistic stimuli further investigations are required.

Our stimuli were 2D contours with convex projections radiating out from the center. In future experiments, these results need to be extended to shapes containing curve segments at other orientations as well. To derive more precise tuning functions of V4 neurons, boundary curvature should be more densely sampled. In this study, stimuli were constructed with a few constant curvature segments. To investigate the dependence of V4 responses on the rate of change or other functions of curvature, responses to stimuli with continuously varying contour curvature needs to be studied. For the stimuli used here, contour representation in terms of curvature and angular position was sufficient (since the stimuli were constructed by combining constant curvature segments) and economical (since it resulted in a small set of points representing each stimulus). For other stimuli, however, representation based on constant curvature segments may be uneconomical (for e.g. stimuli with a high frequency of convex projections along the contour) or insufficient (for e.g. stimuli with long segments of changing curvature). In such cases, it is possible that neurons would encode information in terms of additional stimulus dimensions — a question to be addressed in future studies.

Several other questions need to be addressed — how do V4 neurons respond to more realistic stimuli such as objects with internal structure, simultaneous presentation of contours of multiple objects, stimuli with texture and 3D information or stimuli without explicit contours such as photographic images? In such cases, a different, higher dimensionality would be required to describe the responses of neurons. Finally, only one-third of the neurons that we studied showed systematic tuning for contour characteristics. Amongst the other neurons many were selective for previously described properties such as edge orientation, contrast polarity, oriented bars, etc. A few neurons showed no discernible pattern in their responses or

did not respond to any stimuli presented here. These cells may encode shape in terms of other stimulus dimensions not tested here, such as texture patterns (Hanazawa and Komatsu, 2001), 3D stimuli (Hinkle and Connor, 2002; Watanabe et al., 2002; Hegde and Van Essen, 2005; Hinkle and Connor, 2005; Tanabe et al., 2005), kinetic boundaries (Mysore et al., 2006) or other properties that have not been previously discovered in area V4 — all questions to be addressed in future experiments.

Conclusion

Our results demonstrate that area V4 encodes 2D contours as aggregates of boundary fragments, with individual neuron encoding the characteristics of a specific fragment. This result has provided a first step toward deciphering the bases of shape representation in area V4. The general method used here could be extended to investigate the shape dimensions that dictate V4 responses to more naturalistic stimuli. Targeted investigation of shape dimensions, parametric design of stimuli combined with appropriate analytical techniques will help us identify the other bases of representation in the ventral pathway. This will bring us closer to solving the puzzle of how the primate brain parses the 2D representation of the visual world into 3D objects that are perceived and recognized.

Acknowledgements

Work described in this review was done in the laboratory of C. E. Connor at Zanvyl Krieger Mind/Brain Institute, Johns Hopkins University, Baltimore, MD. Ideas presented here were developed in collaboration with C. E. Connor. Many thanks to Mark H. Histed for useful suggestions on the manuscript and Kristin J. MacCully for assistance in its preparation.

References

Andrews, D.P., Butcher, A.K. and Buckley, B.R. (1973) Acuities for spatial arrangement in line figures: human and ideal observers compared. Vision Res., 13: 599–620.

Attneave, F. (1954) Some informational aspects of visual perception. Psychol. Rev., 61: 183–193.

Baker, C.I., Behrmann, M. and Olson, C.R. (2002) Impact of learning on representation of parts and wholes in monkey inferotemporal cortex. Nat. Neurosci., 5: 1210–1216.

Baizer, J.S., Robinson, D.L. and Dow, B.M. (1977) Visual responses of area 18 neurons in awake, behaving monkey. J. Neurophysiol., 40: 1024–1037.

Besl, P.J. and Jain, R.C. (1985) Three dimensional object recognition. Comput. Surv., 17(1): 75–145.

Biederman, I. (1987) Recognition-by-components: a theory of human image understanding. Psychol. Rev., 94: 115–147.

Biederman, I. and Cooper, E.E. (1991) Priming contour-deleted images: evidence for intermediate representations in visual object recognition. Cogn. Psychol., 23(3): 393–419.

Booth, M.C. and Rolls, E.T. (1998) View-invariant representations of familiar objects in the inferior temporal visual cortex. Cereb. Cortex, 8: 510–523.

Braunstein, M.L., Hoffmann, D.D. and Saidpour, A. (1989) Parts of visual objects: an experimental test of the minima rule. Perception, 18: 817–826.

Brincat, S.L. and Connor, C.E. (2004) Underlying principles of visual shape selectivity in posterior inferotemporal cortex. Nat. Neurosci., 7(8): 880–886.

