

Visual Functions of Primate Area V4

Anitha Pasupathy,^{1,2} Dina V. Popovkina,³
and Taekjun Kim^{1,2}

¹Department of Biological Structure, University of Washington, Seattle, Washington 98195, USA; email: pasupat@uw.edu, taekjun@uw.edu

²Washington National Primate Research Center, University of Washington, Seattle, Washington 98121, USA

³Department of Psychology, University of Washington, Seattle, Washington 98105, USA; email: dina4@uw.edu

Annu. Rev. Vis. Sci. 2020. 6:363–85

First published as a Review in Advance on
June 24, 2020

The *Annual Review of Vision Science* is online at
vision.annualreviews.org

<https://doi.org/10.1146/annurev-vision-030320-041306>

Copyright © 2020 by Annual Reviews.
All rights reserved

**ANNUAL
REVIEWS CONNECT**

www.annualreviews.org

- Download figures
- Navigate cited references
- Keyword search
- Explore related articles
- Share via email or social media

Keywords

shape perception, object recognition, texture perception, macaque area V4, goal-oriented representation

Abstract

Area V4—the focus of this review—is a mid-level processing stage along the ventral visual pathway of the macaque monkey. V4 is extensively interconnected with other visual cortical areas along the ventral and dorsal visual streams, with frontal cortical areas, and with several subcortical structures. Thus, it is well poised to play a broad and integrative role in visual perception and recognition—the functional domain of the ventral pathway. Neurophysiological studies in monkeys engaged in passive fixation and behavioral tasks suggest that V4 responses are dictated by tuning in a high-dimensional stimulus space defined by form, texture, color, depth, and other attributes of visual stimuli. This high-dimensional tuning may underlie the development of object-based representations in the visual cortex that are critical for tracking, recognizing, and interacting with objects. Neurophysiological and lesion studies also suggest that V4 responses are important for guiding perceptual decisions and higher-order behavior.

INTRODUCTION

Human and nonhuman primates rely heavily on their visual system to make critical decisions in daily life. When a monkey swings through the rainforest, or when humans manipulate objects, interpret facial expressions, or look skyward to assess the rain clouds, detailed processing of a visual scene is important. Success depends on perceiving the shape of objects and the textural quality of surfaces, what Adelson & Bergen (1991) refer to as the things and the stuff, respectively, in a visual scene. In the primate brain, this is based on information processing along the ventral visual pathway, which runs from cortical area V1 through V2 and V4 to the subregions of the inferotemporal (IT) cortex. In this review, we focus on V4, a mid-level processing stage along this pathway. Our thesis is that area V4 is important for parsing objects from visual scenes and for generating goal-oriented representations of objects to facilitate perception, decision making, and behavior. In other words, V4 is where a representation for “things” begins to emerge. Based on supporting anatomical, neurophysiological, behavioral, and lesion literature, we argue that:

1. V4 is important for the emergence of object-based representations in the ventral visual cortex.
2. V4 representations guide perceptual decisions and behavior. In this sense, they are goal-oriented representations (as opposed to simply representing the retinal image).

ANATOMICAL CONNECTIONS OF AREA V4

V4 is characterized as a mid-level visual processing stage that receives inputs primarily from area V2 and sends outputs to IT cortex. However, this characterization belies V4's extensive connectivity (**Figure 1**). While much of peripheral V4 receives feedforward projections primarily from V2, central representations receive projections from V1 that bypass V2 (Nakamura et al. 1993). This connectivity could explain the short latencies of many V4 neurons (**Figure 2**). V4 also sends feedback projections to both V1 and V2. V4 is reciprocally interconnected with areas in the dorsal stream, including V3, the middle temporal area (MT), the medial superior temporal area (MST), and the lateral intraparietal area (LIP), and thus may be important for processing spatial information and dynamic stimuli (**Figure 1**). V4 is also directly interconnected with frontal areas, frontal eye fields (FEF), and the ventrolateral prefrontal cortex (vLPFC) (Barbas & Mesulam 1985, Ninomiya et al. 2012b, Ungerleider et al. 2008), consistent with its proposed role in attentional processing and decision making. Recent studies with trans-synaptic rabies virus have demonstrated disynaptic and trisynaptic feedback to V4 from the medial temporal lobe structures and the hippocampus (Ninomiya et al. 2012a). Anterograde and retrograde tracing studies have identified afferent, efferent, and bidirectional connections between V4 and several subcortical structures, including the caudate and putamen of the basal ganglia, the amygdala, the pulvinar, the lateral geniculate nucleus, the claustrum, and the superior colliculus (Gattass et al. 2014). Among the visual cortical areas that have been well studied (in terms of anatomical connectivity), V4 is the most interconnected, with robust pathways to 21 areas, suggesting that V4 may play a pivotal role in different aspects of cortical processing (Felleman & Van Essen 1991).

This extensive connectivity supports a comprehensive role for V4 not only in representing visual stimuli, but also in recognition and attention. V4 is uniquely positioned to execute computations associated with both an early processing stage and a high-level stage. This is reflected in the broad range of response latencies observed in V4 (**Figure 2**): Many neurons respond soon after stimulus onset with latencies that rival neurons in MT and FEF, while many others respond with long latencies comparable to neurons in IT cortex.

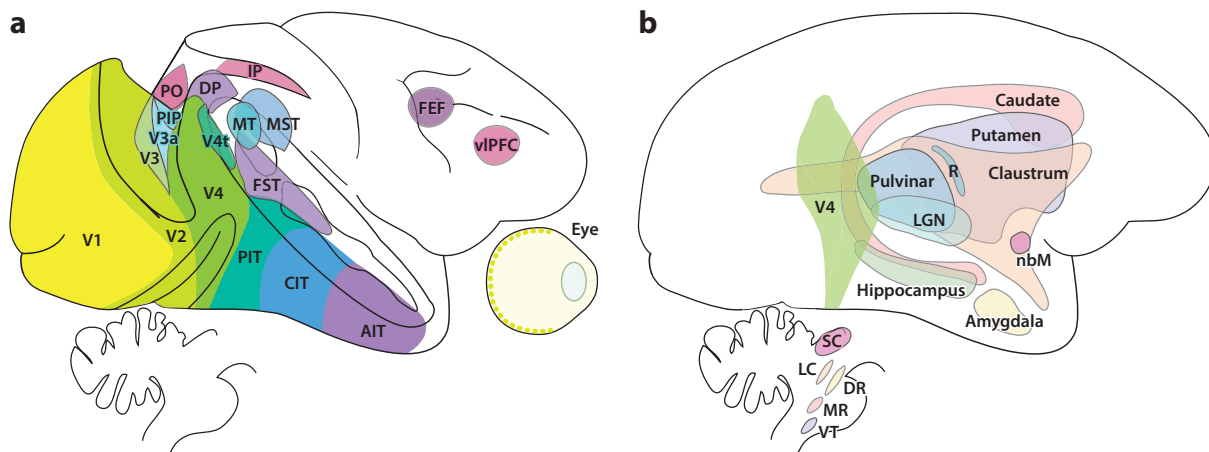


Figure 1

Interconnections between V4 and other brain regions. Areas connected with V4 are highlighted on the lateral surface of the macaque monkey brain. (a) Visual cortical areas along the ventral stream, many along the dorsal visual stream and frontal areas, and (b) the medial temporal lobe and many subcortical structures are interconnected with V4. Figure created using data from Felleman & Van Essen (1991), Parker (2007), Ungerleider et al. (2008), Gattass et al. (2014), and Ninomiya et al. (2012a,b). Part of the figure was adapted with permission from Pasupathy et al. (2018). Abbreviations: AIT, anterior inferotemporal area; CIT, central inferotemporal area; DP, dorsal prelunate area; DR, dorsal raphe; FEF, frontal eye field; FST, fundus of the superior temporal sulcus area; IP, intraparietal areas; LC, locus coeruleus; LGN, lateral geniculate nucleus; MR, medial raphe; MST, medial superior temporal area; MT, middle temporal area; nbM, basal nucleus of Meynert; PIP, posterior intraparietal area; PIT, posterior inferotemporal area; PO, parieto-occipital area; R, thalamic reticular formation; SC, superior colliculus; V3a, visual complex V3 part A; V4t, V4 transition zone; vIPFC, ventrolateral prefrontal cortex; VT, ventral tegmentum.

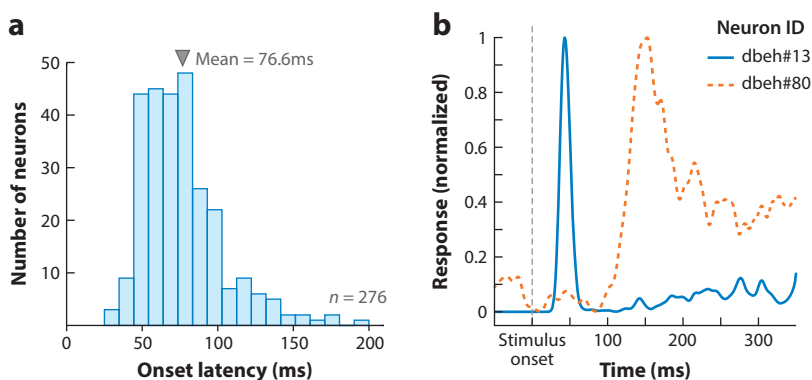


Figure 2

Distribution of response onset latencies of V4 neurons. (a) Onset latencies of 276 V4 neurons from three monkeys. Latency measurements were based on the responses of single V4 neurons in animals engaged in a passive fixation task. As animals fixated a variety of shape stimuli were presented within the receptive field of the cell under study. Peristimulus time histograms (PSTHs) were constructed based on evoked responses of each neuron and onset latencies were calculated based on the half-height method, i.e., as the time point at which the PSTH exceeded the mean of the peak and baseline rates. Onset latencies ranged from 25 ms to 200 ms; mean latency was 76.6 ms. For further details, see Zamarashkina et al. (2020). (b) PSTHs of two example V4 neurons studied during passive fixation, displaying approximately 100 ms difference in onset latency. PSTHs were aligned on stimulus onset and smoothed with a Gaussian kernel ($\sigma = 4$ ms).

ENCODING IMAGE STRUCTURE IN A VISUAL SCENE

Studies over decades have demonstrated that, in V1, the first stage of cortical visual processing, neurons encode visual stimuli in terms of local orientation and spatial frequency (Albrecht et al. 1980; Hubel & Wiesel 1959, 1968; Movshon et al. 1978a,b). Specifically, each small patch of the retinal image is represented by a subpopulation of V1 neurons that signals the orientation and scale of image features at a particular position in visual space, and thousands of such V1 subpopulations that tile the visual field represent the full retinal image. In the next processing stage, in V2, many neurons may be specialized for encoding the structure in visual textures (Movshon & Simoncelli 2014). Texture may be thought of as an aggregate of similar elements in a somewhat ordered spatial arrangement (e.g., a crowd of people, a bunch of leaves, etc.) that together give the impression of a homogeneous surface despite lacking uniformity in color or luminance (Van Gool et al. 1985). Because of its ordered structure, texture can be described in terms of higher-order correlational statistics across position, scale, and orientation (Portilla & Simoncelli 2000). V2 is the first stage in visual cortex to exhibit sensitivity to multipoint correlations in binary pixel images (Victor & Conte 2012, Yu et al. 2015) and may thus be well-suited for encoding naturalistic texture (Freeman et al. 2013, Okazawa et al. 2017, Ziemba et al. 2016). Neurons in V2, unlike those in V1, respond more strongly to image patches that contain the correlational statistics found in natural visual texture than to images that lack them. This preferential encoding in V2 is also evident in human functional magnetic resonance imaging (fMRI) results and is predictive of human ability to detect naturalistic texture. V2 neurons also respond similarly to a family of textures that share correlational statistics, despite variations in local image detail, providing a neuronal basis for texture classification. Taken together, these results support the idea that V2 neurons selectively and differentially encode naturalistic texture and may provide the neuronal basis for its perception.

