

Perceptual Texture Dimensions Modulate Neuronal Response Dynamics in Visual Cortical Area V4

✉ Taekjun Kim,^{1,2} Wyeth Bair,^{1,2} and ✉ Anitha Pasupathy^{1,2}

¹Department of Biological Structure, and ²Washington National Primate Research Center, University of Washington, Seattle, Washington 98195

Texture is an important visual attribute for surface pattern discrimination and therefore object segmentation, but the neural bases of texture perception are largely unknown. Previously, we demonstrated that the responses of V4 neurons to naturalistic texture patches are sensitive to four key features of human texture perception: coarseness, directionality, regularity, and contrast. To begin to understand how distinct texture perception emerges from the dynamics of neuronal responses, in 2 macaque monkeys (1 male, 1 female), we investigated the relative contribution of the four texture attributes to V4 responses in terms of the strength and timing of response modulation. We found that the different feature dimensions are associated with different temporal dynamics. Specifically, the response modulation associated with directionality and regularity was significantly delayed relative to that associated with coarseness and contrast, suggesting that the latter are fundamentally simpler feature dimensions. The population of texture-selective neurons could be grouped into multiple clusters based on the combination of feature dimensions encoded, and those subpopulations displayed distinct temporal dynamics characterized by the weighted combinations of multiple features. Finally, we applied a population decoding approach to demonstrate that texture category information can be obtained from short temporal windows across time. These results demonstrate that the representation of different perceptually relevant texture features emerge over time in the responses of V4 neurons. The observed temporal organization provides a framework to interpret how the processing of surface features unfolds in early and midlevel cortical stages, and could ultimately inform the interpretation of perceptual texture dynamics.

Key words: area V4; object recognition; rhesus monkey; temporal dynamics; texture perception; ventral visual pathway

Significance Statement

To delineate how neuronal responses underlie our ability to perceive visual textures, we related four key perceptual dimensions (coarseness, directionality, regularity, and contrast) of naturalistic textures to the strength and timing of modulation of neuronal responses in area V4, an intermediate stage in the form-processing, ventral visual pathway. Our results provide the first characterization of V4 temporal dynamics for texture encoding along perceptually defined axes.

Introduction

Texture carries critical visual information about object surfaces, and many past studies have reported that neurons in ventral visual pathway respond sensitively to different visual texture features (Knierim and Van Essen, 1992; Lamme, 1995; Kastner et al., 2000; Merigan, 2000; El-Shamayleh and Movshon, 2011).

Received May 6, 2021; revised Nov. 17, 2021; accepted Nov. 22, 2021.

Author contributions: T.K., W.B., and A.P. designed research; T.K. performed research; T.K. analyzed data; T.K. wrote the first draft of the paper; T.K., W.B., and A.P. edited the paper; T.K., W.B., and A.P. wrote the paper.

This work was supported by National Eye Institute Grant R01 EY018839 and National Science Foundation CRCNS Grant IIS-1309725 to A.P.; National Eye Institute R01 EY027023 to W.B.; National Eye Institute Center Core Grant for Vision Research P30 EY01730 to the University of Washington; and National Institutes of Health/ORIP Grant P51 OD010425 to the Washington National Primate Research Center. We thank all members of A.P. laboratory for helpful discussions and comments on the manuscript; and Amber Fyall for assistance with animal training.

The authors declare no competing financial interests.

Correspondence should be addressed to Anitha Pasupathy at pasupat@u.washington.edu.

<https://doi.org/10.1523/JNEUROSCI.0971-21.2021>

Copyright © 2022 the authors

However, making progress on the question of how texture is processed has been slow, partly because it is difficult to know how to parameterize the vast space of all possible textures and how to choose stimuli. Early studies on texture representation were typically conducted with spatially homogeneous patterns composed of separated elements, such as lines or forms, but more recent studies have used naturalistic textures. Portilla and Simoncelli (2000) proposed that perceptually equivalent textures could be synthesized from summary statistics of a natural texture image represented by correlated outputs of multiscale-oriented linear (V1 cell-like) filters. The summary statistics include correlations between spatially neighboring filters, correlations between filters with neighboring orientations, and so on. Subsequent studies inspired by this idea have shown that neuronal responses in areas V2 and V4, but not V1, represent naturalistic image structure, which can be better explained on the basis of higher-order statistics than by simple orientation/spatial frequency (SF) tuning (Freeman et al., 2013; Okazawa et al., 2015, 2017; Ziembka et al., 2016). These studies have provided important insights into

the nature of cortical representation for naturalistic textures. However, because of the high dimensionality of these statistics, it is still poorly understood which texture feature combinations give rise to the perceptual distinctions between coarse and fine textures, regular and irregular textures, etc., and how multiple features are dynamically encoded in neuronal activity.

