

Responses to Orientation Discontinuities in V1 and V2: Physiological Dissociations and Functional Implications

Anita M. Schmid, Keith P. Purpura, and Jonathan D. Victor

Brain and Mind Research Institute, Division of Systems Neurology and Neuroscience, Weill Cornell Medical College, New York, New York 10065

Segmenting the visual image into objects is a crucial stage of visual processing. Object boundaries are typically associated with differences in luminance, but discontinuities in texture also play an important role. We showed previously that a subpopulation of neurons in V2 in anesthetized macaques responds to orientation discontinuities parallel to their receptive field orientation. Such single-cell responses could be a neurophysiological correlate of texture boundary detection. Neurons in V1, on the other hand, are known to have contextual response modulations such as iso-orientation surround suppression, which also produce responses to orientation discontinuities. Here, we use pseudorandom multiregion grating stimuli of two frame durations (20 and 40 ms) to probe and compare texture boundary responses in V1 and V2 in anesthetized macaque monkeys. In V1, responses to texture boundaries were observed for only the 40 ms frame duration and were independent of the orientation of the texture boundary. However, in transient V2 neurons, responses to such texture boundaries were robust for both frame durations and were stronger for boundaries parallel to the neuron's preferred orientation. The dependence of these processes on stimulus duration and orientation indicates that responses to texture boundaries in V2 arise independently of contextual modulations in V1. In addition, because the responses in transient V2 neurons are sensitive to the orientation of the texture boundary but those of V1 neurons are not, we suggest that V2 responses are the correlate of texture boundary detection, whereas contextual modulation in V1 serves other purposes, possibly related to orientation “pop-out.”

Key words: contextual modulation; cue invariance; pop-out; spatial nonlinearities; surround suppression; texture segmentation

Introduction

To analyze a scene, the visual system needs to find the boundaries of objects. Therefore, image segmentation is crucial to vision. Object boundaries are typically associated with differences in luminance, but differences of texture or color can also carry information about boundaries. Nonluminance cues may play an important role in visual processing because identification of boundaries based solely on luminance differences is likely to be confounded by edges of shadows (Derrington et al., 2002; Kingdom, 2003; Kingdom et al., 2004). Color cues facilitate image segmentation (Gegenfurtner and Rieger, 2000) by eliminating such false positives. Texture cues may play a similarly important role in reducing the chance of wrongly inferring the presence of object boundaries in complex visual scenes.

It is unclear which parts of the visual system are involved in texture segmentation. The underlying neural mechanisms are also unknown. Some studies suggest that texture segmentation begins in primary visual cortex (V1; Grosfof et al., 1993; Sillito et

al., 1995; Nothdurft et al., 2000; Schmid, 2008). Specifically, iso-orientation surround suppression is observed in many neurons in V1. These neurons' receptive fields have a suppressive surround that is tuned to the same orientation as the center. Therefore, suppression is reduced when the orientation in the surround differs from the preferred orientation, in particular when it is perpendicular to it (Blakemore and Tobin, 1972; Fries et al., 1977; Nelson and Frost, 1978). Neurons with iso-orientation surround suppression respond especially well when the preferred orientation covers the center and a contrasting orientation covers part of the surround (Schmid, 2008). Such a response property in V1 could be one mechanism by which the visual system detects texture boundaries.

Some neurons in secondary visual cortex (V2) also have properties that may play an important role in texture segmentation (Leventhal et al., 1998; Mareschal and Baker, 1998a, 1998b; Marcicar et al., 2000; von der Heydt et al., 2000; Zhan and Baker, 2006; Song and Baker, 2007; Schmid et al., 2009). We have shown recently that responses to orientation discontinuities are present in one of two subpopulations of V2, the “transient” V2 neurons (Schmid et al., 2009). We also showed that whereas responses of transient V2 neurons to luminance boundaries can be understood as a feedforward combination of V1 responses, responses to texture boundaries cannot; they must be generated by spatial interactions within V2.

In this study, we concentrate on the role of these response properties in V1 and V2 in texture segmentation. Specifically, the question addressed here is how iso-orientation surround suppression in

Received May 28, 2013; revised Jan. 23, 2014; accepted Jan. 28, 2014.

Author contributions: A.M.S. and J.D.V. designed research; A.M.S. and J.D.V. performed research; A.M.S. analyzed data; A.M.S., K.P., and J.D.V. wrote the paper.

This work was supported by the National Institutes of Health (Grant EY9314 to J.D.V.). We thank Qin Hu, Ferenc Mechler, Eyal Nitzany, Ifije E. Ohiorhenuan, and Yunguo Yu for help in performing the experiments.

The authors declare no competing financial interests.

Correspondence should be addressed to Anita M. Schmid, Brain and Mind Research Institute, Division of Systems Neurology and Neuroscience, Weill Cornell Medical College, 1300 York Avenue, New York, NY 10065. E-mail: ams2031@med.cornell.edu.

DOI:10.1523/JNEUROSCI.2293-13.2014

Copyright © 2014 the authors 0270-6474/14/343559-20\$15.00/0

V1 and responses to orientation discontinuities in V2 relate to each other and how they might contribute to texture segmentation. To make a direct comparison, we modified the stimuli used in our previous study (Schmid et al., 2009) by extending each stimulus frame to 40 ms. This allowed signs of iso-orientation surround suppression to emerge in V1. Measuring response dynamics in this temporal range allowed us to observe fundamental differences between texture boundary responses in V1 and V2.

Materials and Methods

The Materials and Methods used in this study were similar to those described previously (Schmid et al., 2009) and are summarized here.

Physiological preparation. Standard acute preparation techniques were used for electrophysiological recordings from single units in the V1 and V2 of cynomolgus monkeys (*Macaca fascicularis*). All procedures were in accordance with institutional and National Institutes of Health guidelines for the care and experimental use of animals and under an approved protocol from the Weill Cornell Medical College Institutional Animal Care and Use Committee.

Experiments were performed on 12 adult animals weighing 2.2–10 kg (11 males and 1 female). The preparation was similar to what was described previously (Mechler et al., 2002; Victor et al., 2006; Schmid et al., 2009) and is summarized here. After an overnight fast, animals were premedicated with atropine (0.05 mg/kg, i.m.; Henry Schein) and then anesthetized with ketamine (Ketaset, 10 mg/kg, i.m.; Fort Dodge Animal Health) or Telazol (4 mg/kg, i.m.; Fort Dodge Animal Health) and xylazine (Rompun, 0.5 mg/kg, i.m.; Bayer). Under anesthesia with isoflurane (1–2%; Hospira) during the surgery, an endotracheal tube was placed and catheters put in both femoral veins, one femoral artery, and the urethra. During recording, anesthesia was maintained with propofol (PropoFlo, 2–20 mg/kg/h, i.v.; Abbott) and sufentanil (Sufenta, 1–10 μ g/kg/h, i.v.; Janssen) and neuromuscular blockade was induced (after all surgical procedures) and maintained with vecuronium bromide (0.25 mg/kg, i.v. bolus, 0.25 mg/kg/h, i.v.; Bedford Laboratories) or rocuronium bromide (1.5 mg/kg, i.v. bolus, 1.5 mg/kg/h, i.v.; Mylan Institutional). Heart rate and rhythm, arterial blood pressure, body temperature, end-expiratory pCO₂, urine output, and EEG were monitored during the course of the experiment. Animal maintenance included intravenous fluids (lactated Ringer's solution or Normosol-R with 5% glucose, 2–4 cm³/kg/h), administration of supplemental O₂ every 6 h, antibiotics (procaine penicillin G, 75,000 U/kg, i.m. prophylactically; King Pharmaceuticals; gentamicin, 5 mg/kg i.m. daily if evidence of infection; Abbott; dexamethasone, 1 mg/kg, i.m. daily; AmTech Teva Animal Health), application of 0.5% bupivacaine (Marcaine; Hospira) to wounds, and ocular instillation of atropine (1%; Bausch & Lomb), and flurbiprofen (Ocufer, 0.03%; Allergan), and periodic cleaning of the contact lenses. With these measures, the preparation remained physiologically stable for 4–5 d.

Recording. After a craniotomy near P10, L15, the lunate sulcus was located and a small durotomy performed over V1, V2, or both. Extracellular recordings were made with 3 or 6 tetrodes (quartz-coated platinum-tungsten fibers; Thomas Recording). The penetration angle varied and we used different geometrical layouts for the tetrode array to accommodate different geometries of the brain and the blood vessels. The analog signal from each tetrode channel was amplified, filtered (0.3–6 kHz), and digitized (25 or 30.303 kHz). Once spiking activity from one or more units was encountered, the region of the receptive field(s) was hand mapped and then centered on the display of a ViewSonic G225f 21-inch monitor (displaying a 1280 × 1024 raster at 100 Hz, mean luminance 47 cd/m², gamma corrected) or a Sun GDM5410 21-inch monitor (displaying a 1280 × 1024 raster at 100 Hz, mean luminance 46 cd/m², gamma corrected) at a distance of 114 cm. Control signals for the CRT display were provided by a PC-hosted system optimized for OpenGL (NVIDIA GeForce3 chipset) programmed in Delphi. Multiple single units were isolated by cluster analysis of spike waveforms via Klusters (Hazan et al., 2006). We used 17 features for the cluster analysis: the peak values of the spikes on the four channels, the trough values on the four channels, the first eight scores of the principal component analysis over all spikes, and the spike

time. After first performing automatic clustering (via KlustaKwik), we merged clusters and reclustered repeatedly (in Klusters) until we were confident that the spikes in isolated clusters came from an isolated single neuron based on waveform shape and its gradual change over time. Spikes that were not confident in assigning to a single neuron were discarded.

Histology. After the completion of the recording, the tetrodes were moved back toward the cortical surface and at three locations bracketing the recording sites lesions were made by passing constant current from the tetrode leads to a grounding screw in the skull. Initially, we used a constant current of typically 12 μ A for 6 s, electrode negative. Because the lesions were not always reliable using this method, we tested using pulsed current similar to what other researchers have successfully used with similar tetrodes (Adams and Horton, 2006). We achieved satisfactory results with 20 μ A for 10 s at 10 Hz (square-wave pulse, electrode negative). After a waiting period of 1 h, the animal was deeply anesthetized with propofol (PropoFlo, 5–10 ml, 100 mg/ml, i.v.; Abbott) and pentobarbital (8–32 mg/kg, i.v.; Virbac) and perfused (4% paraformaldehyde; EMS) in PBS. A block of brain tissue containing the tetrode tracks was then removed. In some of the experiments, we did frozen sectioning; in others, we used a vibratome without freezing. For the frozen sections, we let the brain tissue sink in 10%, 20%, and 30% sucrose solution in 4% paraformaldehyde. The sections were then cut parallel to the tetrode tracks, mounted and stained for Nissl in thionin staining solution (1%; Sigma-Aldrich) or cresyl violet staining solutions (0.2%; Sigma-Aldrich). The border between V1 and V2 was identified (the border is readily visible because of the distinct appearance of layer 4 in V1, which disappears in V2). Only recording locations that we could unambiguously place within V1 or V2 were used for further analysis. Recording sites that were within ~250 μ m from the V1-V2 border were excluded because the tetrodes we used can pick up signals in a range of ~130 μ m (Mechler et al., 2011) and there is an additional uncertainty in localizing the center of the lesions and the reconstruction of recording sites between lesion.

Visual stimulation. The pupils were covered with gas-permeable contact lenses (Metro Optics). Artificial pupils (2 mm) and corrective lenses were used to focus the stimulus on the retina. Foveae and the receptive fields of multineuron activity were mapped on a tangent board for each tetrode. Optical correction was established initially by use of an ophthalmoscope and adjusted to maximize the responses of isolated single units to high- spatial-frequency visual stimuli.

Beginning with the parameters determined by qualitative characterization and hand mapping, computer-controlled stimulation paradigms were used to characterize multineuron activity quantitatively with sine-wave gratings. Orientation tuning was determined by the mean response (F0) and the fundamental modulated response (F1) to drifting gratings at orientations spaced in steps of 11.25 degrees presented at a contrast $c = (L_{max} - L_{min}) / (L_{max} + L_{min})$ of 1.0, with spatial and temporal frequency determined by the initial assessment. Next, spatial frequency tuning was determined by responses to drifting gratings at a 16-fold range of spatial frequencies at the orientation determined by the orientation tuning run and a temporal frequency determined by the initial assessment. Temporal tuning was then assessed by responses to 1, 2, 4, 8, and 16 Hz drifting gratings at the optimal orientation and spatial frequency. Finally, a contrast response function was determined by responses to drifting gratings at contrasts of 0, 0.0625, 0.125, 0.25, 0.5, and 1.0, with orientation, spatial frequency, and temporal frequency determined by the previous quantitative runs. The position of the receptive field (RF) was first determined by auditory assessment of spiking activity using a laser pointer on the monitor screen displaying concentric rings around the currently selected center. After adjusting the center accordingly, the size of the classical RF (CRF) was determined from responses to a drifting grating (all parameters optimized) presented in discs of increasing diameter. Centering the RF was checked by recording the responses to a series of annuli that had a fixed outer radius at the size of the CRF and decreasing inner radii. If the responses did not peak for the annulus with zero inner radius, the stimuli were recentered and the outer and inner diameter runs repeated until the centering was satisfactory. The length and width of the CRF were determined by recording responses to the optimal drifting grating presented in a rectangular window and varying the length and width.

Tuning properties. After the experiment and offline cluster cutting of the recorded spikes, the tuning properties were assessed again, analogous to the assessment for the online multiunit activity. If there was more than one tuning run with close to optimal parameters, we chose the one with the largest χ^2 deviation from random. We chose either the F0 or F1 response, depending on which component had the largest χ^2 deviation from random for the particular tuning run. The preferred orientation, spatial frequency, and temporal frequency were defined as the parameters that elicited the maximal responses.

For quantifying surround suppression, we used the difference-of-Gaussians model (DeAngelis et al., 1994; Sceniak et al., 1999) and fit the size-tuning curves by modeling the excitatory and suppressive sensitivity profiles as 2 d Gaussians as follows:

$$S(x, y) = k_c e^{-((2x)^2 + (2y)^2)/a_c^2} - k_s e^{-((2x)^2 + (2y)^2)/a_s^2} \quad (1)$$

where a_c and a_s are the space constants of the classic and suppressive components and k_c and k_s are their peak strengths.

For length and width suppression measurements, the stimuli were narrow strips, so we fitted the response to a stimulus strip of length x_0 to 1D integrals of the sensitivity profile in Equation 1 as follows (adapted from DeAngelis et al., 1994):

$$\begin{aligned} R(x_0) &= \left[R_0 + \int_{-x_0/2}^{x_0/2} S(x, 0) dx \right] \\ &= \left[R_0 + k_c \int_{-x_0/2}^{x_0/2} e^{-(2x)^2/a_c^2} dx - k_s \int_{-x_0/2}^{x_0/2} e^{-(2x)^2/a_s^2} dx \right] \end{aligned} \quad (2)$$

where R_0 is the spontaneous firing rate and $\lfloor R \rfloor$ denotes R or 0, whichever is larger. This can be rewritten using error functions (erf) as follows:

$$R(x_0) = \left[R_0 + \frac{\sqrt{\pi}}{2} a_c k_c \operatorname{erf}\left(\frac{x_0}{a_c}\right) - \frac{\sqrt{\pi}}{2} a_s k_s \operatorname{erf}\left(\frac{x_0}{a_s}\right) \right] \quad (3)$$

We fit all datasets both to a model with suppression (k_s allowed to vary) and one without (forcing $k_s = 0$) by minimizing the sum of the squared deviations between the modeled response (Equation 3) and the original data. To avoid overfitting the model with suppression (often manifest as a high peak of the modeled response in between measured values), we added a penalty equal to the square of the difference between the predicted maximum and the measured maximum.

Because the model with suppression has two more parameters than the model without suppression, we calculated the Akaike Information Criterion (Akaike, 1974) to determine whether inclusion of these parameters was statistically justified. The AIC is given by:

$$AIC = \chi^2 + 2P \quad (4)$$

where P is the number of parameters ($P = 3$ for the model without suppression, $P = 5$ for the model with suppression). χ^2 is given by:

$$\chi^2 = \sum_{i=1}^n \frac{(R_i^{obs} - R_i^{model})^2}{\sigma^2} \quad (5)$$

where R_i^{obs} is the observed response to the i th stimulus and R_i^{model} is the modeled response (from Equation 3), and σ^2 is the root-mean-squared deviation of the measured responses across replicate runs.