Burkhalter, A. and Van Essen, D.C. (1986) Processing of color, form, and disparity information in visual area VP and V2 of ventral extrastriate cortex in the macaque monkey. J. Neurosci., 6: 2327–2351.

Cadieu, C., Kouh, M., Riesenhuber, M. and Poggio, T. (2005) Shape representation in V4: investigating position-specific tuning for boundary conformation with the standard model of object recognition. Cosyne Abstracts.

Chen, S. and Levi, D.M. (1996) Angle judgment: is the whole the sum of its parts? Vision Res., 36: 1721–1735.

De Weerd, P., Desimone, R. and Ungerleider, L.G. (1996) Cue-dependent deficits in grating orientation discrimination after V4 lesions in macaques. Vis. Neurosci., 13: 529–538.

Desimone, R. and Gross, C.G. (1979) Visual areas in the temporal cortex of the macaque. Brain Res., 178: 363–380.

Desimone, R., Albright, T.D., Gross, C.G. and Bruce, C. (1984) Stimulus-selective properties of inferior temporal neurons in the macaque. J. Neurosci., 4: 2051–2062.

Desimone, R. and Schein, S.J. (1987) Visual properties of neurons in area V4 of the macaque: sensitivity to stimulus form. J. Neurophysiol., 57: 835–868.

Desimone, R., Schien, S.J., Moran, J. and Ungerleider, L.G. (1985) Contour, color and shape analysis beyond the striate cortex. Vision Res., 25: 441–452.

Dickenson, S.J., Pentland, A.P. and Rosenfeld, A. (1992) From volumes to views: an approach to 3-D object recognition. CVGP: Image Understanding, 55: 130–154.

Dobbins, A., Zucker, S.W. and Cynader, M.S. (1987) End-stopped neurons in the visual cortex as a substrate for calculating curvature. Nature, 329: 438–441.

Edelman, S. and Intrator, N. (2000) (Coarse coding of shape fragments) + (retinotopy) approximately = representation of structure. Spat Vis. 13(2–3): 255–264.

Edelman, S. and Intrator, N. (2003) Towards structural systematicity in distributed, statistically bound visual representations. Cogn. Sci., 27: 73–109.

Felleman, D.J. and Van Essen, D.C. (1991) Distributed hierarchical processing in the primate cerebral cortex. Cereb. Cortex, 1: 1–47.

Fujita, I., Tanaka, K., Ito, M. and Cheng, K. (1992) Columns for visual features of objects in monkey inferotemporal cortex. Nature, 360: 343–346.

Gallant, J.L. (2000) The neural representation of shape. In: DeValois, K.K. and DeValois, R.L. (Eds.), Seeing. Academic Press, San Diego, CA, pp. 311–333.

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.

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.

Gallant, J.L., Shoup, R.E. and Mazer, J.A. (2000) A human extrastriate area functionally homologous to macaque V4. Neuron, 27: 227–235.

Gattass, R., Gross, C.G. and Sandell, J.H. (1981) Visual Topography of V2 in the macaque. J. Comp. Neurol., 201: 519–539.

Gross, C.G., Rocha-Miranda, C.E. and Bender, D.B. (1972) Visual properties of neurons in inferotemporal cortex of the macaque. J. Neurophysiol., 35: 96–111.

Hanazawa, A. and Komatsu, H. (2001) Influence of the direction of elemental luminance gradients on the responses of V4 cells to textured surfaces. J. Neurosci., 21: 4490–4497.

Heeley, D.W. and Buchanan-Smith, H.M. (1996) Mechanisms specialized for the perception of image geometry. Vision Res., 36: 3607–3627.

Hegde, J. and Van Essen, D.C. (2000) Selectivity for complex shapes in primate visual area V2. J. Neurosci., 20(5): RC61.

Hegde, J. and Van Essen, D.C. (2005) Role of primate visual area V4 in the processing of 3-D shape characteristics defined by disparity. J. Neurophysiol., 94: 2856–2866.

Heggelund, P. and Hohmann, A. (1975) Responses of striate cortical cells to moving edges of different curvatures. Exp. Brain Res., 31: 329–339.

Heywood, C.A. and Cowey, A. (1987) On the role of cortical area V4 in the discrimination of hue and pattern in macaque monkeys. J. Neurosci., 7: 2601–2617.

Hinkle, D.A. and Connor, C.E. (2002) Three-dimensional orientation tuning in macaque area V4. Nat. Neurosci., 5(7): 665–670.

Hinkle, D.A. and Connor, C.E. (2005) Quantitative characterization of disparity tuning in ventral pathway area V4. J. Neurophysiol., 94: 2726–2737.