IMPORTANCE OF OBJECT-BASED REPRESENTATIONS

Selectivity for naturalistic texture in V2 has typically been characterized with homogeneous, circular texture patches. Thus, these results do not reveal how the presence of a boundary within the receptive field (RF), defined by a texture contrast or, more generally, a heterogeneous texture, might influence V2 encoding. However, based on psychophysical studies (Freeman & Simoncelli 2011, Rosenholtz et al. 2012), the dominant view has been that V2 neuronal responses largely reflect the summary statistics within the image patch, without an explicit representation of contrasts that signify object boundaries (but see Wallis et al. 2019). Object boundaries may influence responses insofar as they influence the correlational statistics, but there is no proposed mechanistic basis for a distinction between object boundaries and texture.

In natural vision, however, segmenting object boundaries and representing the shape of component objects are critically important for several reasons. First, they facilitate shape perception, which is critical for object recognition. Even though familiar objects can be recognized based on color and texture, shape appears to be the fastest and most important cue (Elder & Velisavljevic 2009). Shape perception is also important for guiding motor plans for manipulating or interacting with objects (Cooke et al. 2007, Schaffelhofer & Scherberger 2016).

More generally, object segmentation, recognition, and tracking could provide a framework for the visual system to interpret the abundant information in crowded and dynamic scenes. Object-based representations may be critical for attentional selection (Kahneman et al. 1992, Treisman et al. 1983), for continuity of scene perception across saccades, for saccade planning, and for overall scene understanding. For example, in natural scenes, preferential viewing locations and corrective saccades are associated more strongly with object center than with salient image features

(Nuthmann & Henderson 2010, Schut et al. 2017). Object representations can also serve as anchors to enable interpretation in conditions where visual information is missing or ambiguous: Attention can be directed to illusory as well as real objects (Martinez et al. 2007), and inferred scene structure, not raw image content, drives reconstruction of corrupted natural scenes (Neri 2017). Additionally, objects, rather than object features, may function as fundamental units in visual working memory (Balaban et al. 2019, Luria & Vogel 2011, Markov et al. 2019).

Overall, the establishment of object-based representations in area V4 could enable several functions that maintain a stable and unambiguous percept of the visual world.

PRECURSOR TO AN OBJECT-BASED CODE-BOUNDARY ENHANCEMENT

Two characteristics of V1 surround modulation—collinear facilitation and flexible surround suppression—may be important first steps for enhanced encoding of object boundaries in subsequent processing stages in visual cortex. Several studies have demonstrated that collinear elements outside the RF could facilitate V1 neuronal responses, especially in the presence of randomly oriented distractors (Bauer & Heinze 2002, Kapadia et al. 1995, Polat et al. 1998). Collinear facilitation processes in V2 appear to integrate depth information consistent with the segmentation of surfaces and partially occluded boundaries (Bakin et al. 2000). Many neurons exhibit tuned surround suppression when the RF center and surround are activated by stimuli with similar characteristics (see, e.g., Blakemore & Tobin 1972, Cavanaugh et al. 2002, Kapadia et al. 1995, Knierim & Van Essen 1992, Levitt & Lund 1997, Nelson & Frost 1985, Nothdurft et al. 1999), which could enable the preferential encoding of texture boundaries that may be compared to uniform texture regions (Nothdurft et al. 1999), observed even in an anesthetized animal. Furthermore, signals related to texture boundaries emerge at roughly the same time in V1 and V4 (Poort et al. 2012), and they emerge in V1 even when V2 is inactivated (Hupé et al. 2001), suggesting that edge enhancement is based on local computations rather than on feedback mechanisms. Coen-Cagli and colleagues (2015) recently demonstrated that the strength of surround suppression in V1 for natural images depends on image similarity in the center and surround of the RF—suppression is stronger for homogeneous images and weaker for images that differentially stimulate the center and surround. Taken together, these results support the idea that contextual modulations in the visual cortex serve to detexturize images—enhancing contour representations at the expense of textures—and thus facilitate the signaling of object boundaries (Gheorghiu et al. 2014). Such enhanced encoding of contrast boundaries (as opposed to texture elements) may be a critical first step for the emergence of object-based representations.

EVIDENCE FOR OBJECT-BASED REPRESENTATIONS IN V4

Several lines of evidence support the hypothesis that object-based representations begin to emerge in V4. In a series of experiments, first with isolated angles and curves (Pasupathy & Connor 1999) and then with simple shapes (Pasupathy & Connor 2001), Pasupathy & Connor demonstrated that V4 neurons are selective for specific boundary conformations. Two examples are shown in **Figure 3a**. Both neurons respond strongly to some shapes and weakly to others. For the top neuron, the preferred shapes included a sharp convexity to the upper right, while shapes with a broad convexity at that same location evoked weak responses. Such shape selectivity could be based on an RF hotspot, i.e., stimulation of a specific RF subregion to the upper right could underlie strong responses. To rule out this possibility, several past studies have investigated whether shape

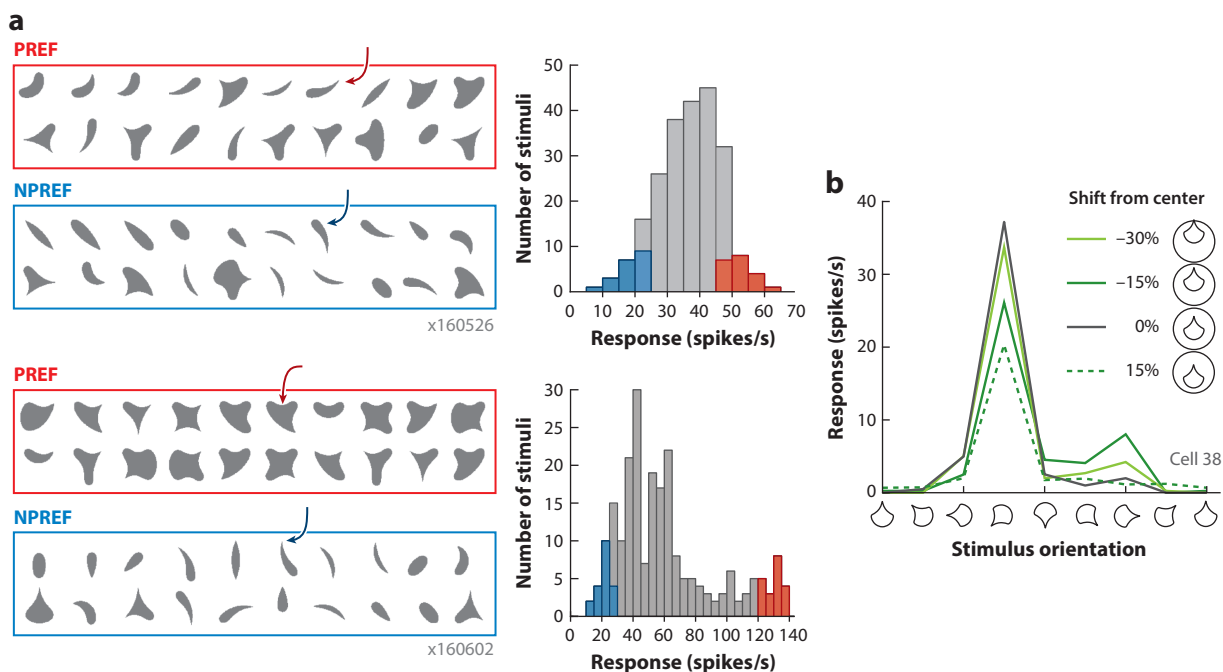


Figure 3

Position-invariant tuning for boundary conformation in V4. (*a*) Examples of tuning for boundary conformation. Responses of two example V4 neurons (top and bottom) to simple 2D shapes recorded in a fixating animal (for further details, see Kim et al. 2019a). Response frequency histograms (*right*) show that the shape stimuli evoked a broad range of responses from both neurons. The 20 shapes that evoked the strongest (PREF, *red*) and weakest (NPREF, *blue*) responses are shown for each neuron (*left*). For the top neuron, all of the preferred shapes included a medium-sharp convex feature to the upper right; shapes that evoked a weak response did not include this feature (compare PREF and NPREF shapes). For the second neuron, preferred but not nonpreferred shapes included a concavity at the top of the shape. Responses of both neurons can be well-explained by a two-dimensional angular position $\times$ boundary curvature model that captures the conformation of the shape boundary (Kim et al. 2019a). Panel adapted from Kim et al. (2019a) (CC BY 4.0). (*b*) Example of position-invariant shape tuning. This neuron shows strong narrow tuning for the orientation of a shape stimulus, responding best when the sharp convex projection is to the lower left. Shape selectivity was consistent irrespective of the absolute position of the shape within the receptive field (compare tuning for the different lines). The response magnitude varied across position, but the preference remained the same.

preference was position invariant. **Figure 3b** shows the results of such a test, where the rotation tuning of a neuron is probed at different positions within the RF. Responses were strongest at a central position and weaker when the stimuli were shifted in either direction from the center. However, the shape preference was maintained across all positions—the shape with the sharp convexity pointing to the lower left evoked the strongest responses. This pattern of responses was typical across V4—many V4 neurons exhibited shape tuning that is independent of spatial position (El-Shamayleh & Pasupathy 2016, Gallant et al. 1993)—thereby disproving the possibility that shape preference is due to RF hotspots.