To gain traction on the question of how neuronal responses underlie texture perception, we sought to represent textures in terms of a small set of perceptual attributes; and for this, we turned to literature on human texture perception and computer vision. The consensus across several studies (e.g., Tamura et al., 1978; Liu and Picard, 1996; Rao and Lohse, 1996) was that the four texture attributes that we term coarseness, contrast, directionality, and regularity (but referred to by a variety of different names in the literature, e.g., line-likeness, repetitiveness) were critical to explain texture similarity judgments in human subjects. Furthermore, these perceptual dimensions were also consistent with the mutually orthogonal subfields derived from an image model based on 2D Wold decomposition of homogeneous random fields (Liu and Picard, 1996). We parameterized natural texture stimuli along these four dimensions (Kim et al., 2019) and analyzed whether differences in temporal dynamics of macaque V4 neuronal responses could be described by selectivity for these perceptually relevant texture features. With these experiments, we sought to gain insights into the complexity of computations associated with the different texture dimensions.

We found that many V4 neurons exhibited distinct temporal dynamics in response to texture stimuli, and the response dynamics were closely associated with texture feature dimensions that individual cells encode. Specifically, the processing of directionality and regularity was significantly delayed relative to that of coarseness and contrast dimensions. Furthermore, through population analysis, we showed that it is possible to cluster neurons that produce similar texture selectivity and temporal dynamics and to dynamically decode texture feature information. Our results are the first to demonstrate the temporal dynamics of V4 neuronal responses involved in the emergence of a cortical surface-texture encoding along perceptually defined axes.

Materials and Methods

The results presented here are based on new analyses of data previously published in Kim et al. (2019). Methods related to animal preparation, data collection, and visual stimuli are identical to those in Kim et al. (2019) but are included here for completeness.

Animal preparation

Two healthy adult macaque monkeys (1 male and 1 female) were used in the study. All animal procedures conformed to National Institutes of Health guidelines and were approved by the Institutional Animal Care and Use Committee at the University of Washington. Animals were surgically implanted with custom-built head posts attached to the skull with orthopedic screws. A V4 recording chamber was placed over the left pre-lunate gyrus on the basis of structural MRI scans. A craniotomy was performed in a subsequent surgery, a few days before the first recording date.

Animals were seated in front of a CRT monitor (at a distance of 57 cm) and were trained to hold their gaze within 1° of a small central fixation spot (0.1° diameter), while a series of stimuli were presented in the visual periphery during a simple passive fixation task.

Data collection

Recordings were performed using an epoxy-insulated tungsten microelectrode (FHC). The microelectrode was lowered into cortex through an 8-channel acute Microdrive system (Gray Matter Research). Signals were amplified, bandpass filtered (150 Hz and 8 kHz) and digitized

(sampling rate 32 kHz) with a Plexon MAP system. Spike waveforms were sorted offline using principal component analysis. Time stamps of single unit spiking activity, eye positions (Eyelink 1000; SR Research), and stimulus events (verified with photodiode signal) were stored at 1 kHz sampling rate for later analysis.

Once a well-isolated single unit was identified, we determined the center of the receptive field (RF) by a hand-mapping procedure (Kim et al., 2019). This was followed by the main experiment in which a series of 4 or 5 visual stimuli were presented, each for 300 ms, separated by a 300 ms interstimulus interval, as the animals maintained fixation. Each stimulus was presented at the center of the RF of the neuron and was scaled such that all parts of all stimuli were within the estimated RF (estimated RF diameter = 1.0 + 0.625° × RF eccentricity [°]; based on Gattass et al., 1988). We studied the responses of 127 well-isolated V4 neurons with RF eccentricities within the central 10° of the visual field (mean ± SD: 5.30 ± 1.48° for Monkey 1; 5.09 ± 1.47° for Monkey 2). For all neurons, we collected a minimum of 6 repetitions for each stimulus condition.