If the model with suppression included had the lower AIC, the receptive field length or width x_{RF} was defined as the length or width that yielded the maximum response of the model function. The suppression strength was then defined as the amount of attenuation observed at the largest length or width measured x_{max} as a fraction of the peak response amplitude (DeAngelis et al., 1994) as the suppression index:

$$SI = \frac{R(x_{RF}) - R(x_{max})}{R(x_{RF})} \quad (6)$$

If the model without suppression had the lower AIC, we concluded that inclusion of the suppression term was not statistically justified. In this

case, the receptive field length or width x_{RF} was defined as the size for which the model response reached 95% of the maximum response and the suppression strength was set to 0.

Stimulation with orientation-discontinuity stimuli. A 4×5 or 6×6 grid of adjacent rectangular regions was positioned to cover the classical and part of the nonclassical RF. Each rectangular region was filled with a sinusoidal grating of either the preferred orientation or the orthogonal nonpreferred orientation controlled by an m-sequence. The stimulus was positioned and sized such that a central subset of regions covered the aggregate CRF of the multineuron activity. For the 4×5 grid, we targeted the CRF with a 2×1 block of regions (Fig. 1A) or a 2×3 block; for the 6×6 grid, we targeted the CRF with a 2×2 block. The spatial frequency was chosen in the upper portion of the passband of the multineuron activity so that each region typically contained 1–2 cycles. The stimulus was shown at contrasts ranging between 50% and 100% (median 90%). The orientation in each region was assigned by a binary m-sequence of order 12 (length 4095) changing every 20 ms or of order 11 (length 2047) changing every 40 ms, so that the total length of the experiment was about the same for the two stimulus frame durations, typically a total of 45 min each. The two experiments with the different frame durations were run in succession. Usually, the 40 ms frame duration experiment was run first followed by the 20 ms experiment. The same m-sequence was used for all regions, but with different starting positions (“taps”) in the sequence for each.

As has been described previously (Golomb, 1981; Sutter, 1992; Benardete and Victor, 1994), the main advantage of an m-sequence, that it has combinatorial properties that guarantee its near-perfect autocorrelation (and thus, facilitate extraction of first-order kernels from the cross-correlation), has a drawback: second- and higher-order kernels can contribute to overlapping bins of the cross-correlation. To separate first- and second-order kernels from each other and from higher-order kernels in the face of this “overlap problem,” we used the same approach introduced in our previous study (Schmid et al., 2009). This consisted of three independent strategies. First, we chose the m-sequence taps to reduce the overlaps of the second-order kernels with each other and with the first-order kernels. To achieve this design, we considered a 200 ms response window and analyzed overlaps from the following sources of kernels: first-order responses (within each region and frame), temporal interactions between adjacent or next-to-adjacent frames (within each region), spatial interactions between adjacent regions (within each frame), and pairwise spatiotemporal interactions between adjacent regions and adjacent frames. With the 4×5 grid, the taps we used yielded 4 overlaps for the stimulus with 20 ms frames and 5 overlaps for the stimulus with 40 ms frames; this is of a possible 11,628 overlaps (between 153 pairs of kernels; $11,628 = 153 \times 152/2$). With the 6×6 grid, the most commonly used taps yielded 60 overlaps for the 20 ms stimulus and 84 overlaps for the 40 ms stimulus; this is of a possible 41,328 overlaps (between 288 pairs of kernels; $41,328 = 288 \times 287/2$). Second, we ran inverse-repeat sequences (i.e., we alternated each m-sequence with a sequence in which the assignment of -1 's and $+1$'s to grating patches was inverted); adding and subtracting the estimates from these sequences yields a separation of odd- and even-order kernels (Sutter, 1992) even if they occupy overlapping bins of the cross-correlation. Finally, to exclude responses due to overlaps of third- and higher-order kernels and the overlaps among second-order kernels that the tap sequence selection did not eliminate, we ran each m-sequence with two different assignments of taps to the regions. This allows for identification of responses that are due to overlaps: overlap responses will occur in the different parts of the m-sequence when the taps are changed, whereas the sought-after second-order responses will not. As described in the Statistics section below, putative second-order responses that were significantly different between the two m-sequence taps were excluded from analysis.

We usually ran 32 repeats of the stimulus (eight repeats of two inverted runs at two different taps). The phases at which the oriented sinusoidal gratings were displayed within each region were chosen pseudorandomly on each frame; the same spatial phase was used for all regions that had the same orientation so that gratings in regions of the same orientation were seamlessly aligned (Fig. 1A). The pseudorandom sequence used for assigning 1 of 4 phases ($0, \pi/2, \pi, 3\pi/2$) was a combination of two binary

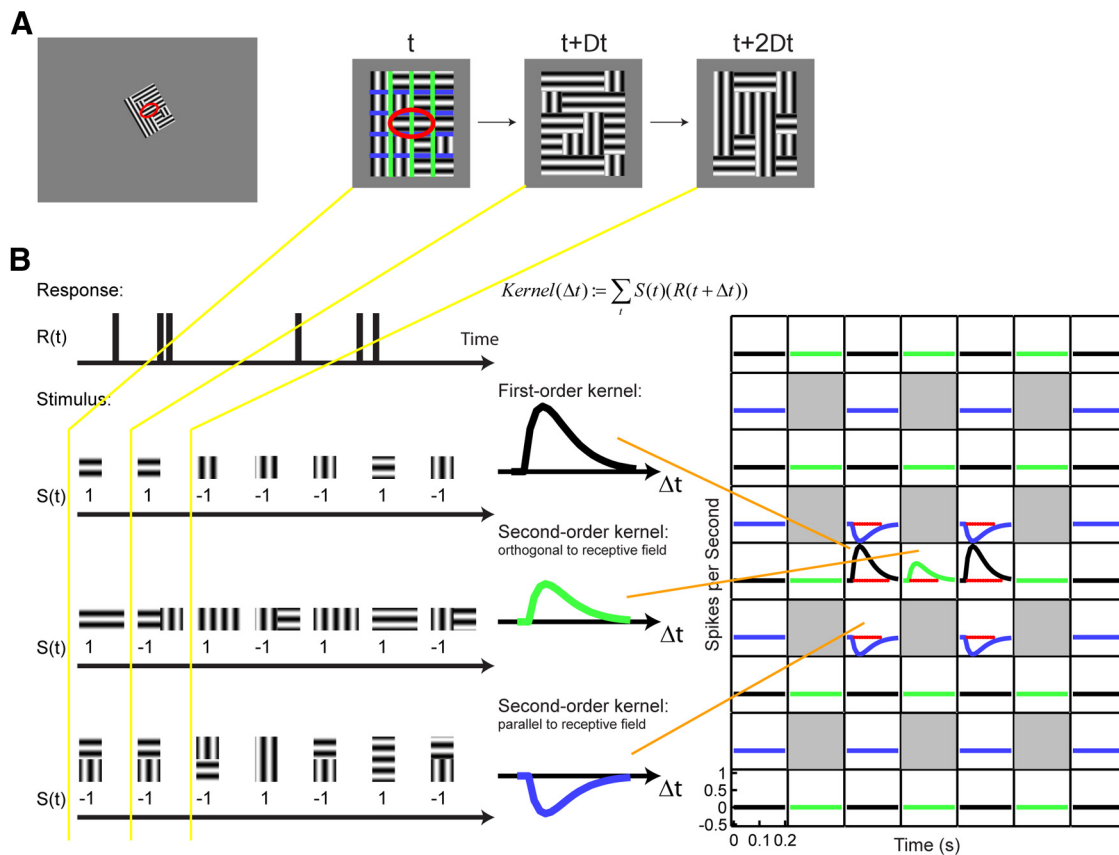


Figure 1. Orientation-discontinuity stimulus and kernel computation. **A**, Stimulus setup: a 4×5 (or 6×6 , not shown here) grid of rectangular regions covered the classical (red ellipse) and nonclassical receptive field. The stimulus was aligned with the preferred orientation of the receptive field. Each region contained a static sinusoidal grating with either the preferred or the orthogonal, nonpreferred orientation. The orientation in each region was randomly reassigned every 20 or 40 ms. Blue and green lines show the region boundaries parallel (blue) and orthogonal (green) to the receptive field; these lines were not part of the stimulus. **B**, Computation of first- and second-order kernels. For each region in the stimulus, the neuron's spike response, in 10 ms bins, was cross-correlated with the stimulus sequence, coded as +1 for the preferred orientation and -1 for the orthogonal orientation. Cross-correlation of this sequence against the neural response yielded the first-order kernel in that region. For the computation of each spatial second-order kernel, the response was correlated with the product of the values of the stimulus in the two neighboring regions: +1 if the grating orientation in the two regions was equal and -1 if they were different. Therefore, positive kernel values indicate an enhanced response to texture continuity and negative values indicate an enhanced response to texture discontinuity. All first-order kernels are then visualized in a grid corresponding to the stimulus layout (black traces). All second-order kernels are visualized between the first-order kernels of the corresponding two stimulus regions that interact to produce them (green traces for second-order kernels orthogonal to the receptive field; blue traces for second-order kernels parallel to the receptive field). Time points for which a kernel is significantly different from zero (see Materials and Methods for details) are indicated by red dots on the zero line.

m-sequences of order 15 and 16 and was thus independent of the m-sequences that determined the orientations themselves.

First-order kernels were computed by reverse correlating the spike response with the m-sequence used for assigning the orientation and then averaging over all spatial phases. For this calculation, the orientation parallel to the receptive field axis—the neuron's preferred orientation—was represented by a +1 token and the orthogonal orientation—the neuron's nonpreferred orientation—was represented by a -1 token (Fig. 1B). We implemented this by calculating a single reverse correlation between the entire stimulus cycle and the response. The first-order kernel for an individual region was then located within the reverse correlation function at a lag corresponding to the tap used for that region. The correlation was scaled so that its amplitude indicated the contribution of a 10 ms segment of the stimulus to the firing rate. Positive values thus signify greater firing following presentation of the neuron's preferred orientation and negative values signify greater firing after presentation of the neuron's nonpreferred orientation.

To calculate spatial second-order kernels, we computed the lags at which they are located within the single reverse correlation between the entire stimulus cycle and the response. This yields the same result as—but is computationally more efficient than—correlating the neural response with the product of the tokens presented in the two neighboring regions of interest (Fig. 1B). Specifically, a +1 in the product of the two sequences represents a frame in which the orientation was uniform

across the two regions, whereas a -1 represents a frame in which the orientation was different in the two regions. Orientation discontinuities are a type of texture boundary. Therefore, a negative second-order kernel signifies a greater response when this type of a texture boundary is present, whereas a positive second-order kernel signifies a greater response when no texture boundary is present. Note that these second-order kernels are terms that isolate interactions between orientations in pairs of regions and cannot be generated by linear summation of orientation-selective responses in individual regions (see icons in Fig. 6).

The 2D grid of grating regions allows us to analyze nearest-neighbor interactions of two kinds. In one kind of interaction, the boundary between the regions is orthogonal to the neuron's preferred orientation; in the other kind, the boundary is parallel. Figure 1B illustrates the computation of a second-order kernel across a boundary orthogonal to the preferred orientation of the neuron, as illustrated by the green lines in Figure 1A. Figure 1B also illustrates the computation of a second-order kernel across a boundary parallel to the preferred orientation of the neuron, as illustrated by the blue lines in Figure 1A.

For comparing the response kernels between neurons, we normalized the kernels (measured between 0 and 200 ms at 10 ms intervals, i.e., 21 time points) by dividing the kernel by the mean of the absolute values, which yields a curve with an average deflection of 1 within the first 200 ms.

Statistics. To determine significant responses for each type of response kernel, we proceeded as follows. First, we only processed kernels with a

signal-to-noise ratio of 1 or higher, with signal-to-noise of a kernel defined as the peak of its absolute response divided by the jackknife estimate of the SD, averaged across all kernels. Next, for each time point and each stimulus region or region combination, we performed a two-tailed one-sample t test to compare the responses to zero, the expected kernel value for a random response. For this test, we used the jackknife estimate of the SD across the 32 repeats for each time point. The significance level was 0.01 and was corrected for multiple comparisons using the Benjamini–Hochberg method, which controls the false discovery rate when test statistics are independent or have positive correlations (Benjamini and Hochberg, 1995, 2001). Then, we performed a two-tailed two-sample t test ($\alpha = 0.01$, corrected for multiple comparisons) at each time point to compare the kernels obtained with the two different taps (combining inverse repeats). This test was done to exclude kernels with a significant dependence on taps, because this indicates the contribution of an overlap with another kernel as described in Stimulation with orientation-discontinuity stimuli, above; 37 of the 653 kernels that had at least one significant time point were excluded for this reason. Note that, independent of the overlap issue, this test is a form of cross-validation because it tests whether model parameters fit from one dataset are consistent with the same parameters fit from another. Finally, we only analyzed second-order kernels between two regions if at least one of them had a significant first-order kernel out of concern that second-order kernels between regions that did not elicit a first-order kernel could be a false positive despite the multiple comparison correction. Finding a second-order kernel or interaction without finding a first-order kernel in any of the regions of interest was a rare occurrence (8 of 104 significant second-order kernels).

Data selection criteria. Because the m-sequence stimulus parameters were established using the collective tuning of the multineuron activity, they were sometimes not optimized for all single neurons identified in the postexperiment spike sorting. Therefore, we restrict the analysis to single neurons with spatial tuning (as determined from grating stimuli) that was sufficiently close to the stimulus parameters used. More specifically, the spatial frequency passband had to include the spatial frequency used in the m-sequence stimulus and the orientation preference had to be no more than 22.5 degrees away from the orientation used in the m-sequence stimulus (this necessarily meant that we only analyzed neurons with well defined orientation tuning). A total of 58 of 122 neurons recorded in V1 and 61 of 90 V2 neurons were aligned in this way. To allow for a direct comparison between the responses to the 20 ms frame duration and those to the 40 ms frame duration stimuli, we only used the data from neurons that showed at least one positive or biphasic significant first-order kernel for both frame durations. This left us with 43 V1 neurons and 45 V2 neurons. Finally, we removed three neurons from the database: two in V1 with double-peaks in the largest first-order kernel and one in V2 with the largest first-order kernel having a negative peak before a positive peak, suggesting that the initial grating characterization of its orientation tuning was misleading. None of those three neurons had any significant second-order kernels, which are the focus of this study.

Classification of V2 neurons into transient and sustained. We showed previously that there are two subpopulations of neurons in V2, one with transient first-order kernels and one with sustained first-order kernels (Schmid et al., 2009), based on principal component analysis (PCA). First-order kernels of transient V2 neurons are biphasic in that a negative peak follows the initial positive peak. We used this criterion here, classifying V2 neurons as transient if a significant (as defined in Statistics, above) negative peak was present in any first-order kernel at either frame duration within the first 200 ms of the response. Categorization based on the shape of the responses using PCA yielded similar results (data not shown here).