Another possibility is that the shape preference illustrated in **Figure 3a** is based on a simple preference for feature conjunctions. For example, preference for shapes with a sharp convexity pointing up could be built by pooling signals from V1 or V2 neurons sensitive to 45° and 135° orientations and then repeating such a template across positions to build position-invariant tuning for feature conjunctions (**Figure 4a**). The hierarchical MAX pooling (HMAX) model for

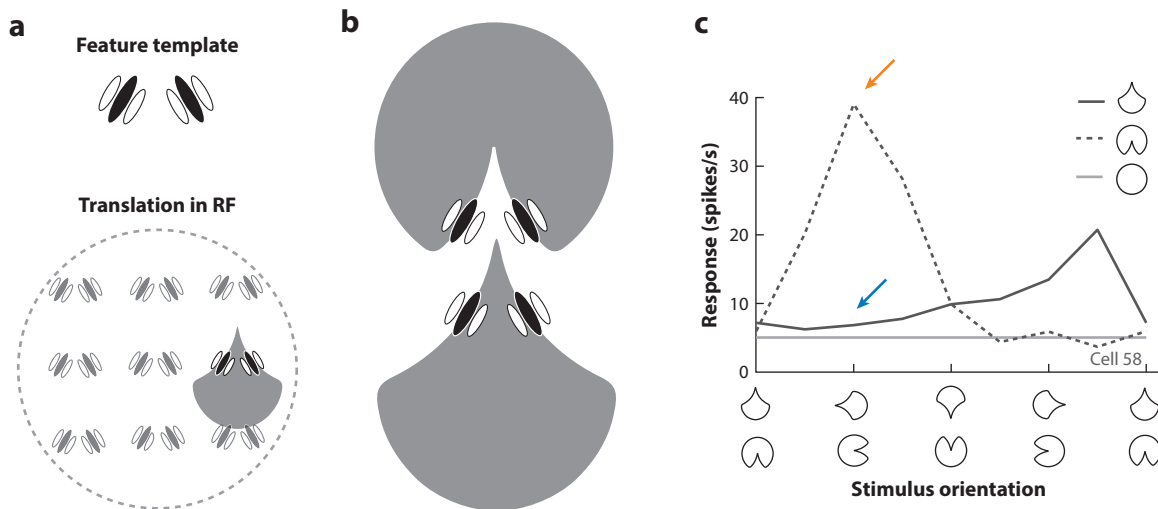


Figure 4

Object-centered tuning for boundary conformation in V4. (a) Preference for a sharp convexity pointing up could be achieved by pooling signals from units tuned to an appropriate combination of orientations (feature template) and repeating the template at multiple positions [translation in the receptive field (RF)]. (b) This model would predict similar responses to the two shapes shown, one with a convex projection at the top of the shape and the other with a concave contour at the bottom of the shape, and would therefore not represent an object-centered code. (c) Example of object-centered tuning for contour conformation. Responses of this neuron show clear rotation tuning for a shape with a concave feature along the boundary. Responses are strongest (orange arrow) when the concavity is to the right of shape center. When the same boundary conformation forms a convex projection to the left of shape center, responses are weak (blue arrow). Thus, responses are not dictated by contour conformation alone.

invariant object recognition proposed by Poggio and colleagues (Cadieu et al. 2007, Riesenhuber & Poggio 1999, Serre et al. 2005) is based on such a strategy to build selectivity and invariance in alternate stages. However, if this were the case, V4 neurons would respond similarly to shapes with a convex projection and those with a concave indentation where the convex and concave parts have matched contour conformations (Figure 4b). This is not what we find. Instead, V4 responses are sensitive both to the contour conformation and to its position relative to object center. For example, in Figure 4c, the neuron responds strongly to stimuli with a concavity to the right of the shape (orange arrow) but not to shapes with a convexity to the left (blue arrow), even though the boundary conformations are identical. This is true even when the local contrast is flipped. Thus, V4 neurons encode information about a feature and its position along the object boundary, i.e., they represent contour features in an object-centered reference frame. Pasupathy & Connor (2001) characterized such object-centered preference for shape in terms of tuning in a 2D shape space for boundary curvature and object-centered angular position. In this scheme, the top neuron in Figure 3a encodes a sharp convexity to the upper right of the shape, and the example in Figure 4c encodes a sharp concavity to the right. Implicit in this model is the idea that the responses of a V4 neuron reflect the shape of the boundary as it relates to the rest of the object—whether a specific contour is a convex projection or a concave indentation requires knowledge about the rest of the object—and where it sits along the object boundary.

To test this idea further, Bushnell et al. (2011b) investigated whether V4 neurons respond strongly to the preferred contour feature when it is formed by the accidental juxtaposition of an occluded and an occluding object boundary. In Figure 5a, the convexity, Θ , is a real contour

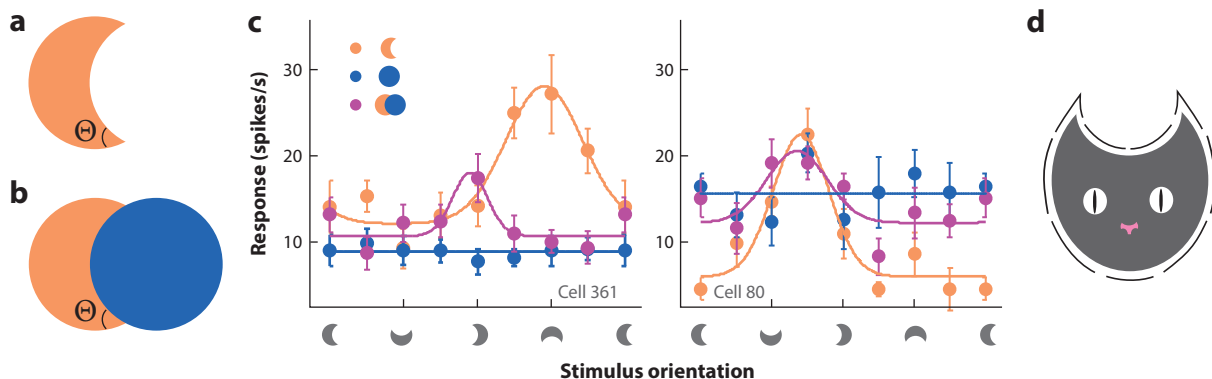


Figure 5

Representation of real but not accidental object boundaries in V4. (*a, b*) Accidental versus real contours. The angle Θ in panel *a* is a real contour, while in panel *b*, it is formed by the accidental occlusion of one circle by another. Accidental contours bound the occluded object, but they carry no information about its true shape, and they are perceptually discounted. (*c*) Responses of two example neurons to crescent shapes in eight orientations (x axis), either in isolation (*orange*) or in combination with a circle (*magenta*); the circle was also presented in isolation (*blue*) at the same eight positions in the receptive field as in the combination stimulus. The neuron on the left responds preferentially to shapes with a sharp convexity at the bottom, as reflected in the tuning curve for the crescent alone (*orange*). When the crescent is formed by partial occlusion in the case of the combination stimulus, responses are suppressed (*magenta*) because the preferred sharp convexity is now an accidental feature. The neuron on the right responds best to shapes with a broad convexity to the upper right, consistent with the strong responses to the circle alone (*blue*). In this case, responses are not suppressed with the addition of a circle because the preferred broad convexity remains a real contour (for further details, see Bushnell et al. 2011b). (*d*) A population of V4 neurons can provide a complete representation of isolated shapes based on the component contour features. For example, the cat may be encoded by V4 neurons selective for sharp convexities to the upper right and upper left; concavity to the top; and broad convexities to the right, left, and bottom (see Pasupathy & Connor 2002). Panels *a–c* adapted from Bushnell et al. (2011b) (CC BY-NC-SA 3.0).

feature that is part of a crescent-shaped object. In **Figure 5b**, however, the same convexity is the result of partial occlusion of an orange circle by a blue circle. Bushnell et al. found that many V4 neurons that respond preferentially to a real contour feature are unresponsive if that feature is the result of partial occlusion, as illustrated in **Figure 5c**. Such suppression cannot be explained on the basis of low-level factors of local chromatic or luminance contrast, number of stimuli, etc. (Bushnell et al. 2011b). These results support the hypothesis that V4 neurons encode only those features in a visual scene that provide information about the true shape of objects. Such suppressed encoding of accidental contours may also form the basis for the emergence of border ownership signals in earlier stages (Pasupathy et al. 2018, Zhou et al. 2000).

Finally, across the V4 population, a greater number of neurons are tuned to sharp convexities and concavities than to shallow convex or concave curves. This is consistent with a sparse object-coding scheme where uncommon but diagnostic regions of acute contour curvature are preferentially encoded (Carlson et al. 2011). Such encoding of object boundaries emerges first in V4 and later in V1, possibly via feedback (Chen et al. 2014).

Taken together, these results support the idea that neurons in area V4 encode parts of an object in terms of the contour features using an object-centered code. In the image in **Figure 5d**, some neurons encode the sharp convexity to the upper right, others encode the concavity to the top, etc. Together, a population of V4 neurons could provide a complete and accurate representation of the entire shape (Pasupathy & Connor 2002) and could thus contribute to its recognition and perception.

JOINT ENCODING OF BOUNDARY AND SURFACE PROPERTIES TO FACILITATE OBJECT SEGMENTATION

Not all V4 neurons are sensitive to object shape. Many neurons are sensitive to surface properties including hue, saturation and luminance contrasts (Bushnell et al. 2011a, Conway et al. 2007, Namima et al. 2014, Schein & Desimone 1990, Zeki 1973), luminance gradients (Hanazawa & Komatsu 2001), and surface texture. Previous studies have documented V4 selectivity for Cartesian and non-Cartesian gratings (Gallant et al. 1993) and in response to brief, sequential presentations of small oriented elements that may be considered dynamic texture (Nandy et al. 2013). Studies have also demonstrated selectivity in V4 for homogeneous naturalistic textures (Arcizet et al. 2008), and such selectivity can be explained on the basis of higher-order correlational statistics (Okazawa et al. 2015). In addition, stimulus dimensions critical for human texture perception, namely, coarseness, directionality, and regularity (Kim et al. 2019b), dictate the texture responses of many V4 neurons. Such V4 sensitivity to surface characteristics is expected, since much of our visual environment is indeed composed of surfaces rather than object boundaries, and processing and accurately perceiving such information are critical for decision making in everyday life (Adelson 2001).

Information about object shape and surface characteristics is often multiplexed in V4 neurons. While some V4 neurons exhibit exclusive tuning for shape or texture, a majority exhibit joint, independent tuning for both stimulus attributes (Kim et al. 2019a). In fact, many neurons that exhibit shape-selective responses for stimuli defined by luminance contrast relative to the background exhibit weak, nonselective responses to shapes defined solely by an outline without an interior fill, and vice versa. These results support the hypothesis that shape selectivity in V4 is dictated both by the boundary and by surface characteristics (Popovkina et al. 2019). Roughly 15% of shape-selective V4 neurons respond best to shapes with blurred boundaries (Oleskiw et al. 2018), and roughly 20% respond selectively to shapes defined solely by a chromatic contrast (Bushnell et al. 2011a). Altogether, these results suggest that individual V4 neurons are tuned in a high-dimensional space that allows the joint encoding of shape and surface characteristics of object parts. In natural vision, object boundaries are defined by contrasts in luminance, texture, and/or chromaticity, and three-dimensional objects that curve away from the plane of fixation may have blurred boundaries. The diverse V4 response properties described above—sensitivity to luminance, texture and chromatic contrasts, blurry boundaries, and luminance gradients—may be well-suited to segment from the background object parts defined by a variety of stimulus cues (Figure 6).