Visual stimuli

Three types of stimuli were used in the original study: shape, texture, and natural scenes. Details of each stimulus type are described in Kim et al. (2019), where we analyzed the relationship between shape selectivity and texture selectivity. Here, we focus on the responses to texture stimuli alone.

We used a set of 21 textures (Fig. 1) to examine how V4 responses encode four texture dimensions (coarseness, directionality, regularity, and contrast) thought to be critical for perception of textures in human subjects (Tamura et al., 1978; Liu and Picard, 1996; Rao and Lohse, 1996).

We devised simple metrics to characterize each texture in terms of each of the four dimensions. Coarseness and directionality indices were computed from the power spectrum, S , of the two-dimensional Fourier transform of an image, $S(sf, \theta) = |F(sf, \theta)|^2$, where sf and θ represent the SF and angle in polar coordinates. The power spectrum was first low-pass filtered to cut off frequencies higher than 60 cycles per image (~10 cycles/°), then normalized so that the sum of all values was 1. The coarseness index was computed by subtracting the sum of sf -weighted power magnitude from 60. The coarseness index ranges from 0 to 60, with the higher values representing the coarser textures (Fig. 1A). To quantify directionality, we calculated the average power magnitude from eight evenly divided ranges of angle (180°/8 = 22.5°). The sum of average magnitude in the top two preferred angular bins divided by that in all other bins was taken as the directionality index, which ranged from 0.25 to 1.00 (Fig. 1B).

The regularity index was devised to quantify the repetitive nature of a texture image. This feature was captured from the two-dimensional autocorrelation map, ρ , of an image described as follows:

$$\rho(x, y) = \sum_{m=1}^M \sum_{n=1}^N I(m, n) \times I(m + x, n + y) / \sum_{m=1}^M \sum_{n=1}^N I^2(m, n)$$

Where $I(m, n)$ denotes an $M \times N$ texture image. The map is converted from cartesian to polar coordinates as follows.

$$\rho(r, \theta) = \rho\left(\sqrt{x^2 + y^2}, \tan^{-1}\left(\frac{y}{x}\right)\right)$$

From the polar representation of the autocorrelation, we obtained one-dimensional autocorrelation functions for all possible directions in 1° steps. An autocorrelation function from a regular pattern has clear peaks and valleys. The regularity measure was given by the height of the most prominent peak in the one-dimensional autocorrelation function across all directions. The regularity index ranges from 0.04 (the most irregular) to 0.96 (the most regular) in our texture stimulus set (Fig. 1C).

The contrast index, which reflects dynamic range (σ) and polarization (α_4) of the gray level distribution across image pixels, was computed with the following equation (Tamura et al., 1978):