Variance explained by adding second-order kernels. To quantify the strength of the second-order kernels, we computed the contribution (i.e., the variance explained) of the second-order kernel to the overall response. For each significant second-order kernel, we first computed the responses to each relevant stimulus configuration (see icons in Figs. 7, 8): the preferred orientation in both stimulus regions (r_{++}), the nonpreferred orientation in both regions (r_{--}), and the two possible combina-

tions of the preferred orientation in one region and the nonpreferred orientation in the other (r_{+-} and r_{-+}) as follows:

$$\begin{aligned} r_{++}(t_i) &= \mu + f_1(t_i) + f_2(t_i) + s(t_i) \\ r_{--}(t_i) &= \mu - f_1(t_i) - f_2(t_i) + s(t_i) \\ r_{+-}(t_i) &= \mu + f_1(t_i) - f_2(t_i) - s(t_i) \\ r_{-+}(t_i) &= \mu - f_1(t_i) + f_2(t_i) - s(t_i) \end{aligned} \quad (7)$$

where μ is the mean firing rate across all stimulus sequences, $f_1(t_i)$ is the first-order kernel for one stimulus region, $f_2(t_i)$ is the first-order kernel for the other stimulus region, $s(t_i)$ is the second-order kernel, and $t_i = 0, 10, \dots, 200$ ms. Equation 7 can be summarized in more compact form as follows:

$$r_{\sigma_1\sigma_2}(t_i) = \mu + \sigma_1 f_1(t_i) + \sigma_2 f_2(t_i) + \sigma_1 \sigma_2 s(t_i), \quad \sigma_1, \sigma_2 \in \{+1, -1\} \quad (8)$$

For cross-validation, we split the four stimulus sequences we used to measure the responses into two sets by grouping each m-sequence and its inverse together; we denote the resulting sets by A and B . For each set, the total response variance was defined as follows:

$$\text{var}_{\text{total}}^{(k)} = \frac{1}{4N} \sum_{\sigma_1, \sigma_2} \sum_{i=1}^N (r_{\sigma_1\sigma_2}^{(k)}(t_i) - \mu)^2, \quad k \in \{A, B\} \quad (9)$$

where N is the number of time points (21 in this case). The cross-validated variance explained by the mean and first-order kernels was defined as follows:

$$\text{var}_f = 1 - \left(\frac{\frac{1}{4} \sum_{\sigma_1, \sigma_2} \left(\text{var}_{f_1; \sigma_1\sigma_2}^{(A \leftarrow B)} + \text{var}_{f_2; \sigma_1\sigma_2}^{(B \leftarrow A)} \right)}{\text{var}_{\text{total}}^{(A)} + \text{var}_{\text{total}}^{(B)}} \right) \right) \quad (10)$$

where $\text{var}_{\sigma_1\sigma_2}^{(A \leftarrow B)}$ is the unexplained variance predicting the total response for set A from the mean and first-order kernel for set B :

$$\text{var}_{f_1; \sigma_1\sigma_2}^{(A \leftarrow B)} = \frac{1}{N} \sum_{i=1}^N (r_{\sigma_1\sigma_2}^{(A)}(t_i) - (\mu + \sigma_1 f_1^{(B)}(t_i) + \sigma_2 f_2^{(B)}(t_i)))^2 \quad (11)$$

and $\text{var}_{\sigma_1\sigma_2}^{(B \leftarrow A)}$ is defined similarly. Note that the prediction in Equation 11 using the mean and first-order kernel only is the same as the responses $r_{\sigma_1\sigma_2}(t_i)$ as defined in Equation 8 but without the interaction term $\sigma_1 \sigma_2 s(t_i)$. The cross-validated variance explained by the first- and second-order kernels together is as follows:

$$\text{var}_f = 1 - \left(\frac{\frac{1}{4} \sum_{\sigma_1, \sigma_2} \left(\text{var}_{f_1; \sigma_1\sigma_2}^{(A \leftarrow B)} + \text{var}_{f_2; \sigma_1\sigma_2}^{(B \leftarrow A)} \right)}{\text{var}_{\text{total}}^{(A)} + \text{var}_{\text{total}}^{(B)}} \right) \right) \quad (12)$$

where $\text{var}_{\sigma_1\sigma_2}^{(A \leftarrow B)}$ is the unexplained variance predicting the total response for set A from the total response for set B , as follows:

$$\text{var}_{f_1; \sigma_1\sigma_2}^{(A \leftarrow B)} = \frac{1}{N} \sum_{i=1}^N \left(r_{\sigma_1\sigma_2}^{(A)}(t_i) - r_{\sigma_1\sigma_2}^{(B)}(t_i) \right)^2 \quad (13)$$

and $r_{\sigma_1\sigma_2}^{(k)}(t_i)$ is the prediction of Equation 8 based on kernels measured from set k .

Finally, the cross-validated fraction of the variance explained by adding the second-order kernel is the difference between the fraction of variance explained by the first- and second-order kernels combined (Equation 12) and the fraction explained by just the first-order kernel (Equation 10) as follows:

$$\text{var}_s = \text{var}_{f_1} - \text{var}_f \quad (14)$$

Because each of the measured responses is computed from an average response over the very large number of instances that contain the relevant pair of orientations (511 or 512 per m-sequence for the 40 ms frame duration, times 8–16 m-sequence presentations), we consider most of the variance to be explainable. Including an estimate for unexplainable variance would increase the amount of explained variance attributable to each type of kernel and would not change their relative importance.

Polarity index. We defined a polarity index to quantify the extent to which responses to texture boundaries of opposite polarity (see icons for columns 3 and 4 in Figs. 7, 8) were similar. The polarity index compares the responses r_{+-} and r_{-+} , and normalizes this by the total variance as follows:

$$PI = \frac{\frac{1}{8N} \sum_{i=1}^N (r_{+-}(t_i) - r_{-+}(t_i))^2}{\text{var}_{total}} = \frac{\frac{1}{8N} \sum_{i=1}^N (2f_1(t_i) - 2f_2(t_i))^2}{\text{var}_{total}} \quad (15)$$

$PI = 0$ when the responses to the two texture boundaries are identical; $PI = 1$ when the responses to the texture boundaries differ maximally and together contribute all of the variance. Note that the difference in responses to the texture boundaries and therefore the numerator of the polarity index is independent of the second-order kernels, but depends on the relative sizes of the two first-order kernels.

Results

We analyzed data from 41 neurons in V1 and 44 neurons in V2 of anesthetized monkeys (V1 and V2: three animals; V1 only: five animals, V2 only: four animals). Data from three of these animals (five neurons in V1, 13 neurons in V2) were already used in a previous study (Schmid et al., 2009). Of the 44 V2 neurons, 15 were transient and 29 were sustained based on the classification described in Materials and Methods.

Responses of individual neurons

We begin by presenting examples of neurons that typify the behavior seen across the dataset and then turn to population statistics. Some V1 neurons showed different response characteristics for the 20 and 40 ms frame durations that relate to both the first-order components (the responses to individual stimulus regions) and second-order components (the interactions between two regions). Figure 2A shows one example. For the 20 ms frame duration (Fig. 2A, left), the neuron responded to two individual regions and both of these first-order kernels were positive (two black traces in the middle of the array). This is what one would expect in the classical receptive field because the neuron should respond more to the preferred orientation than to the orthogonal orientation and it confirms that these two regions provide the strongest stimulation to the classical receptive field. The second-order kernel to these two regions (i.e., the interaction across their border, which is orthogonal to the long axis of the receptive field; see Fig. 1A, green lines for illustration) indicates how signals from these two regions interact. In this example, this orthogonal interaction is also positive (Fig. 2A, left, green trace in the middle of the array), meaning that averaging over responses to both the preferred and nonpreferred orientation, the neuron responded more when the two regions had the same orientation than when they differed (see Fig. 1B for illustration). This behavior is consistent with a simple linear-nonlinear model, in which the nonlinearity has the characteristics of a threshold as follows. If the preferred orientation is present only in one of the two regions within the receptive field, then the neuron's linear input may be only modestly above threshold, so the output is small; however, when both regions in the

receptive field get activated by the preferred orientation, the combined linear input is well above the threshold, producing a disproportionately larger output.

For the 40 ms frame duration, the same neuron shows a quite different response profile (Fig. 2A, right). First, there are more significant first-order kernels now, six in total (Fig. 2A, right, second, third, and fourth rows of black traces, with red points indicating significance). The responses to the two regions that were also present for the 20 ms frame duration look similar, but they last longer, as expected for the longer input signal. However, there are now two more regions that produce positive first-order kernels. This increase in receptive field size may reflect a change in the spatial inputs that are active for this frame duration or simply that responses that were subthreshold at 20 ms are now above threshold. There are also two adjacent regions for which the first-order kernels are negative (Fig. 2A, right, downward black traces in second and third rows, with red points indicating significance). This kind of response is consistent with subtractive surround suppression if the suppression is tuned to the same orientation as the center of the receptive field (iso-orientation surround suppression). Such surround suppression would lead to a smaller response when the preferred orientation is present than when the nonpreferred orientation is present in the region in the receptive field surround, leading to a reduction in suppression and therefore producing a negative first-order kernel. In addition, there is one additional significant second-order kernel that was not seen for the 20 ms frame duration (Fig. 2A, right, downward blue trace, in third row of blue traces, with red points indicating significance). This response is negative in sign, meaning that the neuron responded less when the two regions had the same orientation than when they differed. This is also consistent with iso-orientation surround suppression. Depending on whether the surround suppression is subtractive or divisive (Ayaz and Chance, 2009), it will be seen only in the first-order kernel if it is purely subtractive or in both the first-order and second-order kernels if the surround suppression is either purely divisive or a mixture of subtractive and divisive processes. Interestingly, this interaction takes place for two regions that both have a positive first-order kernel, but the first-order kernel in one region is much smaller, indicating that it is further from the center of the receptive field. This could also mean that the center and surround regions partially overlap. Finally, the second-order kernel orthogonal to the receptive field that was present for the 20 ms frame duration remains (Fig. 2A, right, green trace in center of array), but it has more complex dynamics for the 40 ms frame duration: it is biphasic, first positive and then negative. The faster positive response could be due to a threshold nonlinearity just like for the 20 ms frame duration, but the later negative response cannot be produced by a similar mechanism. This later negative response does not seem to be consistent with divisive iso-orientation surround suppression because it is generated between regions that both appear to be central in the receptive field.

Other V1 neurons showed no responses consistent with surround suppression even at 40 ms. Such an example is shown in Figure 2B. This neuron had four positive first-order kernels for the 20 ms frame duration (Fig. 2B, left, black traces, middle four regions) and 5 positive first-order kernels for the 40 ms frame duration (Fig. 2B, right, black traces with red points indicating significance). For both frame durations, there was one region for which the first-order response was significantly negative at a single 10 ms time bin (Fig. 2B, both panels, second row, second column, see single red dot above trace). There was one positive

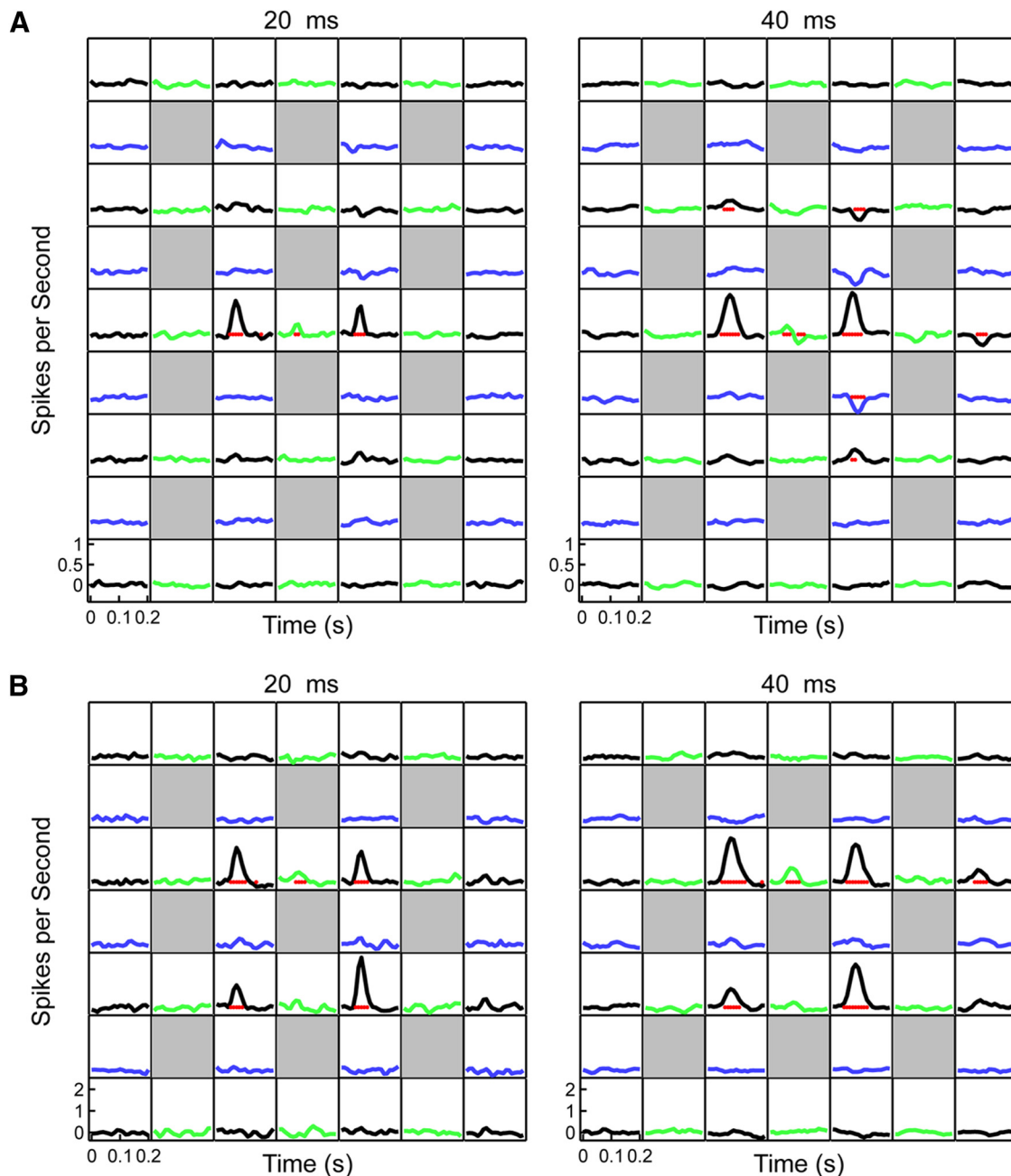


Figure 2. Example V1 neuron response kernels. **A**, First- and second-order kernels of a V1 neuron for the 20 ms frame duration stimulus (left) and the 40 ms frame duration stimulus (right). The responses were measured using the 6×6 layout, but only the responses to 4×5 regions are shown here. Each region was 0.4×0.4 degrees of visual angle. The first-order kernels are plotted in black, second-order kernels for boundaries parallel to the receptive field are plotted in blue, and those for boundaries orthogonal to the receptive field in green (compare with Fig. 1B). Red points on the zero lines mark time points at which the response was significantly different from zero (two-tailed t test, $\alpha = 0.01$, corrected for multiple comparisons). **B**, First- and second-order kernels for another V1 neuron. Layout is as in **A**. Responses were measured using the 6×6 layout, but only the responses to 4×4 regions are shown here. Each region was 0.5×0.5 degrees of visual angle.

second-order kernel between two regions with strong first-order kernels at both frame durations (Fig. 2B, both panels, green trace, middle of second row).