GOAL-ORIENTED REPRESENTATIONS OF OBJECTS

Above, we focus on how simple stimuli—isolated shapes and textures—are encoded in V4 during passive fixation. These experiments (simple stimuli, fixation task) and analysis methods (average activity during stimulus presentation) largely capture feedforward processing in V4. However, V4 is heavily interconnected with other brain areas (see Figure 1), and there are extensive recurrent connections within V4. Experiments that engage these circuits with targeted stimuli and/or behavioral manipulation, and analyses that probe response dynamics, reveal that V4 representations are goal-oriented in several important ways (Figure 7a): (a) They are more closely aligned with the perceived rather than the retinal image; (b) they may guide the choice for the next saccade target; (c) they reflect the targeted processing of stimuli of interest, i.e., those that are the focus of attention; and (d) they may contribute to the sequential comparison of visual

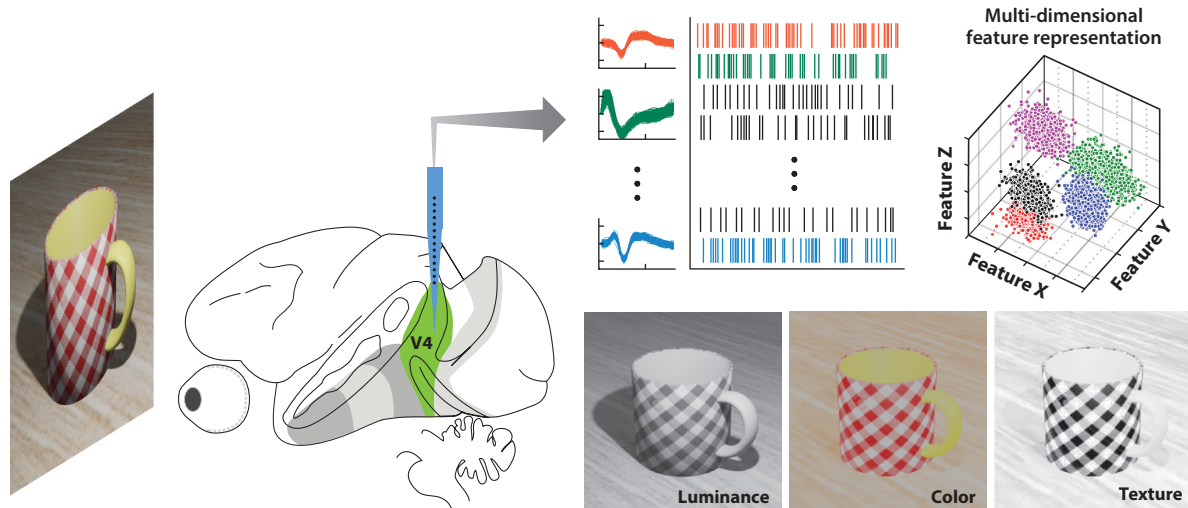


Figure 6

Joint encoding of multiple features for object segmentation. The input image on the left includes a variety of cues; the bottom images are filtered to include information from each cue in isolation and illustrate that form information may be encoded by a contrast in luminance, color, or texture. When viewing such a stimulus, responses of individual V4 neurons are dictated by tuning in a high-dimensional stimulus space defined by shape, luminance, color, texture, blur, depth, etc., which facilitates the effective segmentation of visual objects from the background that may be defined by contrast along a variety of stimulus dimensions.

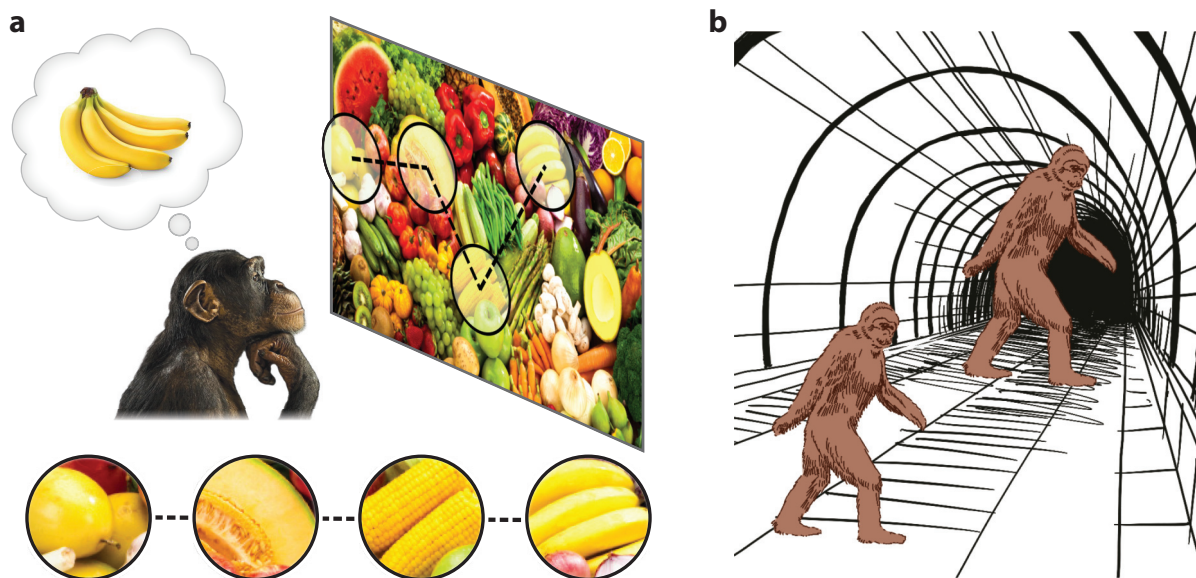


Figure 7

Goal-oriented representations. (a) When confronted with the challenge of spotting bananas in a cluttered produce aisle, the subject may saccade to different locations with yellow objects (*dashed trajectory*) and compare the shape of the object at the attentional focus (*circles*) with a remembered object. Area V4 is thought to be important for all aspects of this process. (b) Size illusion. The retinal sizes of the two sasquatches in this image are identical, but the perceived sizes are dramatically different. This is because the surrounding context suggests that the sasquatch at right is farther away from the observer; thus, the same retinal size would imply a much larger sasquatch farther away.

stimuli, i.e., comparing the memory representation of a stimulus to the visual representation of another. We discuss each of these below.

Building a Percept from the Retinal Input: Solving the Inverse Problem

The image of an object can induce strikingly different percepts depending on the surrounding context. For example, the two sasquatches in **Figure 7b** are identical, yet the one on the right appears to be larger. This simple illusion serves to illuminate the complexity in the computations that transform a retinal image into a percept. The retinal image is a function not just of the properties of the objects and surfaces in the visual world (e.g., shape, size, texture), but also of the viewing conditions (illumination, viewing distance, and angle) and of object–object and object–surface interactions (occlusion). Thus, for example, objects of different sizes at different distances from the observer could cast identical images on the retina. To produce a percept that is informative and actionable from this entangled retinal representation, processing in the successive stages must solve an inverse problem that produces best estimates for the true attributes of the objects and surfaces after accounting for viewing conditions. We do not know how the brain solves this computationally challenging, often ill-defined, inverse problem, but contextual modulations and feedback from higher processing stages are thought to be important (Lamme et al. 1998). Investigations of color constancy and the processing of partial occlusions indicate a role for V4 in this process.

A body of work on human patients, as well as physiological and lesion studies in monkeys, suggests that color constancy—the percept of true color independent of the illuminant—is based on contextual modulations in V4. Several studies have identified patients whose color constancy mechanisms have failed (Clarke et al. 1998, Kennard et al. 1995, Zeki et al. 1999), i.e., their color perception is based on wavelength composition and is not independent of the illuminant. All of these patients suffered damage to regions of the cortex that include human V4 (Walsh 1999). Color constancy is also impaired in monkeys with V4 lesions (Heywood et al. 1992, Schiller 1993, Walsh et al. 1993). In neurophysiological studies, Zeki and colleagues (Kusunoki et al. 2006) demonstrated that V4 color tuning peaks shift in predictable ways when the illumination of a multicolored, Mondrian-like background is changed. Similar shifts in the perceived color were also observed in humans and monkeys when assayed with a match-to-sample task. These results support the hypothesis that regularities in reflected wavelength patterns due to illuminants in the surround influence color tuning functions in V4 and thus contribute to the perception of color constancy.

Geometric regularities in image features caused by occlusions may also be discounted by contextual modulations in V4. V4 neurons tuned to a sharp convexity are more strongly suppressed by adjoining contextual stimuli than are neurons tuned to a broad convexity (**Figure 5c**). This could reflect the geometric prior that curvature discontinuities are a diagnostic feature of partial occlusion but do not carry any information about the true shape of the occluded object. Thus, the greater suppressive influence of adjoining stimuli on the representation of sharp convexities could facilitate the preferential encoding of real object boundaries, rather than accidental contours. Maier and colleagues (Cox et al. 2013) have demonstrated elevated responses in a subpopulation of V4 neurons to illusory Kanizsa surfaces, supporting the hypothesis that active perceptual completion of surfaces and shapes draws on the selective enhancement of activity within V4. This is also consistent with V4 lesion studies demonstrating an impairment in the perception of illusory contours (De Weerd et al. 1996).

Interactions between the visual and frontal cortex may also contribute to solving the inverse problem of deriving a percept from a retinal image. When monkeys discriminate shapes under

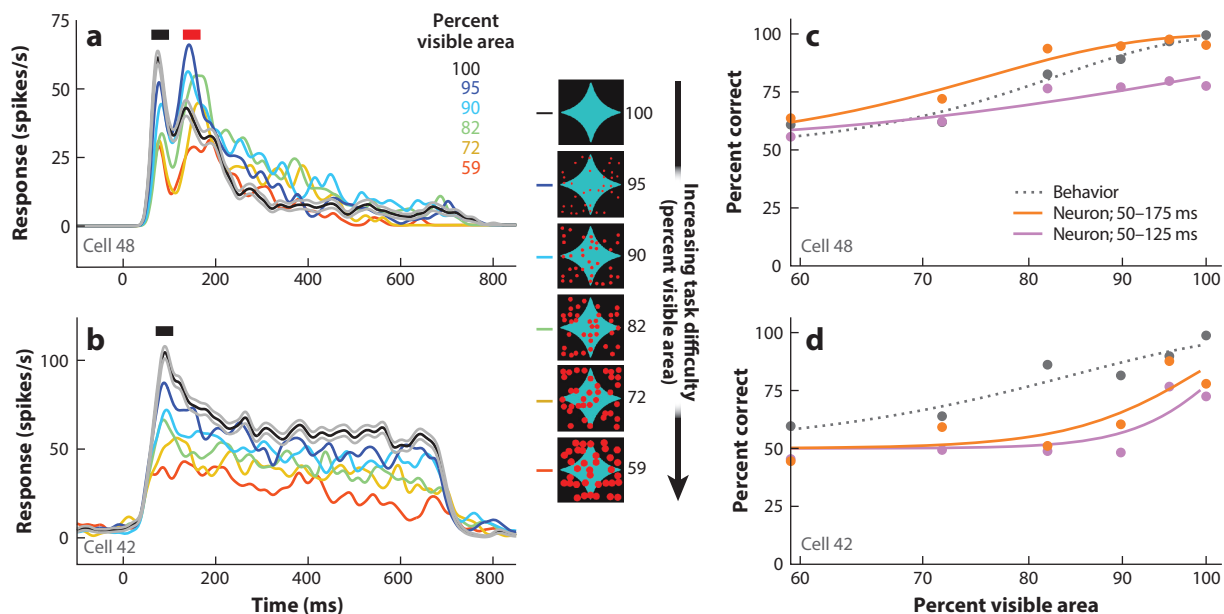


Figure 8

Interactions between the visual and frontal cortex during shape discrimination. (*a, b*) Responses of two example V4 neurons that exhibit different response profiles to partially occluded shape stimuli. PSTHs in panel *a* exhibit two transient peaks (black and red bars), while those in panel *b* exhibit only one transient peak (black bar). During the first transient in both neurons, responses decline with increasing levels of occlusion (line colors; decreasing percent visible area). During the second transient in panel *a*, which may be based on feedback from frontal cortex, responses are stronger for intermediate levels of occlusion (for further details, see Fyall et al. 2017, Kosai et al. 2014). Adapted from Fyall et al. (2017) (CC-BY). (*c, d*) Psychometric and neurometric curves based on the data in panels *a* and *b*, respectively. In both, psychometric performance (dark gray dotted line) declines with increasing levels of occlusion. The neurometric curve based on a larger time window (orange line) shows improved performance at intermediate levels of occlusion for the neuron in panel *a* but not the one in panel *b* due to the enhanced shape selectivity for occluded stimuli during the second transient peak.

partial occlusion, the initial feedforward transient burst in V4 declines with increasing levels of occlusion (**Figure 8a**). Nevertheless, a second transient burst is observed in many neurons that is stronger for occluded stimuli (**Figure 8a**); the net result is that the average response during the first and second transients is less dependent on the occluders and more dependent on the occluded stimulus that needs to be discriminated. This is captured by the neurometric curve, which is much improved when based on activity in the 50–175 ms duration, as opposed to activity in the 50–125 ms duration (**Figure 8c**). A similar improvement is not observed in neurons without a second peak (**Figure 8b,d**). Thus, over time, V4 activity more strongly encodes the stimulus that needs to be discriminated than the occluders, which are irrelevant in this task. The latency of the first and second transient peaks in V4 and timing of peak activity in the ventrolateral prefrontal cortex (vLPFC) support the hypothesis that feedback from the vLPFC may give rise to the second peak in V4 (Fyall et al. 2017). This is also supported by demonstrations of coupling in local field potential (LFP) between V4 and the vLPFC during memory maintenance (Liebe et al. 2012).