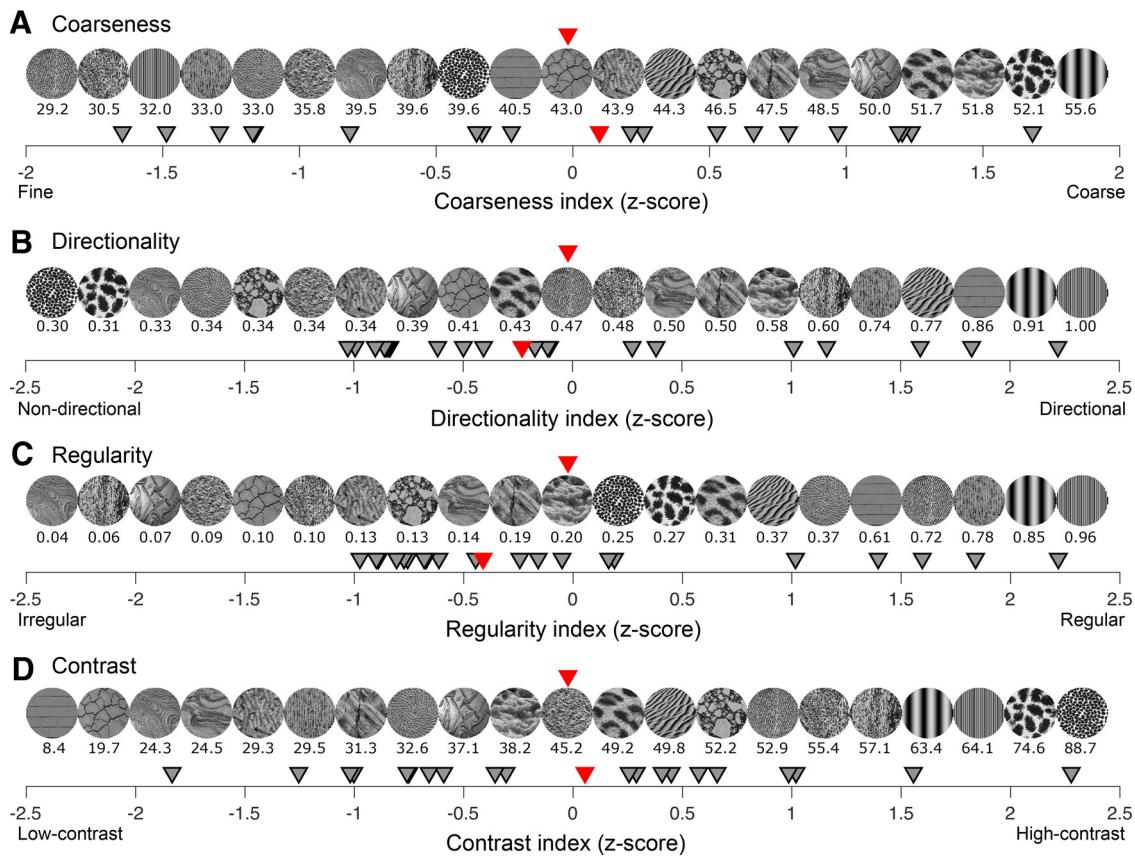


Figure 1. Visual stimuli. We used a set of 21 textures to study responses of V4 neurons. The ordering of textures along the four dimensions is shown: (A) coarseness, (B) directionality, (C) regularity, and (D) contrast. Along each axis, textures are rank-ordered; in A from fine to coarse, etc. Numbers below each texture image indicate the raw index value along each axis. Gray triangles represent the corresponding z-scored value. Red triangles represent the median texture and the corresponding z-scored value along each axis. Each texture was presented in four orientations for a total of 84 stimuli (see Materials and Methods).

$$I_{contrast} = \frac{\sigma}{\alpha_4^n}$$

where σ and α_4 represent the SD and kurtosis, respectively, of the gray-level distribution across image pixels. The exponent $n = 1/4$ was held constant. The lowest and highest contrast indices computed from our texture stimulus set were 8.4 and 88.7, respectively (Fig. 1D).

To choose a small number of texture images that spanned a range of values across these four dimensions, we represented a set of 144 monochrome texture samples, including 112 from the Brodatz album (Brodatz, 1966) and 32 from a commercial library (www.textures.com), in terms of the four texture dimensions described above, and transformed the raw scores of each feature dimension to z scores and chose a set of 21 textures. While our goal was to choose a set of textures that achieved independent sampling across the four dimensions, our stimulus set reflects inherent correlations in naturalistic textures. Across our stimuli, there is a statistically significant correlation between coarseness and regularity ($r = -0.22$, $p < 0.01$ for the 21 chosen textures; $r = -0.30$, $p < 0.01$ for the 144 textures samples) and between regularity and directionality ($r = 0.63$, $p < 0.01$ for the 21 chosen textures; $r = 0.56$, $p < 0.01$ for the 144 textures samples). For some analyses described below, we divided the textures into two categories with a cutoff at $z = 0$ along each dimension; e.g., along the regularity axis, stimuli were categorized as regular ($z > 0$) versus irregular ($z < 0$). Figure 1 shows how the 21 textures chosen for this study are represented along each of the four feature axes.

Each of the 21 textures was presented at four orientations (at 45° intervals) and at two aperture sizes (the estimated RF diameter and twice that size) for a total of 168 stimuli. The small and large aperture textures were identical within the RF, but the large textures extend into the RF surround. Here we report results based on the analysis of spiking

responses to large aperture textures ($N = 84$). Results based on responses to large and small apertures were consistent.

Data analysis

Time course of spiking responses and texture-dependent modulation. Peristimulus time histograms (PSTHs) were constructed by averaging responses across multiple repetitions and convolving with a Gaussian kernel ($\sigma = 5$ ms). To determine the timing of response modulation associated with each of the four texture dimensions, for each texture dimension (e.g., coarseness), we divided the 84 stimulus patches into two equal groups based on their texture indices (Fig. 1) and asked when responses between the two groups (e.g., coarse vs fine) significantly deviated from each other using the Mann–Whitney U test ($p < 0.05$) within a 30 ms sliding window (moving in 1 ms steps).