In contrast to V1 neurons, transient V2 neurons typically had biphasic first-order kernels and negative second-order kernels at both the 20 and 40 ms frame durations. Figure 3A shows the responses of one example of a transient V2 neuron. For the 20 ms frame duration (Fig. 3A, left), the neuron responded to three individual regions (black traces in center of array, with red points indicating significance). Although two first-order kernels were positive, just like for the V1 example neuron in Figure 2A, the third response was biphasic—first positive, then negative—

which is the trademark response of transient V2 neurons. The negative part of the response was more pronounced than what we observe for the V1 example neuron in Figure 2B, where, as noted above, the negative component was only significant at a single time point. As we suggested in our previous study, we believe that the negative component of the biphasic V2 responses are due to feedforward excitation followed by delayed feedforward inhibition, both driven by the preferred orientation (Schmid et al., 2009). This leads to a response that is in essence like a temporal differentiator and could be a mechanism for detecting the sudden onset of the preferred stimulus. For the 40 ms frame duration (Fig. 3A, right), there is one additional biphasic first-order kernel

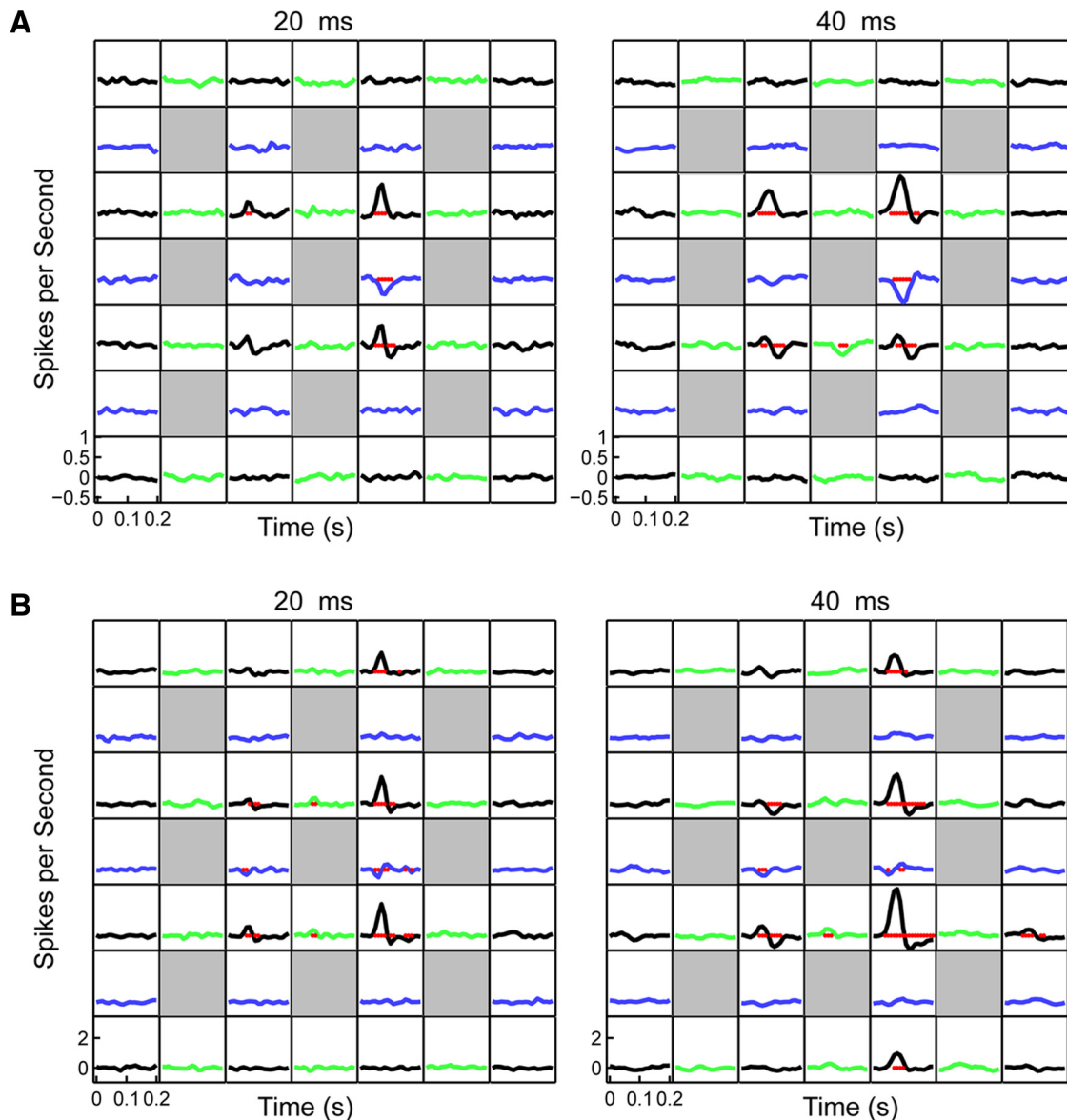


Figure 3. Example transient V2 neuron response kernels. **A**, First- and second-order kernels of a transient V2 neuron for the 20 ms frame duration stimulus (left) and the 40 ms frame duration stimulus (right). The responses were measured using the 4×5 layout, but only the responses to 4×4 regions are shown here. Each region was 0.75×1.0 degrees of visual angle. The first-order kernels are plotted in black, second-order kernels for boundaries parallel to the receptive field are plotted in blue, and those for boundaries orthogonal to the receptive field in green (compare with Fig. 1B). Red points on the zero lines mark time points at which the response was significantly different from zero (two-tailed t test, $\alpha = 0.01$, corrected for multiple comparisons). **B**, First- and second-order kernels for another transient V2 neuron. Layout is as in **A**. The responses were measured using the 6×6 layout, but only the responses to 4×4 regions are shown here. Each region was 1×0.35 degrees of visual angle.

and one biphasic kernel that was only positive for the 20 ms frame duration. More importantly, there is a negative second-order kernel parallel to the receptive field for the 20 ms frame duration (Fig. 3A, left, downward blue trace in middle row) that strengthens at the 40 ms frame duration for this transient V2 neuron (Fig. 3A, right). In addition, there is a negative second-order kernel orthogonal to the receptive field, but only for 40 ms (Fig. 3A, right, downward green trace in third row) and it is much weaker than the one parallel to the receptive field.

Another example of a transient V2 neuron is shown in Figure 3B. For the 20 ms frame duration (Fig. 3B, left), this neuron had biphasic first-order responses for 5 stimulus regions (black traces, top three rows), two positive second-order kernels orthogonal to the receptive field (green traces, middle column) and one negative and one biphasic second-order response parallel to the receptive field (blue traces, middle row).

For the 40 ms frame duration (Fig. 3B, right), the neuron had seven significant first-order kernels (black traces), of which most were also biphasic. There is one stimulus region (Fig. 3B, second row, second column) for which the response at 40 ms is purely negative but was biphasic for the 20 ms frame duration. This negative response has a similar response shape to the biphasic responses, just with a weaker positive peak that does not reach significance, unlike the negative response seen for the example neuron in V1 shown in Figure 2A, which had no matching biphasic responses. The second-order responses were also similar in shape at the 20 and 40 ms frame durations, in particular the negative second-order response parallel to the receptive field (Fig. 3B, both panels, blue trace, middle row, second column).

Sustained V2 neurons had mainly positive first-order and second-order responses. Figure 4 shows the responses of one example of a sustained V2 neuron. For both 20 and 40 ms, there are

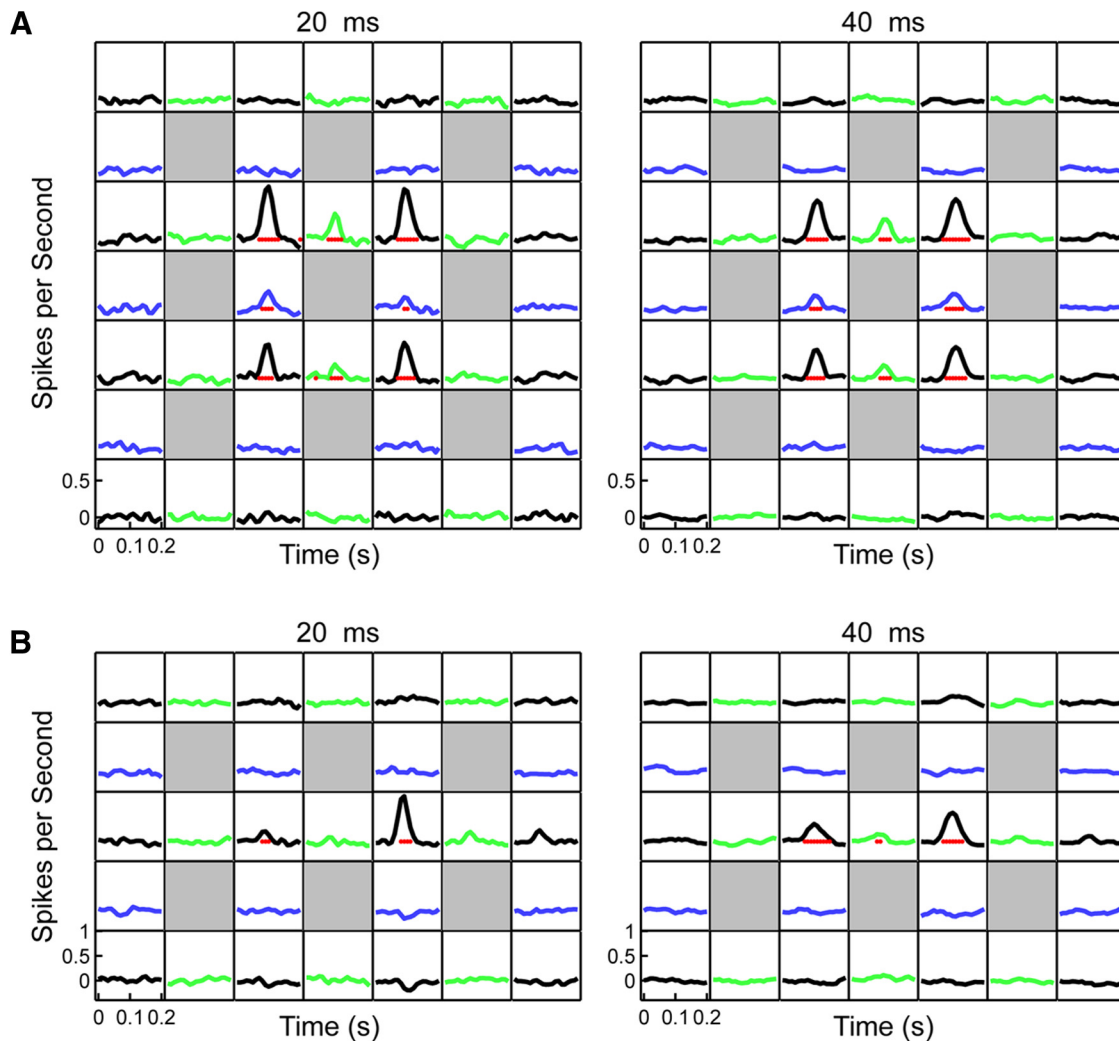


Figure 4. Example sustained V2 neuron response kernels. **A**, First- and second-order kernels of a sustained V2 neuron for the 20 ms frame duration stimulus (left) and the 40 ms frame duration stimulus (right). The responses were measured using the 6×6 layout, but only the responses to 4×4 regions are shown here. Each region was 1.0×1.0 degrees of visual angle. The first-order kernels are plotted in black, second-order kernels for boundaries parallel to the receptive field are plotted in blue, and those for boundaries orthogonal to the receptive field in green (compare with Fig. 1B). Red points on the zero lines mark time points at which the response was significantly different from zero (two-tailed t test, $\alpha = 0.01$, corrected for multiple comparisons). **B**, First- and second-order kernels for another sustained V2 neuron. Layout is as in **A**. The responses were measured using the 6×6 layout, but only the responses to 4×4 regions are shown here. Each region was 0.5×0.25 degrees of visual angle.

four significant first-order kernels, all of which are positive (Fig. 4A, both panels, black traces in middle of array). In addition, all four second-order kernels involving the regions of the stimulus that had first-order kernels are positive and look qualitatively similar for the 20 and 40 ms frame durations (Fig. 4A, both panels, blue and green traces in middle of array). Figure 4B shows another example of a sustained V2 neuron that had two positive first-order kernels at both frame durations (Fig. 4B, both panels, black traces in second row) and one positive second-order kernel orthogonal to the receptive field at the 40 ms frame duration (Fig. 4B, right, green trace, second row).

In sum, whereas frame duration had a large effect on the response dynamics of the example V1 neuron, it had a smaller effect on the response dynamics of the example V2 sustained and transient neurons. In V1, the first- and second-order kernels were exclusively positive or biphasic for the 20 ms frame duration, but some were negative for the 40 ms frame duration. In V2, transient neurons had biphasic first-order kernels and negative second-order kernels at both frame durations, whereas sustained neurons

had mainly positive first- and second-order kernels. As we see in the next section, these features were typical of the population.

Response dynamics

Here, we show that the main response features illustrated by the example cells in Figures 2, 3, and 4 are typical of the population. To compare the shape of the responses independent of the different strengths, we normalized all of the responses by dividing them by the mean of the absolute values (see Materials and Methods for more details) so that the average deflection is 1. Figure 5 shows all significant normalized responses to individual stimulus regions (first-order) for V1, transient V2 neurons, and sustained V2 neurons separately, independent of their position in the receptive field. All responses that reach statistical significance are included (see Materials and Methods for details), allowing for the possibility of more than one response per neuron. The responses for the 20 ms frame duration (Fig. 5A) in V1 are mainly positive (black traces) or biphasic (purple traces), whereas for 40 ms (Fig. 5B), there are also some negative responses (red traces). These nega-

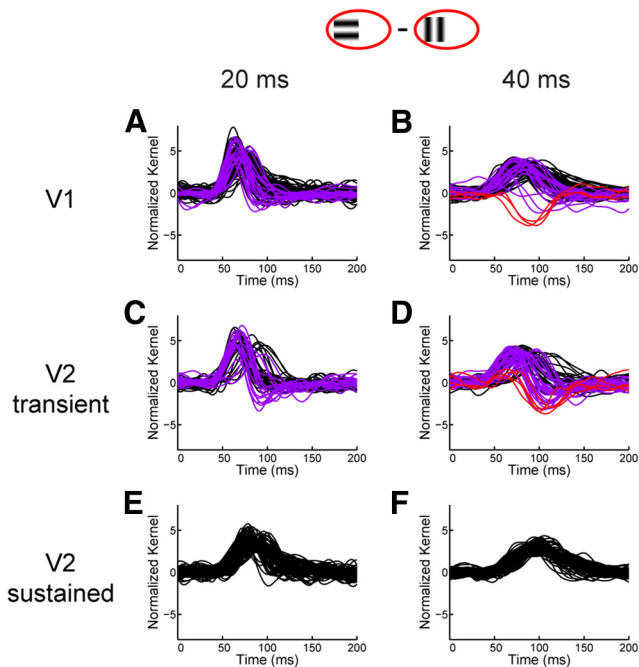


Figure 5. First-order kernels. Icon on top illustrates computation of first-order kernels for any one region of the stimulus: it is the difference in the response to all stimulus frames in which the region contains a grating of the preferred orientation, and the response to all stimulus frames in which the region contains a grating of the nonpreferred orientation. **A**, Normalized first-order kernels of V1 neurons (all significant responses) for 20 ms (see Materials and Methods for details of normalization). Positive responses are plotted in black, biphasic responses in purple, and negative responses in red. Note that a negative kernel value must be statistically significant (see Materials and Methods) to classify the kernel as negative or biphasic. **B**, Same as **A** but for 40 ms. **C**, Normalized first-order kernels of transient V2 neurons for 20 ms. **D**, Same as **C** but for 40 ms. **E**, Normalized first-order kernels of sustained V2 neurons for 20 ms. **F**, Same as **E** but for 40 ms.

tive responses peak at 93 ± 6 ms, only slightly later than the positive responses, which peak at 83 ± 11 ms, but considerably earlier than the negative peak of the biphasic responses, which is at 155 ± 29 ms for the same frame duration. As mentioned above in connection with the responses of a V1 neuron in Figure 2A, we hypothesize that the negative responses are consistent with subtractive iso-orientation surround suppression. If this is true, then this contextual modulation in V1 is apparent only for the longer frame duration.

The first-order kernels of transient V2 neurons are mostly biphasic (first positive, then negative) for both the 20 ms (Fig. 5C) and 40 ms (Fig. 5D) frame durations (purple traces). There are some negative responses for 40 ms (Fig. 5D, red traces), but they have different timing than the negative responses in V1. These negative responses peak at 108 ± 10 ms, which is consistent with the biphasic responses at the same frame duration (which have the negative peak at 130 ± 30 ms), whereas the positive responses peak much earlier, namely at 83 ± 13 ms, a timing very similar to the positive responses in V1 (see last paragraph). The first-order kernels for sustained V2 neurons are positive by definition for both frame durations (Fig. 5E,F, black traces) and the positive responses for the 40 ms frame duration have a peak at 99 ± 11 ms, considerably later than the positive V1 and V2 transient responses. Therefore, overall for V2, there is no qualitative change between the responses for the different frame durations except that the responses are broader in time with the longer frame duration, as one would expect. This is the case for both transient and sustained neurons in V2.

In a similar fashion, we compared the dynamics of the interactions between neighboring regions (second-order kernels). Figure 6 shows all significant second-order kernels across boundaries orthogonal and parallel to the receptive field, as depicted in the icons at the top of the figure, and for the different frame durations separately. In V1 (Fig. 6A–D), these responses are mainly positive for the 20 ms frame duration (Fig. 6A for orthogonal and 6C for parallel, black traces), but for the 40 ms frame duration there are many negative responses (Fig. 6B for orthogonal and 6D for parallel, red traces). As mentioned with regard to the V1 neuron of Figure 2A, we hypothesize that these negative responses are consistent with iso-orientation surround suppression. Over the population, the orientation of the texture boundary does not influence the negative second-order kernels in V1 seen at 40 ms (compare red traces in Fig. 6B,D). This is also consistent with iso-orientation surround suppression, which in some neurons might be more present in the end zones and in others more on the sides of the receptive field, but overall can be present anywhere in the surround of a neuron's receptive field. In our sample, we did not find any individual neurons with negative responses for both orientations of the texture border.

In contrast, the second-order kernels in transient V2 neurons depend on the orientation of the texture boundaries, but not on the frame duration. Specifically, the second-order kernels across orthogonal boundaries are mainly positive for transient V2 neurons at 20 and 40 ms (Fig. 6E,F, black traces), meaning that those neurons fire more if there is no texture boundary orthogonal to the receptive field (consistent with a threshold nonlinearity). However, there were many negative second-order kernels across parallel boundaries for 20 and 40 ms (Fig. 6G,H, red traces), meaning that transient V2 neurons fire more if there is a texture boundary parallel to the receptive field. The single negative second-order kernel orthogonal to the receptive field occurred in a neuron that also had a stronger negative kernel parallel to the receptive field (this is the example neuron shown in Fig. 3A).