Choosing the Next Saccade Target

Several lines of evidence suggest that the V4 representation maintains a saliency map that informs the next saccade. Mazer & Gallant (2003) have demonstrated that, prior to a saccade, a majority of V4 neurons exhibit a presaccadic enhancement that reflects the bottom-up saliency of RF

features, rather than an oculomotor command. A majority of V4 neurons also exhibit a top-down modulation based on target selection. Mazer & Gallant argue that such convergence of bottom-up and top-down processing streams in area V4 results in an adaptive, dynamic map of salience that guides oculomotor planning during natural vision. Burrows & Moore (2009) demonstrated that V4 responses to an oriented, colored bar within the RF were stronger when the surrounding stimuli were designed to produce perceptual popout for color, orientation, or conjunction, but no such enhancement was observed in V1. Schiller & Lee (1991) also demonstrated that V4 lesions impair the ability of monkeys to saccade to the odd-one-out target from an array of distractors, especially if the targets were of low contrast, smaller in size, or slower (see also De Weerd et al. 1999). Thus, V4 responses appear to be critical for identifying stimuli that are different from surrounding stimuli, potentially as targets for the next saccade, even when those stimuli are not the brightest, largest, or fastest.

Attentional Modulations in Area V4

Some studies have shown that behavioral allocation of attention produces robust modulation of the strength of neuronal responses in V4, their synchrony, and their noise correlation (Cohen & Maunsell 2009, Mitchell et al. 2009). V4 attentional effects may be both spatial position-based and feature-based (for a detailed review, see Maunsell 2015). Briefly, studies have demonstrated (a) a shrinking of the V4 RF about the object that is the focus of attention (Moran & Desimone 1985); (b) enhanced responses in a small neighborhood surrounding the focus of attention (spotlight of attention) (Connor et al. 1996); (c) enhanced gamma-band synchronization in V4, which could enhance the postsynaptic impact of V4 neurons driven by the attended stimulus (Fries et al. 2008); and (d) enhanced responses in V4 when a preferred feature in the RF matches the target feature (for a review, see Maunsell & Treue 2006). Feature-based attention could also alter tuning. David and colleagues (2008) demonstrated shifts in orientation and spatial frequency tuning peaks of many V4 neurons toward the direction of the orientation and spatial frequency content of the sought target. Popovkina & Pasupathy (2019) demonstrated that many V4 neurons also exhibit broadening of tuning width for an irrelevant tuning dimension (color) when animals make shape judgements. Such broadening cannot be explained by a gain change alone. Thus, attentional modulations at the level of V4 facilitate the targeted processing of a spatial location, an object, or even a specific stimulus dimension. Feature- or object-based (see Pooresmaeili et al. 2014) attentional mechanisms may be especially important for the maintenance of internal representation of targets and thus for achieving behavioral goals (Hayden & Gallant 2005).

Sequential Comparison of Simple Visual Stimuli

In everyday situations, we often compare items that we are looking at with what we have seen previously, as when we ask, “Have we met before?” or “Have I been here before?”, or when we look for a familiar object in a cluttered environment (see, e.g., **Figure 7a**). Scientists have studied the processes underlying this ability with sequential comparison tasks: Two stimuli are presented in sequence, separated by a delay, and the animal reports whether they are the same or different. To perform this task, animals have to compare the sensory representation of the second stimulus with the memory representation of the first. Because of the involvement of working memory, the prefrontal cortex (PFC) is a plausible site for this comparison (Fuster 1989, Kim & Shadlen 1999, Romo & de Lafuente 2013), but for simple stimuli, V4 may be as well. In a version of the sequential comparison task with simple 2D shape stimuli, Kosai et al. (2014) found that V4 signals carry all task-relevant variables: the sensory representation of the second stimulus, the memory representation of the first stimulus, and the outcome of the computation (**Figure 9**). Crucially,

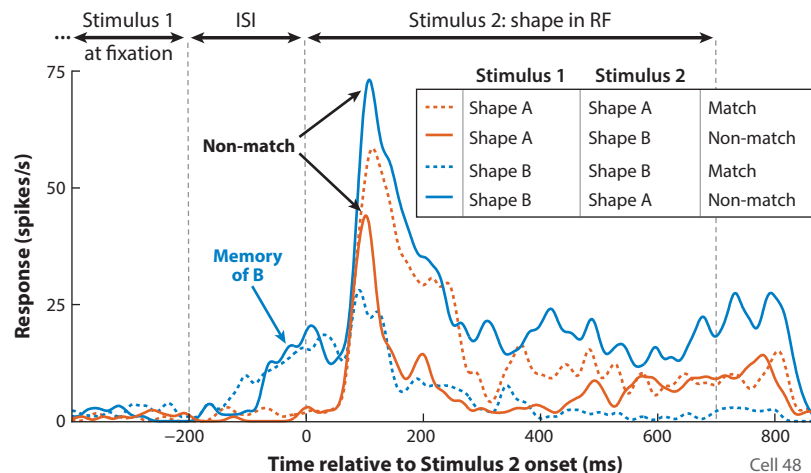


Figure 9

Memory encoding in V4. Responses of an example neuron during the performance of a sequential shape-discrimination task are shown. Stimulus 1 was presented at central fixation (outside the receptive field of the neuron), followed by an inter-stimulus interval (ISI), and stimulus 2 was presented within the receptive field (RF). The animal had to report whether stimuli 1 and 2 were the same or different. Stimulus 1 and stimulus 2 could be shapes A or B, for a total of four conditions (for details, see Kosai et al. 2014). Responses of this neuron include three task-relevant pieces of information. First, the responses provide a sensory representation of stimulus 2 with stronger responses for shape A than shape B (compare the *blue solid line* and *orange dashed line* with the other two lines). Second, a memory representation of stimulus 1 is also evident during the ISI (*blue arrow*), with stronger responses when stimulus 1 was shape B (*blue lines*) rather than shape A (*orange lines*). Finally, responses also reflect whether stimuli 1 and 2 were the same or different with stronger responses during the presentation of stimulus 2 when stimuli 1 and 2 were different (*solid lines, black arrows*).

the memory representation of the first stimulus arises just prior to the onset of the second stimulus, suggesting that it may be communicated from elsewhere in the brain. This idea is consistent with evidence for increased functional connectivity between V4 and the PFC during memory maintenance (Liebe et al. 2012). Furthermore, outcome-related, same-versus-different selectivity emerges soon (approximately 125–200 ms) after second stimulus onset, supporting the hypothesis that V4 may be a site for such a computation (Kosai et al. 2014). Memory- and outcome-related signals have been previously reported in other paradigms (Haenny & Schiller 1988, Hayden & Gallant 2013, Ogawa & Komatsu 2004, Sligte et al. 2009), even when the first stimulus was tactile (Haenny et al. 1988).

FUTURE DIRECTIONS AND OUTSTANDING QUESTIONS

Processing in V4 is likely important for several additional aspects of visual perception that have not been investigated in depth in prior studies. We outline three of these below. We also discuss outstanding questions related to how V4 response properties may be built and how neurons with different response properties may be arranged across V4.

Dynamic Stimuli

Most neurophysiological studies focused on the processing of dynamic stimuli target dorsal stream areas (Manning & Britten 2017). However, 15–30% of V4 neurons exhibit direction bias (Ferrera et al. 1994, Tolia et al. 2005), i.e., they respond more strongly to one motion direction than to the

opposite one, and V4 is strongly interconnected to dorsal stream areas MT and LIP. Bigelow et al. (2019) demonstrated that V4 neurons signal the direction of motion when an object is intermittently displaced across the display with large spatial steps ($dx > 0.5^\circ$), providing the first known neuronal correlate of long-range apparent motion. Thus, V4 may be important for segmentation and tracking of dynamic objects, and future experiments will need to investigate how this might be achieved. In-depth studies are also needed to determine how V4 motion signals are computed and whether form and motion direction signals are multiplexed in the responses of single neurons.

3D Stimuli

V4 is likely critical for building a 3D percept of objects and surfaces, since V4 lesions in monkeys appear to impair this ability (Merigan & Pham 1998); however, this has not been extensively explored. Many V4 neurons are disparity tuned (Hinkle & Connor 2001), and many encode the slant of a bar in depth (Hinkle & Connor 2002). V4 neurons also encode nondisparity cues that are important for building a 3D percept. For example, when an object curves away from the plane of focus, the boundary is blurry, and many shape-selective V4 neurons respond preferentially to such blurry boundaries (Oleskiw et al. 2018). V4 neurons are also selective for surface luminance gradients (Hanazawa & Komatsu 2001), which could contribute to the recovery of 3D information from shading. Consistent with this observation, Arcizet and colleagues (2009) demonstrated that a population of V4 neurons can differentiate between 3D shapes defined by surface shading and corresponding 2D control stimuli. Future studies will need to delineate how 3D information from a variety of cues, e.g., disparity, shading, blur, and occlusions, may be integrated to build a percept that is invariant to illumination direction.

Visual Crowding

Because most previous studies have focused on the encoding of isolated stimuli, we do not know whether V4 encoding is lossy when there are multiple nearby objects. Studies in human subjects have explored the limits of object recognition in crowded displays (Balas et al. 2009, Levi 2008, Pelli & Tillman 2008). For accurate recognition in humans, the minimum spacing between objects scales with eccentricity [Bouma's law (Bouma 1973): approximately $0.3\text{--}0.6 \times \text{RF eccentricity}$]. Because monkey V2 and V4 RF diameters also scale with eccentricity at approximately 0.3 and $0.6 \times \text{RF eccentricity}$ (Gattass et al. 1981, 1988), respectively, the limits of recognition in clutter may be a direct consequence of the representation in the mid-level visual cortex. We do know from prior studies that surround modulation can be strongly suppressive in V4 (Schein & Desimone 1990) and that responses to pairs of stimuli may be modeled as the maximum of the responses to the component stimuli in isolation (Gawne & Martin 2002). More extensive and systematic studies are needed to document how the number, distance, and properties of the neighboring stimuli modulate V4 responses and how this could underlie perceptual limitations due to crowding. Such experiments would be critical for us to begin to understand how natural scenes may be encoded in V4.