Average responses of single neurons. For each stimulus, we quantified the average response magnitude by counting spikes within a window from 50 to 400 ms after each stimulus onset to allow for onset and offset response latency of V4 neurons (Zamarashkina et al., 2020) and averaging across multiple repetitions of a stimulus. These average responses were used for the regression and classification analyses described below.

To identify preferred and nonpreferred textures for representative examples (see Neurons 1–4 in the Figures and Results), we assessed mean responses during chosen time periods based on the temporal dynamics of selectivity for the different texture features: an earlier window (50–150 ms) to visualize preferences for coarseness and contrast, and a later window (100–400 ms) for directionality and regularity.

Regression model for texture selectivity. To ask how the perceptual texture dimensions modulate neuronal responses, for each neuron we conducted a stepwise linear regression analysis to model responses to each texture in terms of a subset of the four texture dimensions as independent variables. Specifically,

$$Z_{\text{response}} = \beta_1 Z_{\text{Coarseness}} + \beta_2 Z_{\text{Directionality}} \\ + \beta_3 Z_{\text{Regularity}} + \beta_4 Z_{\text{Contrast}}$$

where both independent and dependent variables are normalized.

The stepwise fit, implemented with the function *stepwisefit* in MATLAB, begins with an initial constant model and takes forward steps to add a single variable that gives the best improvement to the current model with a coefficient that is significantly different from 0 at $p < 0.05$. In successive steps, coefficients may be removed from the model if they are not significantly different from 0 at $p < 0.1$.

As we described above, because of the inherent correlations in naturalistic textures, our stimulus set showed a statistically significant correlation between coarseness and regularity, and between regularity and directionality indices. In a regression context, a strong collinearity among predictor variables (e.g., a variance inflation factor > 5) may lead to poor estimation of the model parameters (O'Brien, 2007). We verified that all four texture indices had variance inflation factor < 3 , indicating that multicollinearity was not strong. We also examined model fits based on partial least squares and ridge regression methods, and obtained similar model performance results as with stepwise linear regression.

To investigate whether texture selectivity evolves over time, we performed the stepwise linear regression analysis using a 100 ms sliding window (moving in 10 ms steps).

Support vector machine (SVM) for texture classification. To assess how faithfully V4 responses encoded the perceptual texture attributes, we trained a linear SVM to classify texture stimuli into two classes along each of four texture dimensions (coarse vs fine, directional vs nondirectional, regular vs irregular, and high-contrast vs low-contrast) based on the distribution of responses across a population. Decoding performance was tested with population sizes of 10–120 neurons (in steps of 10). For a given population size, we simulated 100 random cell groups by sampling neurons with replacement (i.e., the same neuron could be sampled multiple times in a simulated cell group). The performance of the SVM classifier was evaluated with 10-fold cross-validation in each simulation, and the predictive accuracy was indicated as the proportion correct (chance level: 0.5) averaged from 100 simulation results. To compare the population decoding latencies across texture dimensions, we also performed the SVM classifier using a 30 ms sliding window (in steps of 1 ms). The latency results were consistent regardless of population sizes, and we report the result obtained from the population size of 60 neurons.

Experimental design and statistical analysis

Details of experimental procedure and visual stimuli are described above (see Data collection, Visual stimuli). For all statistical tests presented here, independent group comparisons were performed using a nonparametric Mann–Whitney *U* test, while the multiple-group comparison was done with one-way ANOVA. The strength of the linear relationship