For completeness, we describe the dynamics of sustained V2 neurons, but these neurons in general do not respond to texture boundaries. Figure 6, I and J, show the second-order kernels orthogonal to the receptive field that are mainly positive for both frame durations. Figure 6, K and L, show the second-order kernels parallel to the receptive field and they are all positive too, except for two responses at the 40 ms frame duration (Fig. 6L, red and purple traces). Therefore, sustained V2 neurons mainly have nonlinearities consistent with a simple thresholding.

Size of interactions in V1 and V2

As shown above, nonlinear responses to texture boundaries (negative second-order kernels) are seen both in V1 and in transient V2 neurons at the 40 ms frame duration. To compare the strength of the responses to texture boundaries between V1 and V2, we calculated the cross-validated amount of variance in the overall responses explained by adding the second-order kernels to the first-order kernels (see Materials and Methods for details). The variance explained is close to 0 if the second-order kernel is small compared with the first-order kernels and it is close to 1 if the second-order kernel is very large and the first-order kernels are negligible. Note that this value will necessarily be <1 because we use cross-validation and therefore do not expect to explain all of the variance.

Figure 7 shows the responses that formed the starting point for this calculation in V1, focusing on the seven significant negative second-order kernels recorded in that area corresponding to the red traces in Figure 6, B and D (displayed across the rows in Fig.

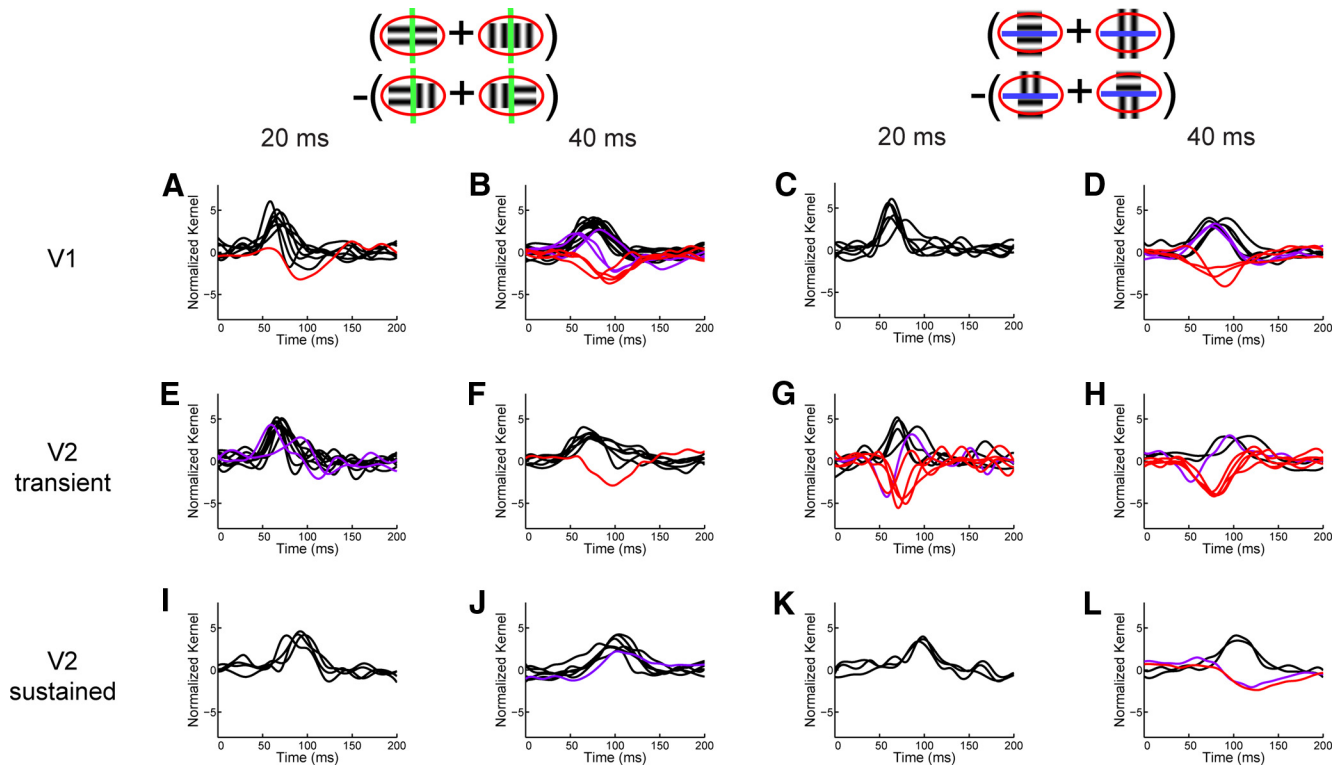


Figure 6. Second-order kernels. Icons on top illustrate computations of second-order kernels orthogonal and parallel to the receptive field: they are the difference in the response to all stimulus frames in which the two regions contain gratings of the same orientation and the response to all stimulus frames in which the regions contain gratings of different orientations. **A**, Normalized second-order kernels orthogonal to the long axis of the receptive field, as illustrated in Figure 1A, of V1 neurons (all significant responses) for 20 ms (see Materials and Methods for details of normalization). Positive responses are plotted in black, biphasic responses in purple, and negative responses in red. **B**, Same as **A** but for 40 ms. **C**, Normalized second-order kernels parallel to the long axis of the receptive field, as illustrated in Figure 1A, of V1 neurons (all significant responses) for 20 ms. **D**, Same as **C** but for 40 ms. **E**, Normalized first-order kernels orthogonal to the receptive field of transient V2 neurons for 20 ms. **F**, Same as **E** but for 40 ms. **G**, Normalized first-order kernels parallel to the receptive field of transient V2 neurons for 20 ms. **H**, Same as **G** but for 40 ms. **I**, Normalized first-order kernels orthogonal to the receptive field of sustained V2 neurons for 20 ms. **J**, Same as **I** but for 40 ms. **K**, Normalized first-order kernels parallel to the receptive field of sustained V2 neurons for 20 ms. **L**, Same as **K** but for 40 ms.

7). Each trace is a linear combination of the mean firing rate, first-order and second-order kernels (see Materials and Methods for further details), but with different weights for the different stimulus combinations, which are depicted as icons in Figure 7. In black in Figure 7 are the responses when the second-order kernel is included, which can be thought of as stimulus-triggered responses with the mean firing rate as the baseline. In gray in Figure 7 are the responses without the second-order kernels included, which can be thought of as the prediction based on the mean firing rate and first-order kernels only. Note that without the second-order kernels, the responses to the first two stimulus combinations are equal and opposite (preferred orientation in both regions and orthogonal orientations in both regions), as are the responses for the last two stimulus combinations (two texture boundaries with opposite polarity). In V1, the predictions without second-order kernels (Fig. 7, solid gray traces) are very similar to the observed responses with second-order kernels (Fig. 7, solid black traces). There is a very modest difference: the responses to the first two stimulus combinations are smaller than predicted by the first-order kernels and the responses to the last two conditions are larger than predicted by the first-order kernels. This difference is the effect of the small negative second-order kernel typically seen in V1.

Figure 8 shows the corresponding analysis of the responses for the transient V2 neurons that had the six significant negative second-order kernels corresponding to the red traces in Figure 6, *F* and *H*. Overall, the difference between the responses with (Fig.

8, solid black traces) and without (Fig. 8, solid gray traces) second-order kernels were substantially more than what was seen in V1. This is most obvious in the last two columns: in some cases, the second-order contribution resulted in a response relative to the mean firing rate baseline that was double the first-order prediction (Fig. 8, third column); in other cases, it eliminated a response that otherwise would have been negative (Fig. 8, fourth column). Note that a response at the baseline level is equal to the average over all four stimulus combinations.

To quantify the comparison between V1 and V2, we computed the amount of variance explained by the second-order kernel cross-validated between the two halves of the data (Figs. 7, 8, dashed and dash-dot lines; see Materials and Methods for details). In V1, there were 7 negative second-order kernels (4 orthogonal and 3 parallel to the receptive field from 6 different neurons) and the variance explained by the first- and second-order kernels together ranged from 0.74 to 0.98 with a median of 0.98, whereas the amount of variance explained by the second-order kernel alone ranged from 0.005 to 0.066 with a median of 0.022. In transient V2 neurons, there were 6 negative second-order kernels (1 orthogonal and 5 parallel to the receptive field from 5 different neurons) and the variance explained by the first- and second-order kernels together ranged from 0.46 to 0.94 with a median of 0.71, whereas the variance explained by the second-order kernel alone ranged from 0.048 to 0.215 with a median of 0.056. The largest value corresponds to the interaction shown in the first row of Figure 8 and to the negative second-order kernel

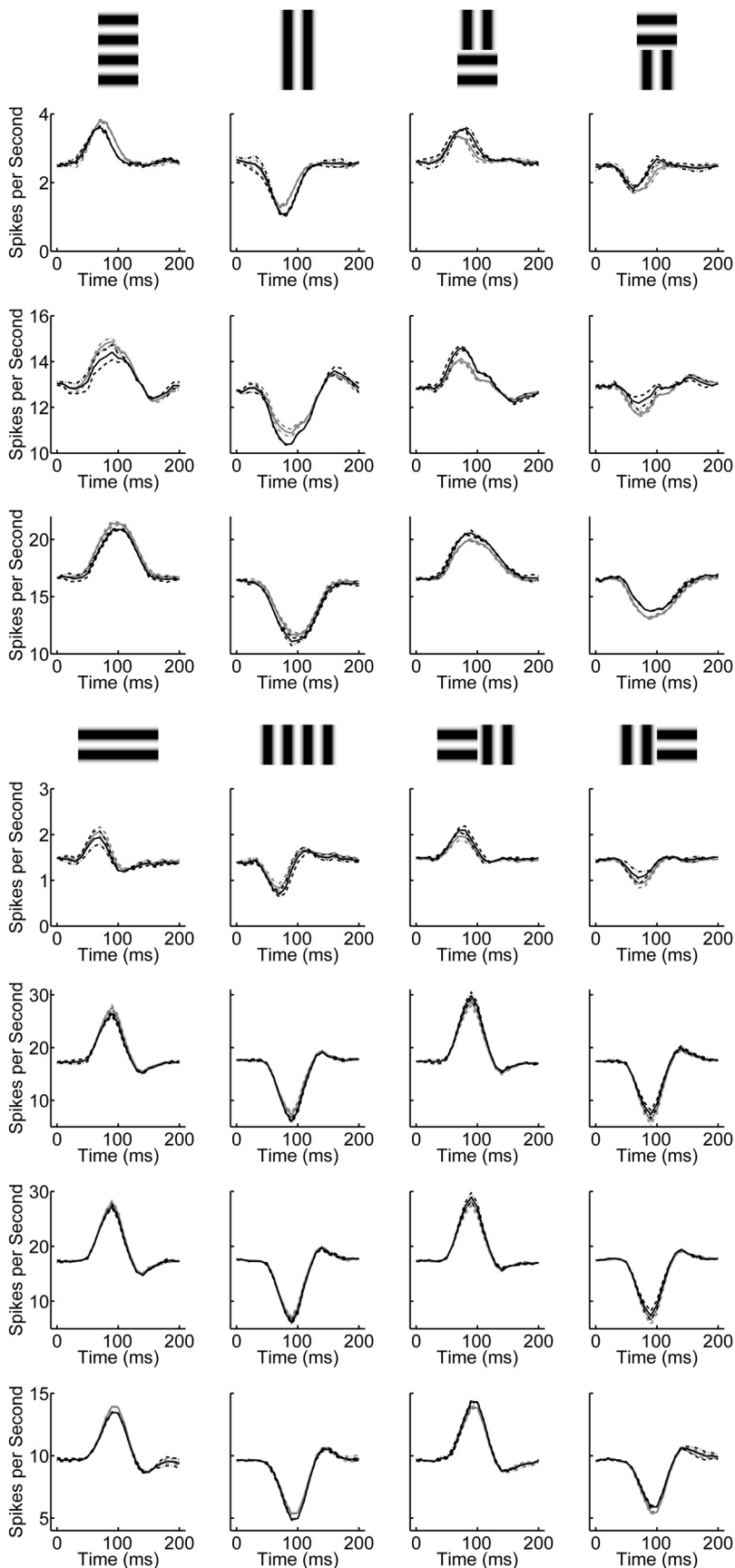


Figure 7. Responses to local stimulus configurations in V1 neurons. Each row shows the responses of V1 neurons to the four possible configurations of orientations in two neighboring regions, represented by the icons at the top of each column. Observed responses are shown in black, predictions without second-order kernels are shown in gray. All interactions that were negative and

parallel to the receptive field of the example neuron shown in Figure 3A (right, blue trace). The negative second-order kernel parallel to the receptive field of the example neuron in Figure 3B (right, blue trace in the center of the second column) has a value close to the median (0.06) and corresponds to the third row in Figure 8. Figure 9 shows the cumulative distribution for the variance explained by second-order kernels (solid black line for negative interactions in V1, solid red line for negative interactions in transient V2). The variance explained by negative second-order kernels was significantly larger in transient V2 neurons than in V1 neurons (two-sample Kolmogorov–Smirnov test, $N_1 = 6$, $N_2 = 7$, $p = 0.037$). This shows that the nonlinear responses to texture boundaries made a stronger contribution in V2 than in V1.

Similarly, we measured the amount of variance explained by the positive second-order kernels (Fig. 9, dashed lines). In V1, there were 18 positive second-order kernels (13 orthogonal and 5 parallel to the receptive field from 13 different neurons) and the variance explained by the first- and second-order kernels together ranged from 0.82 to 0.98 with a median of 0.91, whereas the amount of variance explained by the second-order kernel alone ranged from 0 to 0.064 with a median of 0.021. In transient V2 neurons, there were 8 positive second-order kernels (6 orthogonal and 2 parallel to the receptive field from 6 different neurons) and the variance explained by the first- and second-order kernels together ranged from 0.76 to 0.96 with a median of 0.92, whereas the amount of variance explained by second-order kernels alone ranged from 0.008 to 0.045 with a median of 0.028. The variance explained was significantly larger for negative interactions than for positive interactions in V2 (two-sample Kolmogorov–Smirnov test, $N_1 = 6$, $N_2 = 8$, $p = 0.001$), but not in V1 (two-sample Kolmogorov–Smirnov test, $N_1 = 7$, $N_2 = 18$, $p = 0.99$). The variances explained for positive second-order kernels were not significantly different between V1 and V2 (two-sample Kolmogorov–Smirnov test, $N_1 = 18$, $N_2 = 8$, $p = 0.59$). This shows that the

← statistically significant are shown here. The upper three rows show parallel interactions and the lower four rows show orthogonal interactions. The responses to the texture boundaries were sorted so that the stronger response was in the third column and the weaker response in the fourth column. The dashed and dash-dotted lines are the responses for two halves of the cross-validation (see Materials and Methods for details).