Functional Organization

The distinct patterns of input and output connections of V4 subregions (DeYoe et al. 1994, Felleman et al. 1997, Zeki & Shipp 1989) and the similar preferences of nearby neurons recorded on a single electrode penetration (Gallant et al. 1996, Watanabe et al. 2002) hint that an underlying functional organization exists. Recordings with laminar probes also suggest segregation of

units encoding sensory information and eye-movement planning in the superficial and deep layers (Pettine et al. 2019). Experiments employing functional neuroimaging methods (e.g., fMRI and optical imaging) have demonstrated that V4 contains clear functional domains in terms of preference for color, orientation, spatial frequency, size, motion direction, and disparity (Conway et al. 2007, Fang et al. 2019, Ghose & Ts'O 1997, Li et al. 2013, Lu et al. 2018, Tanigawa et al. 2010). To investigate how multiple feature dimensions may be arranged across V4 (i.e., microarchitecture of functional domains), advanced methods that facilitate high-density sampling of single-unit activity are required. Recent studies using two-photon calcium imaging have revealed neuronal clusters associated with curvature or corners (Jiang et al. 2019) and 3D shape encoding (Nielsen 2019). Future studies with high-density electrode recording and imaging techniques could explain how the neurons that encode various stimulus dimensions or participate in attentional modulation and behavioral influence are laid out across the cortex.

Models

A major outstanding question is how V4 responses are built from responses in V1 and V2. Several models have been proposed, typically to explain responses to one specific class of stimuli, but they often fail to generalize to other classes. For example, the HMAX model for object recognition can provide a good fit for tuning for boundary curvature by pooling from a set of V1 neurons tuned to appropriate orientations (see **Figure 4a**). However, such a contour template model does not capture the object-centered nature of shape coding in V4 and fails to achieve the level of position invariance observed in real neurons (Bair et al. 2015). It also fails to capture the diversity in fill-outline invariance observed in V4 responses (Popovkina et al. 2019). In contrast, the texture statistics model of Okazawa and colleagues (2015) can capture tuning for texture, but it fails to generate the texture-invariant shape tuning observed in V4 (Kim et al. 2019a). Nandy and colleagues (2013) have proposed that V4 tuning for curved contours could be produced by pooling of heterogeneous orientation signals from earlier visual areas. However, such a model cannot produce translation invariance in shape tuning, a fundamental property of many V4 neurons (El-Shamayleh & Pasupathy 2016, Gallant et al. 1993, Pasupathy & Connor 2001). The spectral RF model of Gallant and colleagues (David et al. 2006) can achieve translation invariance but not tuning for boundary curvature (Oleskiw et al. 2014).

In recent years, with the advent of efficient learning algorithms for deep learning neural networks, computer vision has made great strides with regards to object recognition: On some tasks, networks have reached levels that are comparable to humans. Encouraged by this, some studies have begun to compare the internal representations of these models and the responses of visual cortical neurons (e.g., Pospisil et al. 2018, Yamins et al. 2014). While deep networks are not models of the brain (in terms of architecture, cell types, and functionality), analyzing the network can be insightful because the emergence of similar encoding features in models and neurons could imply similar computational strategies. Dissecting the underlying architecture of model units could provide insights into how response properties arise in the brain. Furthermore, detailed study of model units could promote the development of more targeted experiments with neurons; this could be invaluable due to the practical constraints that limit experimental time.

Efforts to date are already beginning to bear fruit. Comparison of responses of single units in deep networks and monkey V4 neurons has identified the current best model for position-invariant tuning for boundary curvature in V4 (Pospisil et al. 2018). Deep networks, in a closed loop with primate physiology experiments, have been used to validate models by testing response predictions to novel synthetic stimuli (see, e.g., Bashivan et al. 2019). Future experiments that compare deep networks and the brain in terms of unit responses and behavior, especially leveraging

their differences, could provide deeper edification about the why and the how of mid-level visual cortical representations and functions.

CONCLUSION

Much remains to be discovered about how mid-level visual cortex represents stimuli and how these representations underlie visual perception and object recognition and guide our behavior. The time is right to make progress—more labs, armed with more experimental and analytical tools, are undertaking this quest. To move forward toward a comprehensive understanding of the functional role of area V4 in vision, we need to conduct more experiments with a variety of parametric, artificial stimuli that serve to constrain, falsify, and update models; these models could then be validated by evaluating their responses to natural stimuli. The biggest obstacle to making progress in this manner is the constraint of experimental time. Thus, the biggest technological advance would be the ability to study the same neuron over days, weeks, and months. With this innovation, we could evaluate and characterize the responses of a single neuron to tens of thousands of stimuli; we could also study the responses of neurons under a variety of behavioral conditions and contexts. These data could then constrain more elaborate and physiologically sound models of form processing that include feedforward, recurrence, and feedback circuits.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

The authors would like to thank Yasmine El-Shamayleh, Yoshito Kosai, and Polina Zamarashkina for data collection. This work was supported by National Eye Institute grants R01 EY018839 and EY029997 to A.P. and the National Institutes of Health/Office of Research Infrastructure Programs grant P51 OD010425 to the Washington National Primate Research Center.

LITERATURE CITED

- Adelson EH. 2001. On seeing stuff: the perception of materials by humans and machines. In *Proceedings of the SPIE Volume 4299: Human Vision and Electronic Imaging VI*, ed. BE Rogowitz, TN Pappas, pp. 1–12. Bellingham, WA: SPIE
- Adelson EH, Bergen JR. 1991. The plenoptic function and the elements of early vision. In *Computational Models of Visual Processing*, ed. M Landy, JA Movshon, pp. 3–20. Cambridge, MA: MIT Press
- Albrecht DG, de Valois RL, Thorell LG. 1980. Visual cortical neurons: Are bars or gratings the optimal stimuli? *Science* 207(4426):88–90
- Arcizet F, Jouffrais C, Girard P. 2008. Natural textures classification in area V4 of the macaque monkey. *Exp. Brain Res.* 189(1):109–20
- Arcizet F, Jouffrais C, Girard P. 2009. Coding of shape from shading in area V4 of the macaque monkey. *BMC Neurosci.* 10(1):140
- Bair W, Popovkina DV, De A, Pasupathy A. 2015. *Modeling shape representation in area V4*. Paper presented at MODVIS Workshop, St. Pete Beach, FL, May 13–15
- Bakin JS, Nakayama K, Gilbert CD. 2000. Visual responses in monkey areas V1 and V2 to three-dimensional surface configurations. *J. Neurosci.* 20(21):8188–98
- Balaban H, Drew T, Luria R. 2019. Neural evidence for an object-based pointer system underlying working memory. *Cortex* 119:362–72

- Balas B, Nakano L, Rosenholtz R. 2009. A summary-statistic representation in peripheral vision explains visual crowding. *J. Vis.* 9(12):13
- Barbas H, Mesulam M-M. 1985. Cortical afferent input to the principals region of the rhesus monkey. *Neuroscience* 15(3):619–37
- Bashivan P, Kar K, DiCarlo JJ. 2019. Neural population control via deep image synthesis. *Science* 364(6439):eaav9436
- Bauer R, Heinze S. 2002. Contour integration in striate cortex. *Exp. Brain Res.* 147(2):145–52
- Bigelow AW, Kim T, Bair W, Pasupathy A. 2019. *Long-range apparent motion tuning in ventral visual area V4*. Paper presented at Society for Neuroscience Meeting, Chicago, Oct. 19–23
- Blakemore C, Tobin EA. 1972. Lateral inhibition between orientation detectors in the cat's visual cortex. *Exp. Brain Res.* 15(4):439–40
- Bouma H. 1973. Visual interference in the parafoveal recognition of initial and final letters of words. *Vis. Res.* 13(4):767–82
- Burrows BE, Moore T. 2009. Influence and limitations of popout in the selection of salient visual stimuli by area V4 neurons. *J. Neurosci.* 29(48):15169–77
- Bushnell BN, Harding PJ, Kosai Y, Bair W, Pasupathy A. 2011a. Equiluminance cells in visual cortical area v4. *J. Neurosci.* 31(35):12398–412
- Bushnell BN, Harding PJ, Kosai Y, Pasupathy A. 2011b. Partial occlusion modulates contour-based shape encoding in primate area V4. *J. Neurosci.* 31(11):4012–24
- Cadiou C, Kouh M, Pasupathy A, Connor CE, Riesenhuber M, Poggio T. 2007. A model of V4 shape selectivity and invariance. *J. Neurophysiol.* 98(3):1733–50
- Carlson ET, Rasquinha RJ, Zhang K, Connor CE. 2011. A sparse object coding scheme in area V4. *Curr. Biol.* 21(4):288–93
- Cavanaugh JR, Bair W, Movshon JA. 2002. Selectivity and spatial distribution of signals from the receptive field surround in macaque V1 neurons. *J. Neurophysiol.* 88(5):2547–56
- Chen M, Yan Y, Gong X, Gilbert CD, Liang H, Li W. 2014. Incremental integration of global contours through interplay between visual cortical areas. *Neuron* 82(3):682–94
- Clarke S, Walsh V, Schoppig A, Assal G, Cowey A. 1998. Colour constancy impairments in patients with lesions of the prestriate cortex. *Exp. Brain Res.* 123(1–2):154–58
- Coen-Cagli R, Kohn A, Schwartz O. 2015. Flexible gating of contextual influences in natural vision. *Nat. Neurosci.* 18(11):1648–55
- Cohen MR, Maunsell JHR. 2009. Attention improves performance primarily by reducing interneuronal correlations. *Nat. Neurosci.* 12(12):1594–600
- Connor CE, Gallant JL, Preddie DC, Van Essen DC. 1996. Responses in area V4 depend on the spatial relationship between stimulus and attention. *J. Neurophysiol.* 75(3):1306–8
- Conway BR, Moeller S, Tsao DY. 2007. Specialized color modules in macaque extrastriate cortex. *Neuron* 56(3):560–73
- Cooke T, Jäkel F, Wallraven C, Bühlhoff HH. 2007. Multimodal similarity and categorization of novel, three-dimensional objects. *Neuropsychologia* 45(3):484–95
- Cox MA, Schmid MC, Peters AJ, Saunders RC, Leopold DA, Maier A. 2013. Receptive field focus of visual area V4 neurons determines responses to illusory surfaces. *PNAS* 110(42):17095–100
- David SV, Hayden BY, Gallant JL. 2006. Spectral receptive field properties explain shape selectivity in area V4. *J. Neurophysiol.* 96(6):3492–505
- David SV, Hayden BY, Mazer JA, Gallant JL. 2008. Attention to stimulus features shifts spectral tuning of V4 neurons during natural vision. *Neuron* 59(3):509–21
- De Weerd P, Desimone R, Ungerleider LG. 1996. Cue-dependent deficits in grating orientation discrimination after V4 lesions in macaques. *Vis. Neurosci.* 13(3):529–38
- De Weerd P, Peralta MR, Desimone R, Ungerleider LG. 1999. Loss of attentional stimulus selection after extrastriate cortical lesions in macaques. *Nat. Neurosci.* 2(8):753–58
- DeYoe EA, Felleman DJ, Van Essen DC, McClendon E. 1994. Multiple processing streams in occipitotemporal visual cortex. *Nature* 371(6493):151–54
- El-Shamayleh Y, Pasupathy A. 2016. Contour curvature as an invariant code for objects in visual area V4. *J. Neurosci.* 36(20):5532–43