Hoffmann, D.D. and Richards, W.A. (1984) Parts of recognition. Cognition, 18: 65–96.

Hubel, D.H. and Weisel, T.N. (1959) RFs of single neurones in the cat's striate cortex. J. Physiol. (Lond.), 148: 574–591.

Hubel, D.H. and Weisel, T.N. (1965) RFs and functional architecture in two nonstriate visual areas (18 and 19) of the cat. J. Neurophysiol., 28: 229–289.

Hubel, D.H. and Weisel, T.N. (1968) RFs and functional architecture of monkey striate cortex. J. Physiol. (Lond.), 195: 215–243.

Hubel, D.H. and Livingstone, M.S. (1987) Segregation of form, color, and stereopsis in primate area 18. J. Neurosci., 7: 3378–3415.

Hummel, J.E. and Biederman, I. (1992) Dynamic binding in a neural network for shape recognition. Psychol. Rev., 99: 480–517.

Ito, M. and Komatsu, H. (2004) Representation of angles embedded within contour stimuli in area V2 of macaque monkeys. J. Neurosci., 24(13): 3313–3324.

Ito, M., Tamura, H., Fujita, I. and Tanaka, K. (1995) Size and position invariance of neuronal responses in monkey inferotemporal cortex. J. Neurophysiol., 73(1): 218–226.

Janssen, P., Vogels, R. and Orban, G.A. (1999) Macaque inferior temporal neurons are selective for disparity-defined three-dimensional shapes. Proc. Natl. Acad. Sci. USA, 96: 8217–8222.

Kaniza, G. and Gerbino, W. (1976) Convexity and symmetry in figure-ground organization. In: Henle, M. (Ed.), Art and Artefacts. Springer, New York, pp. 25–32.

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.

Lennie, P. (2003) Receptive fields. Curr. Biol., 13(6): R216–R219.

Logothetis, N.K. and Sheinberg, D.L. (1996) Visual object recognition. Annu. Rev. Neurosci., 19: 577–621.

Marr, D. and Nishihara, H.K. (1978) Representation and recognition of the spatial organization of three-dimensional shapes. Proc. R. Soc. Lond B Biol. Sci., 200: 269–294.

Merigan, W.H. (1996) Basic visual capacities and shape discrimination after lesions of extrastriate area V4 in macaques. Vis. Neurosci., 13: 51–60.

Merigan, W.H., Nealy, T.A. and Maunsell, J.H.R. (1993) Visual effects of lesions of cortical area V2 in macaques. J. Neurosci., 13: 3180–3191.

Merigan, W.H. and Phan, H.A. (1998) V4 lesions in macaques affect both single- and multiple-viewpoint shape discriminations. Vis. Neurosci., 15: 359–367.

Milner, P.M. (1974) A model for visual shape recognition. Psychol. Rev., 81: 521–535.

Mysore, S.G., Vogels, R., Raiguel, S.E. and Orban, G.A. (2006) Processing of kinetic boundaries in macaque V4. J. Neurophysiol. 95(3): 1864–1880.

Op de Beeck, H., Wagemans, J. and Vogels, R. (2001) Inferotemporal neurons represent low-dimensional configurations of parameterized shapes. Nat. Neurosci., 4: 1244–1252.

Pasupathy, A. and Connor, C.E. (1999) Responses to contour features in macaque area V4. J.Neurophysiol., 82: 2490–2502.

Pasupathy, A. and Connor, C.E. (2001) Shape representation in area v4: position-specific tuning for boundary conformation. J. Neurophysiol., 86: 2505–2519.

Pasupathy, A. and Connor, C.E. (2002) Population coding of shape in area V4. Nat. Neurosci., 5: 1332–1338.

Pentland, A. (1989) Shape information from shading: a theory about human perception. Spatial Vis., 4(2–3): 165–182.

Perrett, D.I., Rolls, E.T. and Caan, W. (1982) Visual neurones responsive to faces in the monkey temporal cortex. Exp. Brain Res., 47: 329–342.

Poggio, T. and Edelman, S. (1990) A network that learns to recognize three-dimensional objects. Nature, 343: 263–266.

Regan, D., Gray, R. and Hamstra, S.J. (1996) Evidence for a neural mechanism that encodes angles. Vision Res., 36: 323–330.

Riesenhuber, M. and Poggio, T. (1999) Hierarchical models of object recognition in cortex. Nat. Neurosci., 2: 1019–1025.

Sary, G., Vogels, R. and Orban, G.A. (1993) Cue-invariant shape selectivity of macaque inferior temporal neuron. Science, 260: 95–997.