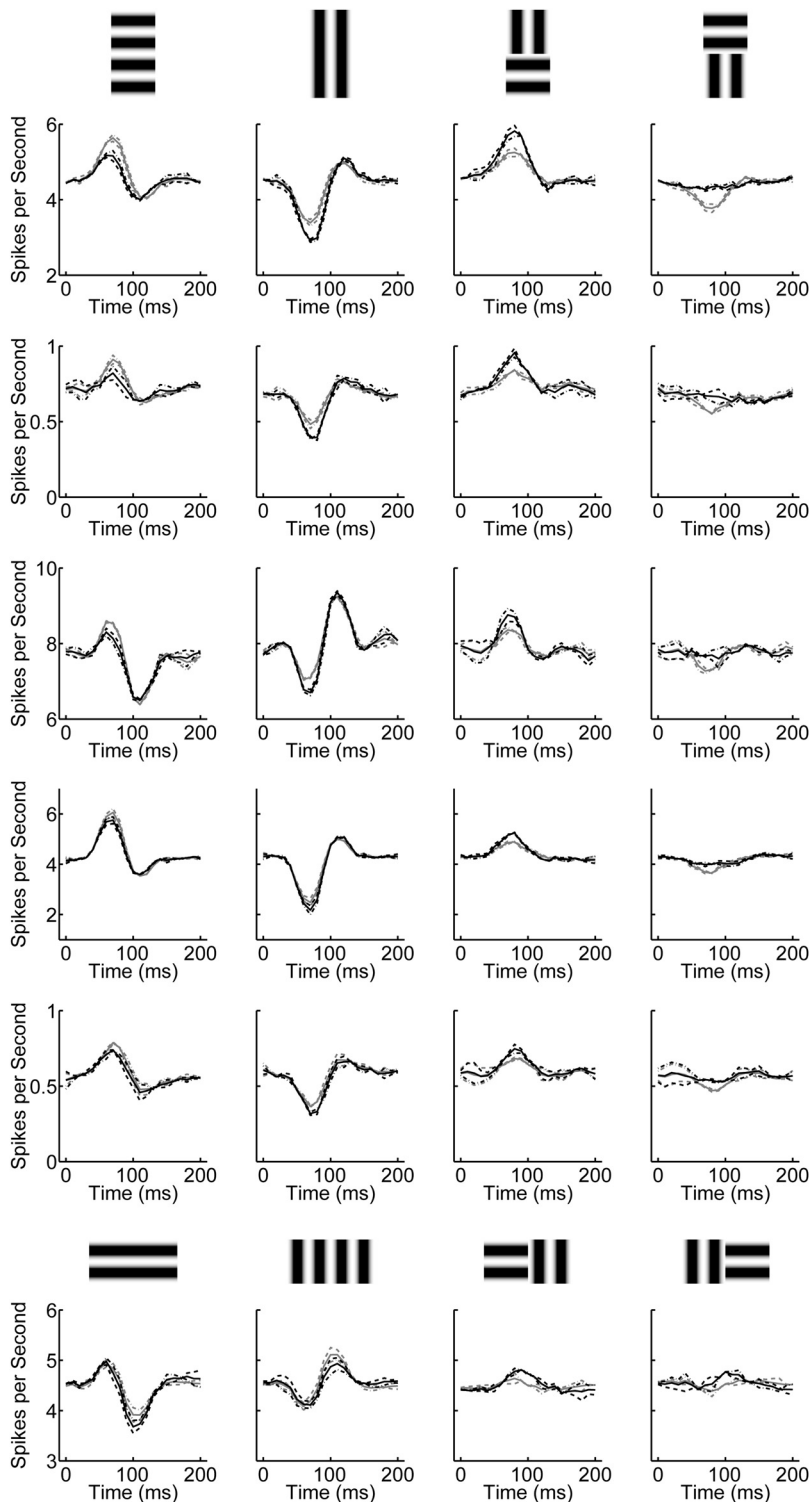


Figure 8. Responses to local stimulus configurations in transient V2 neurons. Each row shows the responses of transient V2 neurons to the four possible configurations of orientations in two neighboring regions, represented by the icons at the top of each column. Observed responses are shown in black, predictions without second-order kernels are shown in gray. All interactions that were negative and significant are shown here. The upper five rows show parallel interactions and the lowest row an orthogonal interaction. Other details are as in Figure 7.

positive interactions, which are probably due to thresholding, are of similar strength in V1 and V2 and that they are weaker than the negative interactions in V2, which correspond to nonlinear responses to texture boundaries. Biphasic second-order kernels (with significant positive and negative peaks) were excluded from

this analysis (five biphasic responses in V1, one in transient V2 neurons).

Spatial properties of interactions in V1 and V2

We hypothesize that the orientation-discontinuity responses in V1 are due to iso-orientation surround suppression, but that in V2, they are generated within the center of the receptive field itself. If this hypothesis is true, the negative second-order kernels in V1 should be produced between a stimulus region within the center of the receptive field and one that is in the surround of the receptive field, whereas those in V2 should occur between two regions that are both within the center of the receptive field. In V1, therefore, the response should be large when the preferred orientation is presented in the stimulus region in the center and the nonpreferred orientation is presented in the surround. But the response should also be very small for the stimulus of the other polarity: nonpreferred orientation in the center and preferred orientation in the surround. In V2, the polarity of the texture boundary should not be crucial because both regions are within the center of the receptive field.

Figures 7 and 8 show that, qualitatively, this expectation is supported. V1 responses shown in Figure 7 depend strongly on the polarity of the texture boundary. It can be seen in Figure 7 that, in V1, the overall responses to the two polarities of the texture boundary are very different from each other (third and fourth columns): it is positive-going for one polarity and negative-going for the other. This implies that one of the regions drove the center much more strongly than the other. In transient V2 neurons, the responses to the two polarities of the texture boundary (Fig. 8, third and fourth column) appeared to be more similar, indicating that each of the two regions involved in the interaction drove the center at a similar level.

To quantify this difference between V1 and V2, we computed the polarity index (see Materials and Methods). The polarity index is close to 0 if the responses are similar for either polarity of the stimulus and it is large if the responses are very different. Note that the polarity index is defined in a way that cancels the contributions of nonlinear interactions between the two

regions (see Materials and Methods) to allow it to focus on a comparison of the orientation preferences of each region individually. For pairs of regions that generated the 7 negative second-order kernels in V1, the polarity index ranged from 0.20 to 0.54 with a median of 0.38. For pairs of regions that generated the 6

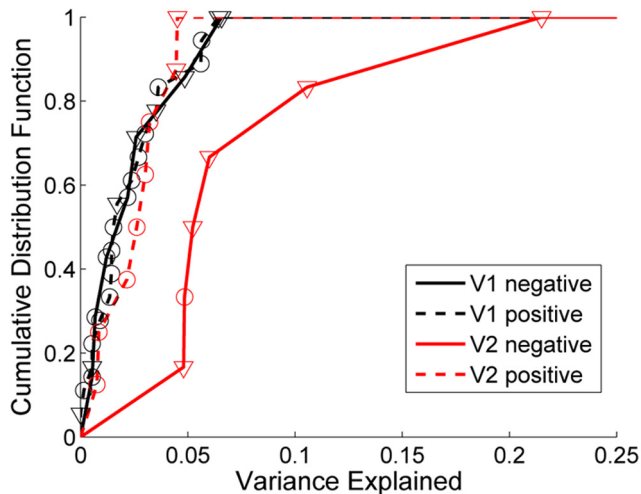


Figure 9. Variance explained by the second-order kernel. Cumulative distribution function of the variance explained for V1 (black lines) and transient V2 neurons (red lines) for significant negative (bold lines) and significant positive interactions (dashed lines). Triangles denote variance explained by second-order kernels parallel to the receptive field and circles variance explained by significant second-order kernels orthogonal to the receptive field.

negative second-order kernels in transient V2 neurons, the polarity index ranged from 0.04 to 0.27 with a median of 0.15. Figure 10 shows the cumulative distribution of the polarity index for negative interactions in V1 (black solid line) and transient V2 neurons (red solid line). As can be seen in Figure 10, the polarity index was smaller in transient V2 neurons. The polarity index is significantly different between V1 and transient V2 neurons (two-sample Kolmogorov–Smirnov test, $N_1 = 7$, $N_2 = 6$, $p = 0.006$).

This analysis shows that the polarity of the texture boundary has less influence on the response in transient V2 neurons than in V1. Because, in transient V2 neurons, both regions that contribute to the interaction have substantial first-order kernels with the same dominant sign, the V2 interaction is not consistent with models based on separate regions of the receptive field with different orientation preferences such as center-surround, or any other spatial arrangement such as suppressive side flanks or end zones. Such models would predict a large polarity index, as seen in V1. Rather, the computation performed by transient neurons is akin to computing the magnitude of the spatial derivative within the center of the receptive field.

To support this interpretation, we consider individual examples. The two interactions shown in the bottom two rows in Figure 7 have a polarity index larger than 0.5. Comparing columns 3 and 4 in that figure shows how the responses for the two texture boundaries is close to symmetric around the mean firing rate. The first row of Figure 7, which shows the same sign flip, corresponds to the negative second-order kernel for the example neuron shown in Figure 2A (right, negative blue trace with red points indicating significance) and has a polarity index of 0.27. Note that the first-order kernels of the interacting regions are both positive, but very different in magnitude. This is consistent with the notion that the larger response is dominated by the center and the smaller response contains antagonistic contributions from the surround, which overlaps with the center. Conversely, an example of a small polarity index (0.11) is the interaction in a V2 neuron shown in the third row of Figure 8. Comparing columns 3 and 4 shows how the response does not flip in sign around the mean firing rate, in contrast to what is observed in V1. This

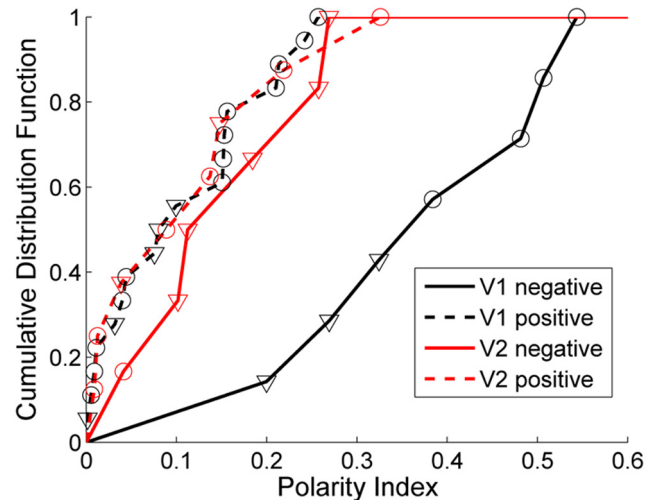


Figure 10. Polarity index. Shown are cumulative distributions of the polarity indexes for V1 (black lines) and transient V2 neurons (red lines) for significant negative (bold lines) and significant positive interactions (dashed lines). Triangles denote polarity indexes measured for second-order kernels parallel to the receptive field and circles those measured for significant second-order kernels orthogonal to the receptive field.

example interaction is also shown in Figure 3B (right, negative blue trace with red points indicating significance), and it can be seen in that figure that the two associated first-order kernels are weak compared with other first-order kernels for that neuron, but of similar strength compared with each other. It is not clear if they both would belong to the “center” or the “surround” of the classical receptive field or their overlap, but, in any case, the similar first-order kernels indicate that this interaction is not consistent with one region being in the center and the other in the surround.

With regard to the positive second-order kernels (seen both in V1 and V2), we hypothesize that they are both due to a simple threshold nonlinearity. Therefore, they should occur well within the receptive field and the polarity index should be small. Across the 18 positive second-order kernels in V1, we found polarity indexes ranging from 0.002 to 0.26 with a median of 0.09. Figure 10 shows their distribution (dashed black line). As can be seen, the polarity index was smaller for the positive second-order kernels in V1 (Fig. 10, black dashed line) than for the negative second-order kernels (black solid line), consistent with the hypothesis (two-sample Kolmogorov–Smirnov test, $N_1 = 7$, $N_2 = 18$, $p = 0.0004$).

In transient V2 neurons, the 8 positive second-order kernels were associated with a polarity index ranging from 0.009 to 0.33 with a median of 0.11. There was no significant difference in polarity indexes between negative and positive second-order kernels in V2 (two-sample Kolmogorov–Smirnov test, $N_1 = 6$, $N_2 = 8$, $p = 0.61$; compare red solid and dashed lines in Fig. 10), which is consistent with the notion that both of those interactions occur at similar positions within the receptive field. There is also no significant difference between positive second-order kernels in V1 versus transient V2 neurons (two-sample Kolmogorov–Smirnov test, $N_1 = 18$, $N_2 = 8$, $p = 0.97$; compare dashed black and red lines in Fig. 10).

In sum, second-order kernels consistent with thresholds (the positive interactions in V1 and V2) and second-order kernels consistent with responses to texture boundaries in V2 (negative interactions across boundaries parallel to the receptive field) occur for responses that have little dependence on the polarity of the

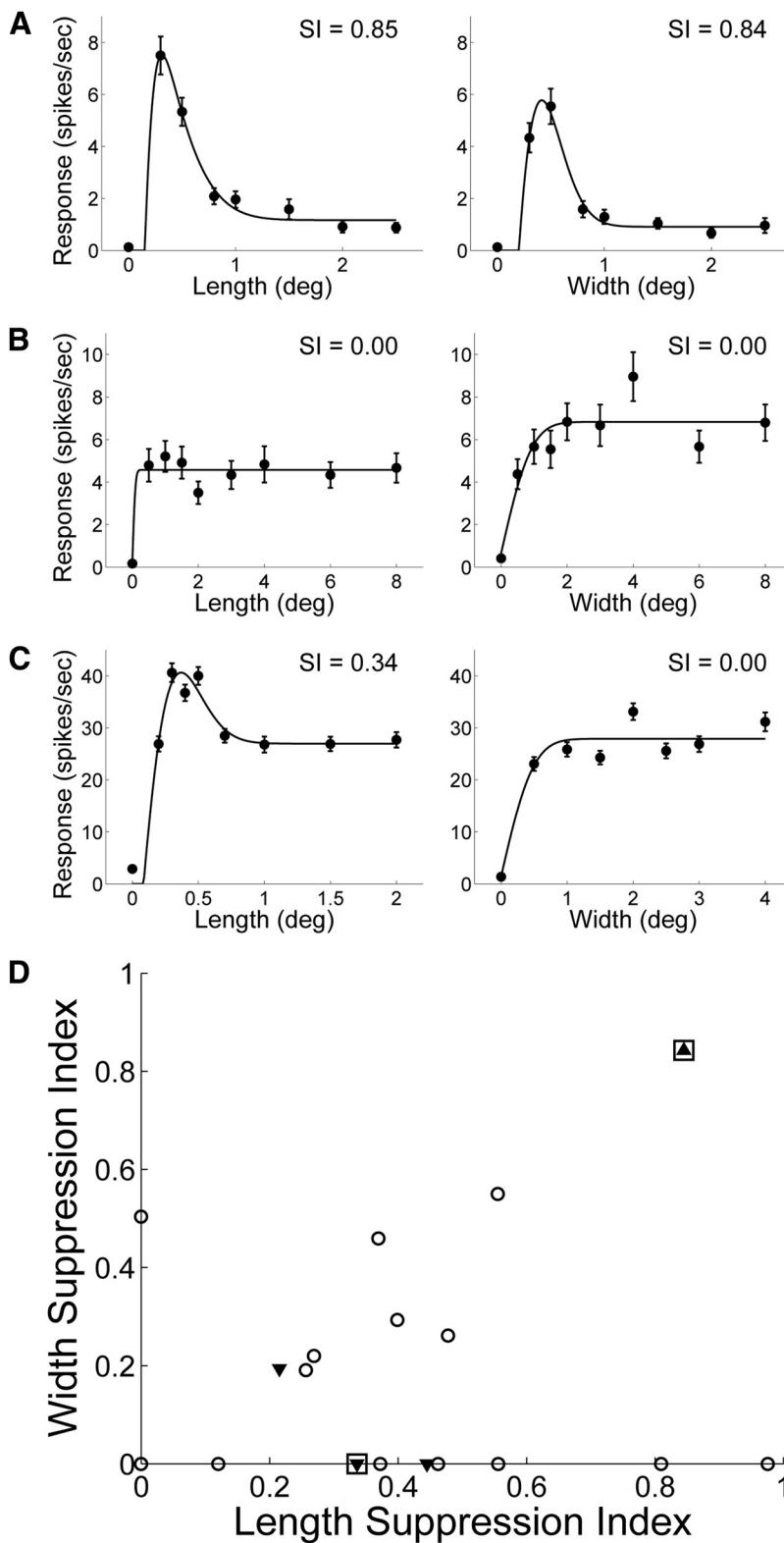


Figure 11. Length- and width-tuning in V1. **A–C**, Length-tuning (left column) and width-tuning (right column) of individual neurons in V1, with their corresponding suppression index (SI). Error bars indicate SE. **D**, SI for length and width for all neurons in V1. Open circles correspond to neurons without any negative kernels for the orientation-discontinuity stimulus with the 40 ms frame duration, squares represent neurons with negative first-order kernels, upward-pointing triangles represent neurons with negative second-order kernels parallel to the receptive field, and downward-pointing triangles represent those with negative second-order kernels orthogonal to the receptive field. Eighteen cells are plotted at the origin; none have negative kernels or width or length suppression.

boundary and thus originate in regions that drive the receptive field center with similar strengths. In contrast, the second-order kernels consistent with responses to orientation discontinuities in V1 (negative interactions either parallel or orthogonal to the receptive field) occur for responses that have a strong dependence on the polarity of the boundary. This in turn means that the two involved stimulus regions have different orientation preferences, which is consistent with one region driving the receptive field center much more strongly than the other or with one region in the center and the other region in the surround.