- Elder JH, Velisavljevic L. 2009. Cue dynamics underlying rapid detection of animals in natural scenes. *J. Vis.* 9(7):7
- Fang Y, Chen M, Xu H, Li P, Han C, et al. 2019. An orientation map for disparity-defined edges in area V4. *Cereb. Cortex* 29(2):666–79
- Felleman DJ, Van Essen DC. 1991. Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* 1(1):1–47
- Felleman DJ, Xiao Y, McClendon E. 1997. Modular organization of occipito-temporal pathways: cortical connections between visual area 4 and visual area 2 and posterior inferotemporal ventral area in macaque monkeys. *J. Neurosci.* 17(9):3185–200
- Ferrera V, Rudolph KK, Maunsell JH. 1994. Responses of neurons in the parietal and temporal visual pathways during a motion task. *J. Neurosci.* 14(10):6171–86
- Freeman J, Simoncelli EP. 2011. Metamers of the ventral stream. *Nat. Neurosci.* 14(9):1195–201
- Freeman J, Ziemba CM, Heeger DJ, Simoncelli EP, Movshon JA. 2013. A functional and perceptual signature of the second visual area in primates. *Nat. Neurosci.* 16(7):974–81
- Fries P, Womelsdorf T, Oostenveld R, Desimone R. 2008. The effects of visual stimulation and selective visual attention on rhythmic neuronal synchronization in macaque area V4. *J. Neurosci.* 28(18):4823–35
- Fuster JM. 1989. *The Prefrontal Cortex*. New York: Raven Press. 2nd ed.
- Fyall AM, El-Shamayleh Y, Choi H, Shea-Brown E, Pasupathy A. 2017. Dynamic representation of partially occluded objects in primate prefrontal and visual cortex. *eLife* 6:e25784
- Gallant JL, Braun J, Van Essen DC. 1993. Selectivity for polar, hyperbolic, and Cartesian gratings in macaque visual cortex. *Science* 259(5091):100–3
- Gallant JL, Connor CE, Rakshit S, Lewis JW, Van Essen DC. 1996. Neural responses to polar, hyperbolic, and Cartesian gratings in area V4 of the macaque monkey. *J. Neurophysiol.* 76(4):2718–39
- Gattass R, Galkin TW, Desimone R, Ungerleider LG. 2014. Subcortical connections of area V4 in the macaque. *J. Comp. Neurol.* 522(8):1941–65
- Gattass R, Gross CG, Sandell JH. 1981. Visual topography of V2 in the macaque. *J. Comp. Neurol.* 201(4):519–39
- Gattass R, Sousa A, Gross CG. 1988. Visuotopic organization and extent of V3 and V4 of the macaque. *J. Neurosci.* 8(6):1831–45
- Gawne TJ, Martin JM. 2002. Responses of primate visual cortical V4 neurons to simultaneously presented stimuli. *J. Neurophysiol.* 88(3):1128–35
- Gheorghiu E, Kingdom FAA, Petkov N. 2014. Contextual modulation as de-texturizer. *Vis. Res.* 104:12–23
- Ghose GM, Ts’O DY. 1997. Form processing modules in primate area V4. *J. Neurophysiol.* 77(4):2191–96
- Haenny PE, Maunsell JH, Schiller PH. 1988. State dependent activity in monkey visual cortex. II. Retinal and extraretinal factors in V4. *Exp. Brain Res.* 69(2):245–59
- Haenny PE, Schiller PH. 1988. State dependent activity in monkey visual cortex. I. Single cell activity in V1 and V4 on visual tasks. *Exp. Brain Res.* 69(2):225–44
- Hanazawa A, Komatsu H. 2001. Influence of the direction of elemental luminance gradients on the responses of V4 cells to textured surfaces. *J. Neurosci.* 21(12):4490–97
- Hayden BY, Gallant JL. 2005. Time course of attention reveals different mechanisms for spatial and feature-based attention in area V4. *Neuron* 47(5):637–43
- Hayden BY, Gallant JL. 2013. Working memory and decision processes in visual area V4. *Front. Neurosci.* 7:18
- Heywood CA, Gadotti A, Cowey A. 1992. Cortical area V4 and its role in the perception of color. *J. Neurosci.* 12(10):4056–65
- Hinkle DA, Connor CE. 2001. Disparity tuning in macaque area V4. *Neuroreport* 12(2):365–69
- Hinkle DA, Connor CE. 2002. Three-dimensional orientation tuning in macaque area V4. *Nat. Neurosci.* 5(7):665–70
- Hubel DH, Wiesel TN. 1959. Receptive fields of single neurones in the cat’s striate cortex. *J. Physiol.* 148(3):574–91
- Hubel DH, Wiesel TN. 1968. Receptive fields and functional architecture of monkey striate cortex. *J. Physiol.* 195:215–43
- Hupé J-M, James AC, Girard P, Bullier J. 2001. Response modulations by static texture surround in area V1 of the macaque monkey do not depend on feedback connections from V2. *J. Neurophysiol.* 85(1):146–63

- Jiang R, Li M, Tang S. 2019. Discrete neural clusters encode orientation, curvature and corners in macaque V4. *bioRxiv* 808907. <https://doi.org/10.1101/808907>
- Kahneman D, Treisman A, Gibbs BJ. 1992. The reviewing of object files: object-specific integration of information. *Cogn. Psychol.* 24(2):175–219
- Kapadia MK, Ito M, Gilbert CD, Westheimer G. 1995. Improvement in visual sensitivity by changes in local context: parallel studies in human observers and in V1 of alert monkeys. *Neuron* 15(4):843–56
- Kennard C, Lawden M, Morland AB, Ruddock KH. 1995. Colour identification and colour constancy are impaired in a patient with incomplete achromatopsia associated with prestriate cortical lesions. *Proc. R. Soc. Lond. B* 260(1358):169–75
- Kim J-N, Shadlen MN. 1999. Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nat. Neurosci.* 2(2):176–85
- Kim T, Bair W, Pasupathy A. 2019a. Neural coding for shape and texture in macaque area V4. *J. Neurosci.* 39(24):4760–74
- Kim T, Bair W, Pasupathy A. 2019b. *Response dynamics in primate V4 are modulated by perceptual dimensions of visual textures*. Paper presented at Society for Neuroscience Meeting, Chicago, Oct. 19–23
- Knierim JJ, van Essen DC. 1992. Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *J. Neurophysiol.* 67(4):961–80
- Kosai Y, El-Shamayleh Y, Fyall AM, Pasupathy A. 2014. The role of visual area V4 in the discrimination of partially occluded shapes. *J. Neurosci.* 34(25):8570–84
- Kusunoki M, Moutoussis K, Zeki S. 2006. Effect of background colors on the tuning of color-selective cells in monkey area V4. *J. Neurophysiol.* 95(5):3047–59
- Lamme VA, Supér H, Spekreijse H. 1998. Feedforward, horizontal, and feedback processing in the visual cortex. *Curr. Opin. Neurobiol.* 8(4):529–35
- Levi DM. 2008. Crowding—an essential bottleneck for object recognition: a mini-review. *Vis. Res.* 48(5):635–54
- Levitt JB, Lund JS. 1997. Contrast dependence of contextual effects in primate visual cortex. *Nature* 387(6628):73–76
- Li P, Zhu S, Chen M, Han C, Xu H, et al. 2013. A motion direction preference map in monkey V4. *Neuron* 78(2):376–88
- Liebe S, Hoerzer GM, Logothetis NK, Rainer G. 2012. Theta coupling between V4 and prefrontal cortex predicts visual short-term memory performance. *Nat. Neurosci.* 15(3):456–62
- Lu Y, Yin J, Chen Z, Gong H, Liu Y, et al. 2018. Revealing detail along the visual hierarchy: neural clustering preserves acuity from V1 to V4. *Neuron* 98(2):417–28.e3
- Luria R, Vogel EK. 2011. Shape and color conjunction stimuli are represented as bound objects in visual working memory. *Neuropsychologia* 49(6):1632–39
- Manning TS, Britten KH. 2017. Motion processing in primates. In *Oxford Research Encyclopedia: Neuroscience*. Oxford, UK: Oxford Res. Encycl.
- Markov YA, Tiurina NA, Utochkin IS. 2019. Different features are stored independently in visual working memory but mediated by object-based representations. *Acta Psychol.* 197:52–63
- Martinez A, Ramanathan DS, Foxe JJ, Javitt DC, Hillyard SA. 2007. The role of spatial attention in the selection of real and illusory objects. *J. Neurosci.* 27(30):7963–73
- Maunsell JHR. 2015. Neuronal mechanisms of visual attention. *Annu. Rev. Vis. Sci.* 1:373–91
- Maunsell JHR, Treue S. 2006. Feature-based attention in visual cortex. *Trends Neurosci.* 29(6):317–22
- Mazer JA, Gallant JL. 2003. Goal-related activity in V4 during free viewing visual search: evidence for a ventral stream visual salience map. *Neuron* 40(6):1241–50
- Merigan WH, Pham HA. 1998. V4 lesions in macaques affect both single- and multiple-viewpoint shape discriminations. *Vis. Neurosci.* 15:359–67
- Mitchell JF, Sundberg KA, Reynolds JH. 2009. Spatial attention decorrelates intrinsic activity fluctuations in macaque area V4. *Neuron* 63(6):879–88
- Moran J, Desimone R. 1985. Selective attention gates visual processing in the extrastriate cortex. *Science* 229:782–84