Sato, T., Kawamura, T. and Iwai, E. (1980) Responsiveness of inferotemporal single units to visual pattern stimuli in monkeys performing discrimination. Exp. Brain Res., 38(3): 313–319.

Schiller, P.H. (1995) Effects of lesions in visual cortical area V4 on the recognition of transformed objects. Nature, 376: 342–344.

Schiller, P.H. and Lee, K. (1991) The role of primate extrastriate area V4 in vision. Science, 251: 1251–1253.

Schwartz, E.L., Desimone, R., Albright, T.D. and Gross, C.G. (1983) Shape recognition and inferior temporal neurons. Proc. Natl. Acad. Sci. USA, 80: 5776–5778.

Sigala, N. and Logothetis, N.K. (2002) Visual categorization shapes feature selectivity in the primate temporal cortex. Nature, 415: 318–320.

Subirana-Vilanova, J.B. and Richards, W. (1996) Attentional frames, frame curves and figural boundaries: the outside/inside dilemma. Vision Res., 36: 1493–1501.

Tanabe, S., Doi, T., Umeda, K. and Fujita, I. (2005) Disparity-tuning characteristics of neuronal responses to dynamic random-dot stereograms in macaque visual area V4. J. Neurophysiol., 94: 2683–2699.

Tanaka, K. (1993) Neuronal mechanisms of object recognition. Science, 262: 685–688.

Tanaka, K. (1996) Inferotemporal cortex and object vision. Annu. Rev. Neurosci., 19: 109–139.

Tanaka, K., Saito, H., Fukada, Y. and Moriya, M. (1991) Coding visual images of objects in the inferotemporal cortex of the macaque monkey. J. Neurophysiol., 66: 170–189.

Triesman, A. and Gormican, S. (1988) Feature analysis in early vision: evidence from search asymmetries. Psychol. Rev., 95: 15–48.

Tsunoda, K., Yamane, ., Nishizaki, M. and Tanifuji, M. (2001) Complex objects are represented in macaque inferotemporal cortex by the combination of feature columns. Nat. Neurosci., 4: 832–838.

Ullman, S. (1989) Aligning pictorial descriptions: an approach to object recognition. Cognition, 32: 193–254.

Ullman, S. (1998) Three-dimensional object recognition based on the combination of views. Cognition, 67: 21–44.

Ungerleider, L. and Mishkin, M. (1982) Two cortical visual systems. In: Ingle, D.J., Goodale, M.A. and Mansfield, R.J.W. (Eds.), Analysis of Visual Behavior. MIT Press, Cambridge, MA, pp. 549–586.

Versavel, M., Orban, G.A. and Lagae, L. (1990) Responses of visual cortical neurons to curved stimuli and chevrons. Vision Res., 30: 235–248.

von der Heydt, R. and Peterhans, E. (1989) Mechanisms of contour perception in monkey visual cortex. I. Lines of pattern discontinuity. J. Neurosci., 9: 1731–1748.

von der Heydt, R., Zhou, H. and Friedman, H.S. (2000) Representation of stereoscopic edges in monkey visual cortex. Vision Res., 40(15): 1955–1967.

Wang, Y., Fujita, I. and Murayama, Y. (2000) Neuronal mechanisms of selectivity for object features revealed by blocking inhibition in inferotemporal cortex. Nat. Neurosci., 3: 807–813.

Wang, G., Tanaka, K. and Tanifuji, M. (1996) Optical imaging of functional organization in the monkey inferotemporal cortex. Science, 272: 1665–1668.

Watanabe, M., Tanaka, H., Uki, T. and Fujita, I. (2002) Disparity-selective neurons in area V4 of macaque monkeys. J. Neurophysiol., 87: 1960–1973.

Wilkinson, F., James, T.W., Wilson, H.R., Gati, J.S., Menon, R.S. and Goodale, M.A. (2000) An fMRI study of the selective activation of human extrastriate form vision areas by radial and concentric gratings. Curr. Biol., 10: 1455–1458.

Wilson, H.R. (1999) Non-Fourier cortical processes in texture, form and motion perception. In: Peters, A. and Jones, E.G. (Eds.), Cerebral Cortex: Models of Cortical Circuitry. Plenum, New York, pp. 445–477.

Wilson, H.R., Wilkinson, F. and Asaad, W. (1997) Concentric orientation summation in human form vision. Vision Res., 37: 2325–2330.

Wolfe, J.M., Yee, A. and Friedman-Hill, S.R. (1992) Curvature is a basic feature for visual search tasks. Perception, 21: 465–480.

Zucker, S.W., Dobbins, A. and Iverson, L. (1989) Two stages of curve detection suggest two styles of visual computation. Neural Comput., 1: 68–81.