Surround suppression

As an additional test of the hypothesis that negative kernels are related to surround suppression in V1 but not in V2, we looked at the suppression in the length- and width-tuning curves of our neuron population. Specifically, we hypothesized that, in V1, surround suppression should be associated with negative first-order and second-order kernels, but no such association should be present in V2.

In V1, we measured length- and width-tuning curves in 35 of the 41 neurons recorded (for the other six neurons, the position at which the tuning curves were recorded did not match the position of the orientation-discontinuity stimuli). Figure 11A shows the length- and width-tuning curves for a V1 neuron that had two negative first-order kernels, one in the end-zone of the receptive field and one on the side, and one negative second-order kernel parallel to the receptive field at the 40 ms frame duration (same example neuron as shown in Fig. 2A). The neuron showed strong suppression both along length and width. Figure 11B shows the length- and width-tuning curves for another neuron that had no negative second-order kernels (same example neuron as shown in Fig. 2B); it had no surround suppression. Figure 11C shows the length- and width-tuning curves for another neuron that had two negative second-order kernels orthogonal to the receptive field and one negative first-order kernel in the end-zone of the receptive field at the 40 ms frame duration. This neuron showed suppression along length, which is the corresponding dimension for a second-order kernel orthogonal to the receptive field (see Fig. 1A for illustration). Finally, Figure 11D shows the suppression indexes for all 35 V1 neurons. Neurons with negative second-order interactions were plotted as filled symbols (upward-pointing

triangles for those parallel to the receptive field and downward-pointing triangles for those orthogonal to the receptive field), those with negative first-order kernels were framed with a square, and those without any negative kernels were plotted as open circles. Some neurons scattered around the diagonal, meaning that they had both length and width suppression and two of those had negative kernels. Some neurons had length but no width suppression and one neuron had width but no length suppression. This result is similar to what has been reported previously (DeAngelis et al., 1994) except for the low percentage of neurons with width but no length suppression, which is similar to the percentage of neurons with length but no width suppression in the previous study (see their Figure 4). The previous study had a larger sample size and predominantly sampled neurons with surround suppression, which might explain this difference. There are 18 neurons plotted at the origin without any length or width suppression and none of them had negative kernels.

In V2, we measured length- and width-tuning curves at the same position as the orientation-discontinuity stimuli in 40 of 44 V2 neurons. Figure 12A shows the length- and width-tuning curves for a transient V2 neuron that had a negative second-order kernel parallel to the receptive field and a negative first-order kernel (same example neuron as shown in Fig. 3B). The neuron showed suppression along length, but not along width; the opposite (suppression along width) would be what one would expect if the negative second-order kernel parallel to the receptive field was related to surround suppression (see Fig. 1A for illustration). Figure 12B shows the length- and width-tuning curves for another neuron that had no negative kernels (same example neuron as shown in Fig. 4B); it showed suppression along width. Figure 12C shows the length- and width-tuning curves for a third neuron that had a negative second-order kernel parallel to the receptive field; this neuron showed no surround suppression. To summarize, Figure 12D shows the length and width suppression indexes for all 40 V2 neurons. Interestingly, only four neurons in V2 showed both length and width suppression (not falling on any axes) and they were not close to the diagonal as in V1 (compare with Fig. 11D). As in Figure 11, neurons with negative second-order interactions were plotted as filled symbols (upward-pointing triangles for those parallel to the receptive field and

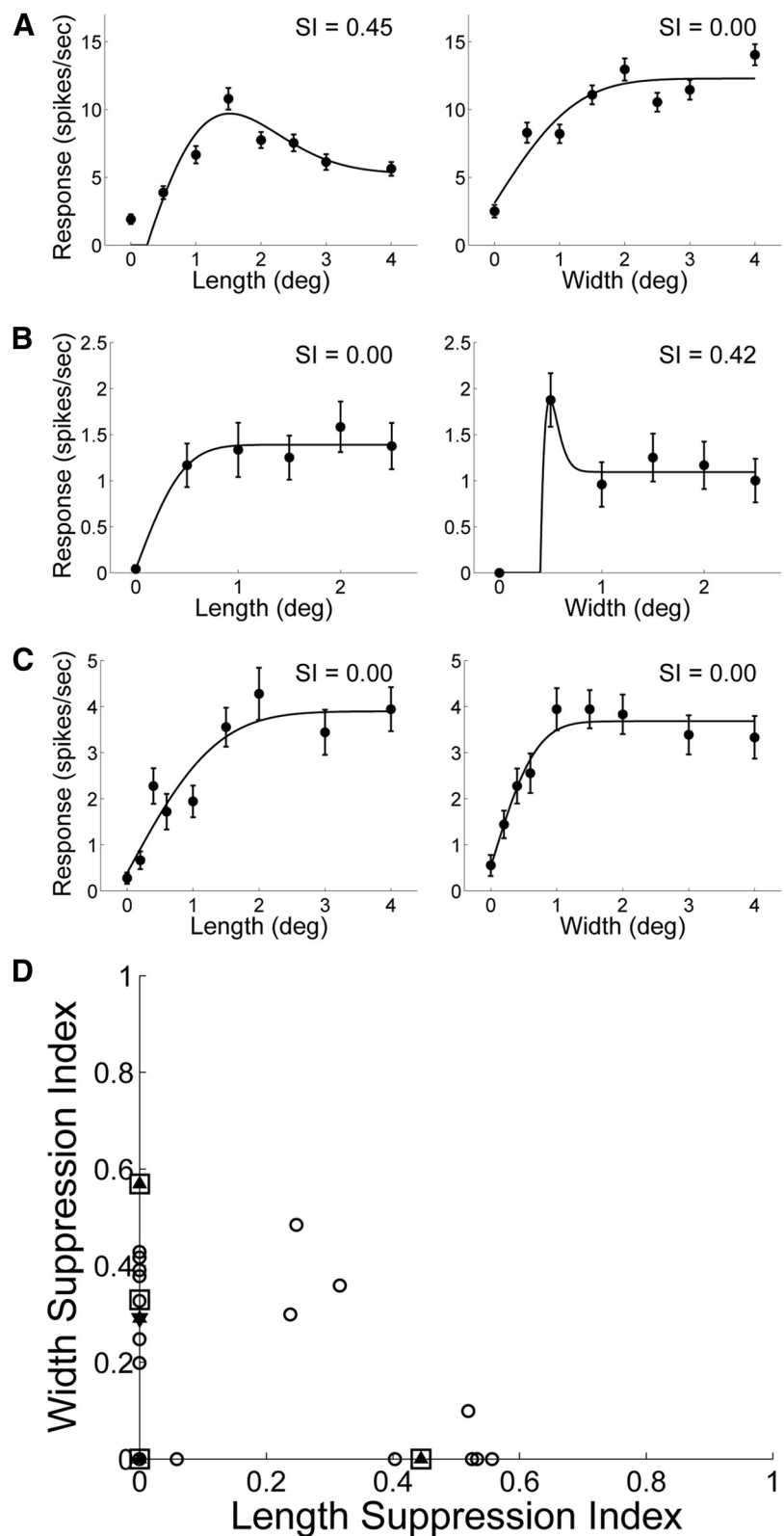


Figure 12. Length- and width-tuning in V2. **A–C**, Length-tuning (left column) and width-tuning (right column) of individual neurons in V2 with corresponding suppression index (SI). Error bars indicate SE. **D**, SI for length and width for all neurons in V1. Open circles correspond to neurons without any negative kernels for the orientation-discontinuity stimulus with the 40 ms frame duration, squares represent neurons with negative first-order kernels, upward-pointing triangles represent neurons with negative second-order kernels parallel to the receptive field, and downward-pointing triangles represent those with negative second-order kernels orthogonal to the receptive field. Twenty cells are plotted at the origin; whereas none have width or length suppression, four have negative kernels.

Table 1. Surround suppression versus negative kernels in V1 and V2

	Surround suppression	Negative kernels		One-sided Fisher's exact test
		Yes	No	
V1	Yes	4	13	$p = 0.046$
	No	0	18	
V2	Yes	4	16	$p = 0.65$
	No	4	16	

Shown are 2×2 contingency tables of number of neurons with and without suppression and with and without negative kernels in V1 and V2. A one-sided Fisher's exact test was performed on each 2×2 contingency table with the hypothesis that surround suppression and negative kernels are associated with each other. Corresponding p -values are listed for each contingency table.

downward-pointing triangles for those orthogonal to the receptive field), those with negative first-order kernels were framed with a square, and those without any negative kernels were plotted as open circles. There were 20 neurons plotted at the origin without any length or width suppression and four of them had negative kernels.

We performed a categorical analysis based on the presence or absence of surround suppression and the presence or absence of negative kernels. Table 1 shows the 2×2 contingency tables for surround suppression and negative kernels in V1 and V2. Of the V1 neurons with negative kernels, all showed surround suppression, whereas only 42% of neurons without negative kernels did; this difference in proportions was significant (Table 1). Interestingly, all three neurons with negative second-order kernels orthogonal to the receptive field had length suppression and the one neuron with a negative second-order kernel parallel to the receptive field had width suppression, which are the corresponding dimensions (compare upward- and downward-pointing triangles in Fig. 11D). In V2, the same analysis revealed very different results. Of the V2 neurons with negative kernels, 50% showed surround suppression, as did 50% of those without negative kernels (Table 1). Therefore, in contrast to the strong association in V1, there was no association between surround suppression and negative kernels in V2.

In sum, negative kernels are associated with surround suppression in V1, but not in V2. Note that surround suppression was similarly prevalent in V1 and V2, so this difference is not simply due to the fraction of neurons with this phenomena (49% of V1 neurons, 50% of V2 neurons, $p = 0.91$, with Yates correction). In addition, the strength of the surround suppression was similar in V1 and V2: in V1, the suppression index ranged from 0 to 0.98 with a mean of 0.16 and, in V2, it ranged from 0 to 0.60 with a mean of 0.11 (two-sample Kolmogorov–Smirnov test, width and length combined, $N_1 = 70$, $N_2 = 80$, $p = 0.86$), consistent with what was reported previously (Shushruth et al., 2009).

Receptive field size

Because V2 receptive fields are known to be larger than those in V1 (Foster et al., 1985; Shushruth et al., 2009), we investigated whether the size difference could explain the differences seen between interactions in V1 and V2. In our sample, there was a trend for larger receptive field sizes in V2 than in V1, but the difference was not very pronounced: the median RF length in V1 was 0.75 (range 0.02–3.06) degrees of visual angle (dva), whereas it was 0.99 (range 0.09–5.68) dva in V2. The median RF width in V1 was 0.85 (range 0.09–4.00) dva and 0.88 (range 0.10–5.1) dva in V2. Reasons for the relatively small difference seen in our study might include definition of receptive field size (see Materials and

Methods), stimulus parameters, and anatomical location. However, we found an interesting difference between transient and sustained V2 neurons: transient neurons had significantly wider receptive fields (median of 1.23, range 0.35–2.56 dva) than sustained neurons (median of 0.56, range 0.1–5.1 dva, two-sample Kolmogorov–Smirnov test, $N_1 = 15$, $N_2 = 25$, $p = 0.007$). This is notable because it supports the idea that the negative second-order responses parallel to the receptive field in transient V2 neurons occur within the receptive field. The receptive fields in transient V2 neurons were also wider than those in V1, even though the difference did not quite reach significance (two-sample Kolmogorov–Smirnov test, $N_1 = 15$, $N_2 = 35$, $p = 0.0501$).

To determine whether the sizing of the orientation-discontinuity stimuli was similar in V1 and V2 relative to the individual receptive field sizes, we computed the ratio of the receptive field length and width to the size of the stimulus region in the corresponding dimension. As indicated in Materials and Methods, we adjusted the size of the stimulus regions to the receptive field; therefore, we anticipated that the ratio should be similar in V1 and V2. In V1, the ratio for the length had a median of 1.77 (range 0.04–3.68) and, in V2, the median was 1.35 (range 0.09–5.68); the difference in distributions was not significant (two-sample Kolmogorov–Smirnov test, $N_1 = 35$, $N_2 = 40$, $p = 0.30$). For the width, the ratio in V1 had a median of 1.8 (range 0.14–5.46) and, in V2, it was 1.2 (range 0.10–5.11); the difference was also not significant (two-sample Kolmogorov–Smirnov test, $N_1 = 35$, $N_2 = 40$, $p = 0.13$). The large range is to be expected because, as mentioned in the Materials and Methods, we sized the stimulus based on multiunit activity at each recording site and not to each individual neuron's receptive field size. The trend was to have more regions of the stimulus within the receptive field in V1 than in V2.

Relative latencies of V1 and V2 receptive field interactions

Finally, we used the notion of “suppression delay” (Bair et al., 2003) to compare the dynamics of responses to orientation discontinuities between V1 and V2. Suppression delay is defined in V1 as the onset latency of the suppression minus the onset latency of the response in the CRF, the area in which stimuli elicit a (positive) response directly. In our stimulus setup, the CRF response corresponds to positive first-order kernels. The suppression is seen as negative first-order or second-order kernels. Without assuming that V2 responses are generated in the same manner, we also measured this quantity in V2 neurons.

In V1, there were three negative first-order kernels in addition to the seven negative second-order kernels. The suppression onset latency measured in those 10 responses had a mean \pm SD of 71 ± 10 ms (range 50–80 ms) and the suppression delay was 23 ± 12 ms (range 0–40 ms). In transient V2 neurons, there were four negative first-order kernels and six negative second-order kernels. The suppression onset latency for those 10 responses was 80 ± 19 ms (range 50–110 ms) and the suppression delay 32 ± 18 ms (range 0–50 ms). Even though both the suppression onset delay and the suppression delay were longer in transient V2 neurons than in V1, there was no significant difference between the two (two-sample Kolmogorov–Smirnov tests, $N_1 = 10$, $N_2 = 10$, $p = 0.11$ for delays; $p = 0.31$ for latencies).

Discussion

Summary of findings

We measured responses in V1 and V2 to orientation discontinuities using two stimulus frame durations, 20 and 40 ms. Comparison of the responses revealed that V1 and V2 responses to

orientation-defined texture boundaries arise independently. In V1, these responses are only present for the 40 ms duration, are not selective for the orientation of the texture boundary, and are associated with surround suppression. In V2, these responses are present at both frame durations, are selective for the orientation of texture boundaries, and are not associated with surround suppression.

Comparisons of V1 and V2

A mechanistic question addressed here is the extent to which responses to texture boundaries are related to surround suppression. In V1, the characteristics of the responses to orientation discontinuities are fully consistent with iso-orientation surround suppression, but this does not appear to be the case in V2. In transient V2 neurons, negative kernels are not associated with surround suppression, whereas there is a strong association in V1 (Table 1). In addition, negative second-order kernels in transient V2 neurons occur between regions of the receptive field with similar strength, accounting for the finding that the polarity of the texture boundary has less influence on the response in V2 than in V1 (Fig. 10).

V1 and V2 responses to orientation discontinuities also differ in their dynamics. Interestingly, this difference is one of temporal resolution, not latency. V2 responses resolve orientation discontinuities that appear for only 20 ms, whereas V1 neurons require 40 ms. These V1 and V2 responses, however, do not differ in either their suppression onset latency or their “suppression delay” (Bair et al., 2003). Our measured suppression delays of 20–30 ms are longer than the delay of ~10 ms reported previously for iso-orientation surround suppression in V1 (Bair et al., 2003). Earlier studies reported a range of 15–25 ms for the delay in V1 (Knierim and van Essen, 1992; Nothdurft et al., 1999). We used much shorter stimulus frame durations than all other studies mentioned here, which could have led to weaker suppression strengths and therefore longer latencies.

Cue invariance

To segment a scene into objects, the orientation of boundaries must be determined from the available cues such as luminance and texture differences between an object and background. Neurons involved in these computations are therefore expected to show cue invariance, that is, a consistent orientation preference independent of cue modality. Transient V2 neurons respond more to texture boundaries parallel to the receptive field than to boundaries that are orthogonal to it, which matches their orientation tuning for luminance boundaries. We therefore suggest that these neurons are involved in image segmentation that includes either luminance or texture boundaries.