- Movshon JA, Simoncelli EP. 2014. Representation of naturalistic image structure in the primate visual cortex. *Cold Spring Harb. Symp. Quant. Biol.* 79:115–22
- Movshon JA, Thompson ID, Tolhurst DJ. 1978a. Spatial summation in the receptive fields of simple cells in the cat's striate cortex. *J. Physiol.* 283(1):53–77
- Movshon JA, Thompson ID, Tolhurst DJ. 1978b. Spatial and temporal contrast sensitivity of neurones in areas 17 and 18 of the cat's visual cortex. *J. Physiol.* 283(1):101–20
- Nakamura H, Gattass R, Desimone R, Ungerleider L. 1993. The modular organization of projections from areas V1 and V2 to areas V4 and TEO in macaques. *J. Neurosci.* 13(9):3681–91
- Namima T, Yasuda M, Banno T, Okazawa G, Komatsu H. 2014. Effects of luminance contrast on the color selectivity of neurons in the macaque area v4 and inferior temporal cortex. *J. Neurosci.* 34(45):14934–47
- Nandy AS, Sharpee TO, Reynolds JH, Mitchell JF. 2013. The fine structure of shape tuning in area V4. *Neuron* 78(6):1102–15
- Nelson JI, Frost BJ. 1985. Intracortical facilitation among co-oriented, co-axially aligned simple cells in cat striate cortex. *Exp. Brain Res.* 61(1):54–61
- Neri P. 2017. Object segmentation controls image reconstruction from natural scenes. *PLOS Biol.* 15(8):e1002611
- Nielsen K. 2019. *Clustering of 3D and 2D shape processing in area V4*. Paper presented at Society for Neuroscience Meeting, Chicago, Oct. 19–23
- Ninomiya T, Sawamura H, Inoue K, Takada M. 2012a. Multisynaptic inputs from the medial temporal lobe to V4 in macaques. *PLOS ONE* 7(12):e52115
- Ninomiya T, Sawamura H, Inoue K, Takada M. 2012b. Segregated pathways carrying frontally derived top-down signals to visual areas MT and V4 in macaques. *J. Neurosci.* 32(20):6851–58
- Nothdurft H-C, Gallant JL, Van Essen DC. 1999. Response modulation by texture surround in primate area V1: correlates of “popout” under anesthesia. *Vis. Neurosci.* 16(1):15–34
- Nuthmann A, Henderson JM. 2010. Object-based attentional selection in scene viewing. *J. Vis.* 10(8):20
- Ogawa T, Komatsu H. 2004. Target selection in area V4 during a multidimensional visual search task. *J. Neurosci.* 24(28):6371–82
- Okazawa G, Tajima S, Komatsu H. 2015. Image statistics underlying natural texture selectivity of neurons in macaque V4. *PNAS* 112(4):E351–60
- Okazawa G, Tajima S, Komatsu H. 2017. Gradual development of visual texture-selective properties between macaque areas V2 and V4. *Cereb. Cortex* 27(10):4867–80
- Oleskiw TD, Nowack A, Pasupathy A. 2018. Joint coding of shape and blur in area V4. *Nat. Commun.* 9(1):466
- Oleskiw TD, Pasupathy A, Bair W. 2014. Spectral receptive fields do not explain tuning for boundary curvature in V4. *J. Neurophysiol.* 112(9):2114–22
- Parker AJ. 2007. Binocular depth perception and the cerebral cortex. *Nat. Rev. Neurosci.* 8(5):379–91
- Pasupathy A, Connor CE. 1999. Responses to contour features in macaque area V4. *J. Neurophysiol.* 82(5):2490–502
- Pasupathy A, Connor CE. 2001. Shape representation in area V4: position-specific tuning for boundary conformation. *J. Neurophysiol.* 86(5):2505–19
- Pasupathy A, Connor CE. 2002. Population coding of shape in area V4. *Nat. Neurosci.* 5(12):1332–38
- Pasupathy A, El-Shamayleh Y, Popovkina D. 2018. Visual shape and object perception. In *Oxford Research Encyclopedia: Neuroscience*, ed. SM Sherman. Oxford, UK: Oxford Res. Encycl.
- Pelli DG, Tillman KA. 2008. The uncrowded window of object recognition. *Nat. Neurosci.* 11(10):1129–35
- Pettine WW, Steinmetz NA, Moore T. 2019. Laminar segregation of sensory coding and behavioral readout in macaque V4. *PNAS* 116(29):14749–54
- Polat U, Mizobe K, Pettet MW, Kasamatsu T, Norcia AM. 1998. Collinear stimuli regulate visual responses depending on cell's contrast threshold. *Nature* 391(6667):580–84
- Pooresmaeili A, Poort J, Roelfsema PR. 2014. Simultaneous selection by object-based attention in visual and frontal cortex. *PNAS* 111(17):6467–72
- Poort J, Raudies F, Wannig A, Lamme VAF, Neumann H, Roelfsema PR. 2012. The role of attention in figure-ground segregation in areas V1 and V4 of the visual cortex. *Neuron* 75(1):143–56

- Popovkina D, Bair W, Pasupathy A. 2019. Modelling diverse responses to filled and outline shapes in macaque V4. *J. Neurophysiol.* 121(3):1059–77
- Popovkina DV, Pasupathy A. 2019. Task context modulates feature-selective responses in area V4. bioRxiv 594150. <https://doi.org/10.1101/594150>
- Portilla J, Simoncelli EP. 2000. A parametric texture model based on joint statistics of complex wavelet coefficients. *Int. J. Comput. Vis.* 40(1):49–70
- Pospisil DA, Pasupathy A, Bair W. 2018. “Artphysiology” reveals V4-like shape tuning in a deep network trained for image classification. *eLife* 7:e38242
- Riesenhuber M, Poggio T. 1999. Hierarchical models of object recognition in cortex. *Nat. Neurosci.* 2(11):1019–25
- Romo R, de Lafuente V. 2013. Conversion of sensory signals into perceptual decisions. *Prog. Neurobiol.* 103:41–75
- Rosenholtz R, Huang J, Raj A, Balas BJ, Ilie L. 2012. A summary statistic representation in peripheral vision explains visual search. *J. Vis.* 12(4):14
- Schaffelhofer S, Scherberger H. 2016. Object vision to hand action in macaque parietal, premotor, and motor cortices. *eLife* 5:e15278
- Schein SJ, Desimone R. 1990. Spectral properties of V4 neurons in the macaque. *J. Neurosci.* 10(10):3369–89
- Schiller PH. 1993. The effects of V4 and middle temporal (MT) area lesions on visual performance in the rhesus monkey. *Vis. Neurosci.* 10(4):717–46
- Schiller PH, Lee K. 1991. The role of the primate extrastriate area V4 in vision. *Science* 251(4998):1251–53
- Schut MJ, Fabius JH, Van der Stoep N, Van der Stigchel S. 2017. Object files across eye movements: previous fixations affect the latencies of corrective saccades. *Atten. Percept. Psychophys.* 79(1):138–53
- Serre T, Wolf L, Poggio T. 2005. Object recognition with features inspired by visual cortex. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 994–1000. Piscataway, NJ: IEEE
- Sligte IG, Scholte HS, Lamme VAF. 2009. V4 activity predicts the strength of visual short-term memory representations. *J. Neurosci.* 29(23):7432–38
- Tanigawa H, Lu HD, Roe AW. 2010. Functional organization for color and orientation in macaque V4. *Nat. Neurosci.* 13(12):1542–48
- Tolias AS, Keliris GA, Smirnakis SM, Logothetis NK. 2005. Reply to “Motion processing in macaque V4.” *Nat. Neurosci.* 8(9):1125
- Treisman A, Kahneman D, Burkell J. 1983. Perceptual objects and the cost of filtering. *Percept. Psychophys.* 33(6):527–32
- Ungerleider LG, Galkin TW, Desimone R, Gattass R. 2008. Cortical connections of area V4 in the macaque. *Cereb. Cortex* 18(3):477–99
- Van Gool L, Dewaele P, Oosterlinck A. 1985. Texture analysis Anno 1983. *Comput. Vis. Graph. Image Process.* 29(3):336–57
- Victor JD, Conte MM. 2012. Local image statistics: maximum-entropy constructions and perceptual salience. *J. Opt. Soc. Am. A* 29(7):1313–45
- Wallis TS, Funke CM, Ecker AS, Gatys LA, Wichmann FA, Bethge M. 2019. Image content is more important than Bouma’s Law for scene metamers. *eLife* 8:e42512
- Walsh V. 1999. How does the cortex construct color? *PNAS* 96(24):13594–96
- Walsh V, Carden D, Butler SR, Kulikowski JJ. 1993. The effects of V4 lesions on the visual abilities of macaques: hue discrimination and colour constancy. *Behav. Brain Res.* 53(1–2):51–62
- Watanabe M, Tanaka H, Uka T, Fujita I. 2002. Disparity-selective neurons in area V4 of macaque monkeys. *J. Neurophysiol.* 87(4):1960–73
- Yamins DLK, Hong H, Cadieu CF, Solomon EA, Seibert D, DiCarlo JJ. 2014. Performance-optimized hierarchical models predict neural responses in higher visual cortex. *PNAS* 111(23):8619–24
- Yu Y, Schmid AM, Victor JD. 2015. Visual processing of informative multipoint correlations arises primarily in V2. *eLife* 4:e06604
- Zamarashkina P, Popovkina DV, Pasupathy A. 2020. Stimulus and task dependence of response latencies in primate area V4. *J. Neurophysiol.* 17(10):476

- Zeki S, Aglioti S, McKeefry D, Berlucchi G. 1999. The neurological basis of conscious color perception in a blind patient. *PNAS* 96(24):14124–29
- Zeki S, Shipp S. 1989. Modular connections between areas V2 and V4 of macaque monkey visual cortex. *Eur. J. Neurosci.* 1(5):494–506
- Zeki SM. 1973. Colour coding in rhesus monkey prestriate cortex. *Brain Res.* 53:422–27
- Zhou H, Friedman HS, von der Heydt R. 2000. Coding of border ownership in monkey visual cortex. *J. Neurosci.* 20(17):6594–611
- Ziomba CM, Freeman J, Movshon JA, Simoncelli EP. 2016. Selectivity and tolerance for visual texture in macaque V2. *PNAS* 113(22):E3140–49



Contents

Fifty Years Exploring the Visual System <i>Joel Pokorny and Vivianne C. Smith</i>	1
Genetic and Environmental Risk Factors for Keratoconus <i>Sionne E.M. Lucas and Kathryn P. Burdon</i>	25
Minimally Invasive Glaucoma Surgery: A Critical Appraisal of the Literature <i>David J. Mathew and Yvonne M. Buys</i>	47
Human Organoids for the Study of Retinal Development and Disease <i>Claire M. Bell, Donald J. Zack, and Cynthia A. Berlinicke</i>	91
Cellular-Scale Imaging of Transparent Retinal Structures and Processes Using Adaptive Optics Optical Coherence Tomography <i>Donald T. Miller and Kazuhiro Kurokawa</i>	115
Microglia Activation and Inflammation During the Death of Mammalian Photoreceptors <i>Sarah J. Karlen, Eric B. Miller, and Marie E. Burns</i>	149
Reprogramming Müller Glia to Regenerate Retinal Neurons <i>Manuela Labne, Mikiko Nagashima, David R. Hyde, and Peter F. Hitchcock</i>	171
Axon Regeneration in the Mammalian Optic Nerve <i>Philip R. Williams, Larry I. Benowitz, Jeffrey L. Goldberg, and Zbigang He</i>	195
Retinal Ganglion Cell Axon Wiring Establishing the Binocular Circuit <i>Carol Mason and Nefeli Slavi</i>	215
Topographic Variations in Retinal Encoding of Visual Space <i>Alina Sophie Heukamp, Rebekah Anne Warwick, and Michal Rivlin-Etzion</i>	237
Organization, Function, and Development of the Mouse Retinogeniculate Synapse <i>Liang Liang and Chinfai Chen</i>	261
Signals Related to Color in the Early Visual Cortex <i>Gregory D. Horwitz</i>	287
Role of Feedback Connections in Central Visual Processing <i>Farran Briggs</i>	313

Linking Neuronal Direction Selectivity to Perceptual Decisions About Visual Motion <i>Tatiana Pasternak and Duje Tadin</i>	335
Visual Functions of Primate Area V4 <i>Anitha Pasupathy, Dina V. Popovkina, and Taekjun Kim</i>	363
Coding Perceptual Decisions: From Single Units to Emergent Signaling Properties in Cortical Circuits <i>Kristine Krug</i>	387
Anatomy and Function of the Primate Entorhinal Cortex <i>Aaron D. Garcia and Elizabeth A. Buffalo</i>	411
Tuning the Senses: How the Pupil Shapes Vision at the Earliest Stage <i>Sebastiaan Mathôt</i>	433
Can We See with Melanopsin? <i>Robert J. Lucas, Annette E. Allen, Nina Milosavljevic, Riccardo Storchi, and Tom Woelders</i>	453
Retinal Image Formation and Sampling in a Three-Dimensional World <i>Larry N. Thibos</i>	469
Image-Computable Ideal Observers for Tasks with Natural Stimuli <i>Johannes Burge</i>	491
Measuring and Modeling Visual Appearance <i>Laurence T. Maloney and Kenneth Knoblauch</i>	519
Visual Search: How Do We Find What We Are Looking For? <i>Jeremy M. Wolfe</i>	539
Rethinking Space: A Review of Perception, Attention, and Memory in Scene Processing <i>Monica S. Castelbano and Karolina Krzys</i>	563

Errata

An online log of corrections to *Annual Review of Vision Science* articles may be found at
<http://www.annualreviews.org/errata/vision>