In area 18, the V2 homolog in cat, researchers have also shown cue invariance for texture-defined boundaries (Mareschal and Baker, 1998a, 1998b; Zhan and Baker, 2006; Song and Baker, 2007). They compared the orientation tuning of “second-order gratings” with that of standard luminance gratings and found the tunings to be similar. Therefore, these neurons in area 18 demonstrate cue invariance for orientation preference. In these studies, the responses to the second-order gratings were isolated by using spatial frequencies for the first-order luminance gratings that were higher than the resolution of the neuron recorded. The present study used a combinatorial approach (based on m-sequences) to separate first- and second-order kernels. With our approach, it is possible to separate first- and second-order responses even within the spatial frequency passband of the neurons.

Despite these differences in approach, both the Baker studies in cat and the present studies in macaque identified a population of neurons in area 18 or V2 that have cue-invariant responses to texture boundaries. The notion that cue-invariant signaling of boundaries emerges in immediate poststriate processing is also supported by a study in macaque V2, in which orientation tuning was measured with abutting luminance gratings with different spatial phases (“anomalous contours”) and found to be similar to the orientation tuning for luminance bars (von der Heydt and Peterhans, 1989). In addition, a very recent study by Baker et al. in macaque V2 found results similar to those in cat area 18 (Baker et al., 2013).

We did not find such cue invariance of orientation preference in macaque V1. This, along with the differences in dynamics reported here, shows that the contextual response modulations are quite different in V2 and V1. Other researchers have failed to find a consistent difference between responses to texture boundaries defined by differences in orientation in macaque V1 and V2 (El-Shamayleh and Movshon, 2011). Yet others have found a higher percentage of neurons responding to contrast-modulated gratings in area 18 than in area 17 in cats, but no other fundamental differences between the two areas in this regard (Zhou and Baker, 1994). In another study, responses to texture boundaries (illusory contours) were only found in primate V2 and not in V1 (von der Heydt and Peterhans, 1989).

Functional implications

In sum, we find that orientation discontinuities are signaled by neurons in V1 and V2, but the responses have rather different properties: as a population, V2 neurons signal texture-defined boundaries and luminance boundaries in a way that is cue invariant; the V1 population signals the presence of a texture-defined boundary, but not its orientation. For parsing of a scene into objects, however, the orientation of boundaries are crucial. Therefore, the V1 population activity is unlikely to play a major role in scene segmentation based on texture differences.

What, then, is a likely role for the contextual modulations in V1? It has been suggested that they are relevant for texture “pop-out” (Knierim and van Essen, 1992; Kastner et al., 1997, 1999; Li, 1999; Nothdurft et al., 1999). In pop-out, a human subject can effortlessly detect a target feature surrounded by distractors if it differs in an elementary way, such as having a different orientation (Treisman and Gelade, 1980; Bergen and Julesz, 1983a, b; Nothdurft, 1992). The stimuli used in pop-out visual search tasks consist of a target surrounded by distractors. It is not necessary to detect the orientation of any texture boundaries in this context, just their position, so the contextual modulation seen in V1 is suitable as a neural correlate.

Texture segmentation has also been studied extensively in psychophysical experiments (Caelli and Moraglia, 1985; Nothdurft, 1985a, 1985b; McIlhagga et al., 1990; Wolfson and Landy, 1995; Gray and Regan, 1997), sometimes in combination with the pop-out effect (Treisman and Gelade, 1980; Nothdurft, 1992), and it has been suggested that the two processes are directly linked (Nothdurft et al., 2000; Thielscher and Neumann, 2005). Our study suggests that the two processes might not be mediated by the same neural mechanism, but rather that V1 contextual modulation is the basis for pop-out whereas transient V2 responses to texture boundaries are the basis of texture segmentation.

This is similar to findings in alert monkeys (Marcus and Van Essen, 2002). Using a stimulus composed of a disk defined by orientation contrast, those investigators found stronger modulation in V2 than in V1; a difference that did not reach significance

at the population level (their Figure 4). They interpreted the modulations in V1 and V2 as part of a distributed, hierarchical scene segmentation process. The differences identified here between V1 and V2 suggest that they have separate and definable roles. In addition, they show that attention had only a modest effect on the enhancement of firing rates in both V1 and V2, indicating that segmentation in V1 and V2 can well be studied in anesthetized animals.

Interestingly, the difference in temporal resolution of the V1 and V2 responses to orientation discontinuities supports the notion that the former is primarily related to pop-out, whereas the latter is primarily related to orientation-based texture discrimination. It was shown in psychophysical experiments in humans that the stimulus duration needed to detect a single line is shorter than the stimulus duration needed for the pop-out effect in the orientation domain (Nothdurft, 2000). Conversely, the temporal resolution for orientation-based texture segregation was found not to be lower than for local orientation coding (Motoyoshi and Nishida, 2001). This is consistent with the idea that pop-out is mediated by contextual modulations in V1 that are slower to develop, whereas texture segmentation is mediated by transient neurons in V2, in which responses to texture boundaries have a higher temporal resolution, possibly mediated by a filter-rectify-filter mechanism (Chubb and Sperling, 1988; Zavitz and Baker, 2013).

Conclusions

We find that texture boundaries defined by a difference in orientation elicit different kinds of responses in V1 and V2. Responses in V1 are consistent with iso-orientation surround suppression, whereas responses in transient V2 neurons are more akin to a spatiotemporal differentiation of V1 inputs (Schmid et al., 2009). Responses in V1 have lower temporal resolution than those in V2. In addition, over the population, responses to texture boundaries in V1 are not tuned to the orientation of the boundary, whereas those in V2 are. Therefore, we hypothesize that the responses in V1 are involved in pop-out phenomena, whereas those in transient V2 neurons are involved in texture segmentation.

References

- Adams DL, Horton JC (2006) Monocular cells without ocular dominance columns. *J Neurophysiol* 96:2253–2264. [CrossRef Medline](#)
- Akaike H (1974) A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19:716–723. [CrossRef](#)
- Ayaz A, Chance FS (2009) Gain modulation of neuronal responses by subtractive and divisive mechanisms of inhibition. *J Neurophysiol* 101:958–968. [Medline](#)
- Bair W, Cavanaugh JR, Movshon JA (2003) Time course and time-distance relationships for surround suppression in macaque V1 neurons. *J Neurosci* 23:7690–7701. [Medline](#)
- Baker C, Li G, Wang Z, Yao Z, Yuan N, Talebi V, Tan J, Wang Y, Zhou Y (2013) Second-order neuronal responses to contrast modulation stimuli in primate visual cortex. *J Vis* 13(9):41.
- Benardete EA, Victor JD (1994) An extension of the m-sequence technique for the analysis of multi-input nonlinear systems. In: *Advanced methods of physiological system modeling* (Marmarelis VZ, ed), pp 87–110. New York: Plenum.
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Part B* 57:289–300.
- Benjamini Y, Hochberg Y (2001) The control of the false discovery rate in multiple testing under dependency. *Ann Stat* 29:1165–1188. [CrossRef](#)
- Bergen JR, Julesz B (1983a) Parallel versus serial processing in rapid pattern discrimination. *Nature* 303:696–698. [CrossRef Medline](#)
- Bergen JR, Julesz B (1983b) Rapid discrimination of visual patterns. *IEEE Transact Syst Man Cyb* 13:857–863.
- Blakemore C, Tobin EA (1972) Lateral inhibition between orientation detectors in the cat's visual cortex. *Exp Brain Res* 15:439–440. [Medline](#)
- Caelli T, Moraglia G (1985) On the detection of Gabor signals and discrimination of Gabor textures. *Vision Res* 25:671–684. [CrossRef Medline](#)
- Chubb C, Sperling G (1988) Drift-balanced random stimuli: a general basis for studying non-Fourier motion perception. *J Opt Soc Am A* 5:1986–2007. [CrossRef Medline](#)
- DeAngelis GC, Freeman RD, Ohzawa I (1994) Length and width tuning of neurons in the cat's primary visual cortex. *J Neurophysiol* 71:347–374. [Medline](#)
- Derrington AM, Parker A, Barraclough NE, Easton A, Goodson GR, Parker KS, Tinsley CJ, Webb BS (2002) The uses of colour vision: behavioural and physiological distinctiveness of colour stimuli. *Philos Trans R Soc Lond B Biol Sci* 357:975–985. [CrossRef Medline](#)
- El-Shamayleh Y, Movshon JA (2011) Neuronal responses to texture-defined form in macaque visual area V2. *J Neurosci* 31:8543–8555. [CrossRef Medline](#)
- Foster KH, Gaska JP, Nagler M, Pollen DA (1985) Spatial and temporal frequency selectivity of neurones in visual cortical areas V1 and V2 of the macaque monkey. *J Physiol* 365:331–363. [Medline](#)
- Fries W, Albus K, Creutzfeldt OD (1977) Effects of interacting visual patterns on single cell responses in cats striate cortex. *Vision Res* 17:1001–1008. [CrossRef Medline](#)
- Gegenfurtner KR, Rieger J (2000) Sensory and cognitive contributions of color to the recognition of natural scenes. *Curr Biol* 10:805–808. [CrossRef Medline](#)
- Golomb SW (1981) Shift register sequences, revised edition. Laguna Hills, CA: Aegean Park.
- Gray R, Regan D (1997) Vernier step acuity and bisection acuity for texture-defined form. *Vision Res* 37:1717–1723. [Medline](#)
- Grosf DH, Shapley RM, Hawken MJ (1993) Macaque V1 neurons can signal 'illusory' contours. *Nature* 365:550–552. [CrossRef Medline](#)
- Hazan L, Zugaro M, Buzsáki G (2006) Klusters, NeuroScope, NDManager: a free software suite for neurophysiological data processing and visualization. *J Neurosci Methods* 155:207–216. [CrossRef Medline](#)
- Kastner S, Nothdurft HC, Pigarev IN (1997) Neuronal correlates of pop-out in cat striate cortex. *Vision Res* 37:371–376. [CrossRef Medline](#)
- Kastner S, Nothdurft HC, Pigarev IN (1999) Neuronal responses to orientation and motion contrast in cat striate cortex. *Vis Neurosci* 16:587–600. [CrossRef Medline](#)
- Kingdom FA (2003) Color brings relief to human vision. *Nat Neurosci* 6:641–644. [CrossRef Medline](#)
- Kingdom FA, Beauce C, Hunter L (2004) Colour vision brings clarity to shadows. *Perception* 33:907–914. [CrossRef Medline](#)
- Knierim JJ, van Essen DC (1992) Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *J Neurophysiol* 67:961–980. [Medline](#)
- Leventhal AG, Wang Y, Schmolesky MT, Zhou Y (1998) Neural correlates of boundary perception. *Vis Neurosci* 15:1107–1118. [Medline](#)
- Li Z (1999) Contextual influences in V1 as a basis for pop out and asymmetry in visual search. *Proc Natl Acad Sci U S A* 96:10530–10535. [CrossRef Medline](#)
- Marcus VL, Raiguel SE, Xiao D, Orban GA (2000) Processing of kinetically defined boundaries in areas V1 and V2 of the macaque monkey. *J Neurophysiol* 84:2786–2798. [Medline](#)
- Marcus DS, Van Essen DC (2002) Scene segmentation and attention in primate cortical areas V1 and V2. *J Neurophysiol* 88:2648–2658. [CrossRef Medline](#)
- Mareschal I, Baker CL Jr (1998a) Temporal and spatial response to second-order stimuli in cat area 18. *J Neurophysiol* 80:2811–2823. [Medline](#)
- Mareschal I, Baker CL Jr (1998b) A cortical locus for the processing of contrast-defined contours. *Nat Neurosci* 1:150–154. [CrossRef Medline](#)
- McIlhagga W, Hine T, Cole GR, Snyder AW (1990) Texture segregation with luminance and chromatic contrast. *Vision Res* 30:489–495. [CrossRef Medline](#)
- Mechler F, Reich DS, Victor JD (2002) Detection and discrimination of relative spatial phase by V1 neurons. *J Neurosci* 22:6129–6157. [Medline](#)
- Mechler F, Victor JD, Ohiorhenuan I, Schmid AM, Hu Q (2011) Three-dimensional localization of neurons in cortical tetrode recordings. *J Neurophysiol* 106:828–848. [CrossRef Medline](#)
- Motoyoshi I, Nishida S (2001) Temporal resolution of orientation-based texture segregation. *Vision Res* 41:2089–2105. [CrossRef Medline](#)

- Nelson JL, Frost BJ (1978) Orientation-selective inhibition from beyond the classic visual receptive field. *Brain Res* 139:359–365. [CrossRef Medline](#)
- Nothdurft H (2000) Saliency from feature contrast: temporal properties of saliency mechanisms. *Vision Res* 40:2421–2435. [CrossRef Medline](#)
- Nothdurft HC (1985a) Sensitivity for structure gradient in texture discrimination tasks. *Vision Res* 25:1957–1968. [CrossRef Medline](#)
- Nothdurft HC (1985b) Orientation sensitivity and texture segmentation in patterns with different line orientation. *Vision Res* 25:551–560. [CrossRef Medline](#)
- Nothdurft HC (1992) Feature analysis and the role of similarity in preattentive vision. *Percept Psychophys* 52:355–375. [CrossRef Medline](#)
- Nothdurft HC, Gallant JL, Van Essen DC (1999) Response modulation by texture surround in primate area V1: correlates of “popout” under anesthesia. *Vis Neurosci* 16:15–34. [Medline](#)
- Nothdurft HC, Gallant JL, Van Essen DC (2000) Response profiles to texture border patterns in area V1. *Vis Neurosci* 17:421–436. [CrossRef Medline](#)
- Sceniak MP, Ringach DL, Hawken MJ, Shapley R (1999) Contrast’s effect on spatial summation by macaque V1 neurons. *Nat Neurosci* 2:733–739. [CrossRef Medline](#)
- Schmid AM (2008) The processing of feature discontinuities for different cue types in primary visual cortex. *Brain Res* 1238:59–74. [CrossRef Medline](#)
- Schmid AM, Purpura KP, Ohiorhenuan IE, Mechler F, Victor JD (2009) Subpopulations of neurons in visual area v2 perform differentiation and integration operations in space and time. *Front Syst Neurosci* 3:15. [CrossRef Medline](#)
- Shushruth S, Ichida JM, Levitt JB, Angelucci A (2009) Comparison of spatial summation properties of neurons in macaque V1 and V2. *J Neurophysiol* 102:2069–2083. [CrossRef Medline](#)
- Sillito AM, Grieve KL, Jones HE, Cudeiro J, Davis J (1995) Visual cortical mechanisms detecting focal orientation discontinuities. *Nature* 378:492–496. [CrossRef Medline](#)
- Song Y, Baker CL Jr (2007) Neuronal response to texture- and contrast-defined boundaries in early visual cortex. *Vis Neurosci* 24:65–77. [Medline](#)
- Sutter EE (1992) A deterministic approach to nonlinear systems analysis. In: *Nonlinear vision: determination of neural receptive fields, function, and networks* (Pinter RB, Nabet B, eds), pp 171–220. Boca Raton, FL: CRC.
- Thielscher A, Neumann H (2005) Neural mechanisms of human texture processing: texture boundary detection and visual search. *Spat Vis* 18: 227–257. [CrossRef Medline](#)
- Treisman AM, Gelade G (1980) A feature-integration theory of attention. *Cogn Psychol* 12:97–136. [CrossRef Medline](#)
- Victor JD, Mechler F, Repucci MA, Purpura KP, Sharpee T (2006) Responses of V1 neurons to two-dimensional hermite functions. *J Neurophysiol* 95:379–400. [Medline](#)
- von der Heydt R, Peterhans E (1989) Mechanisms of contour perception in monkey visual cortex. I. Lines of pattern discontinuity. *J Neurosci* 9:1731–1748. [Medline](#)
- von der Heydt R, Zhou H, Friedman HS (2000) Representation of stereoscopic edges in monkey visual cortex. *Vision Res* 40:1955–1967. [CrossRef Medline](#)
- Wolfson SS, Landy MS (1995) Discrimination of orientation-defined texture edges. *Vision Res* 35:2863–2877. [CrossRef Medline](#)
- Zavitz E, Baker CL (2013) Texture sparseness, but not local phase structure, impairs second-order segmentation. *Vision Res* 91:45–55. [CrossRef Medline](#)
- Zhan CA, Baker CL Jr (2006) Boundary cue invariance in cortical orientation maps. *Cereb Cortex* 16:896–906. [CrossRef Medline](#)
- Zhou YX, Baker CL Jr (1994) Envelope-responsive neurons in areas 17 and 18 of cat. *J Neurophysiol* 72:2134–2150. [Medline](#)