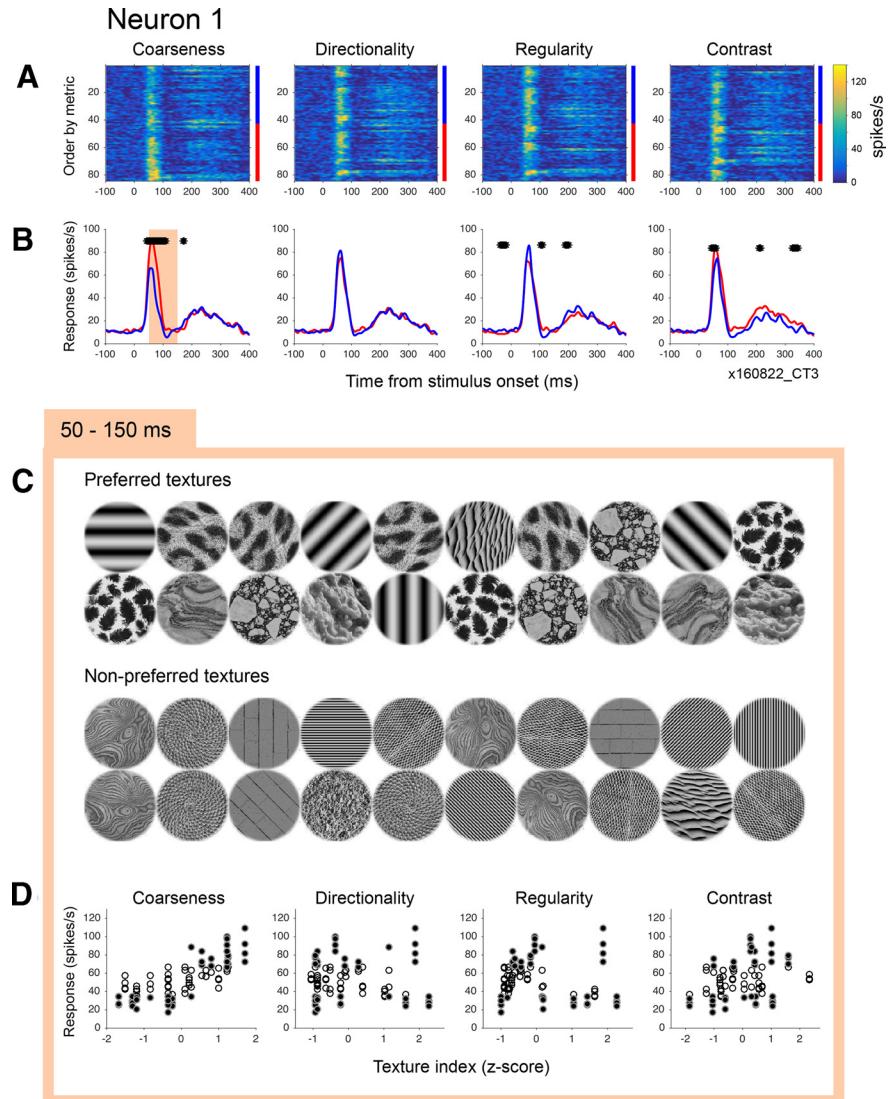


Figure 2. Example neuron selective for coarse textures. **A**, PSTHs are shown for the 84 texture stimuli. All panels represent the same data but differently ordered in accordance with the rank-ordering in Figure 1 (i.e., in ascending order of the index value along each of the four texture dimensions). From top to bottom for each panel, textures run from fine to coarse, nondirectional to directional, irregular to regular, and low-contrast to high-contrast, respectively. Color represents response strength in accordance with the scale bar. Responses are smoothed with a Gaussian of $\sigma = 5$ ms. **B**, Average PSTHs for the top and bottom halves along each texture dimension are shown in blue and red, respectively. For example, in the first panel at left, blue and red represent responses to fine and coarse textures, respectively. Black asterisks indicate time points with significant difference between red and blue curves (Mann–Whitney *U* test in a 30 ms sliding window, $p < 0.05$). Statistically significant difference between responses to coarse and fine textures emerged 41 ms after stimulus onset. **C**, The 20 most (top) and least (bottom) preferred textures based on the number of spikes during the 50–150 ms window after stimulus onset (shading in **B**) are shown. **D**, Scatter plots represent neuronal responses to all texture stimuli during the 50–150 ms window as a function of each texture index. Filled symbols represent the 20 preferred and nonpreferred textures shown in **C**.

between pairs of variables was assessed by Pearson's correlation coefficient. A p value of 0.05 was considered significant.

Data and software availability

The data and analysis code that support the findings of this study are available from the corresponding author on request.

Results

We studied the responses of 127 V4 neurons in 2 macaque monkeys (56 in Monkey 1, 71 in Monkey 2) to a set of 21 naturalistic textures, each presented at four orientations. Our stimuli (Fig. 1) sample four texture attributes (coarseness,

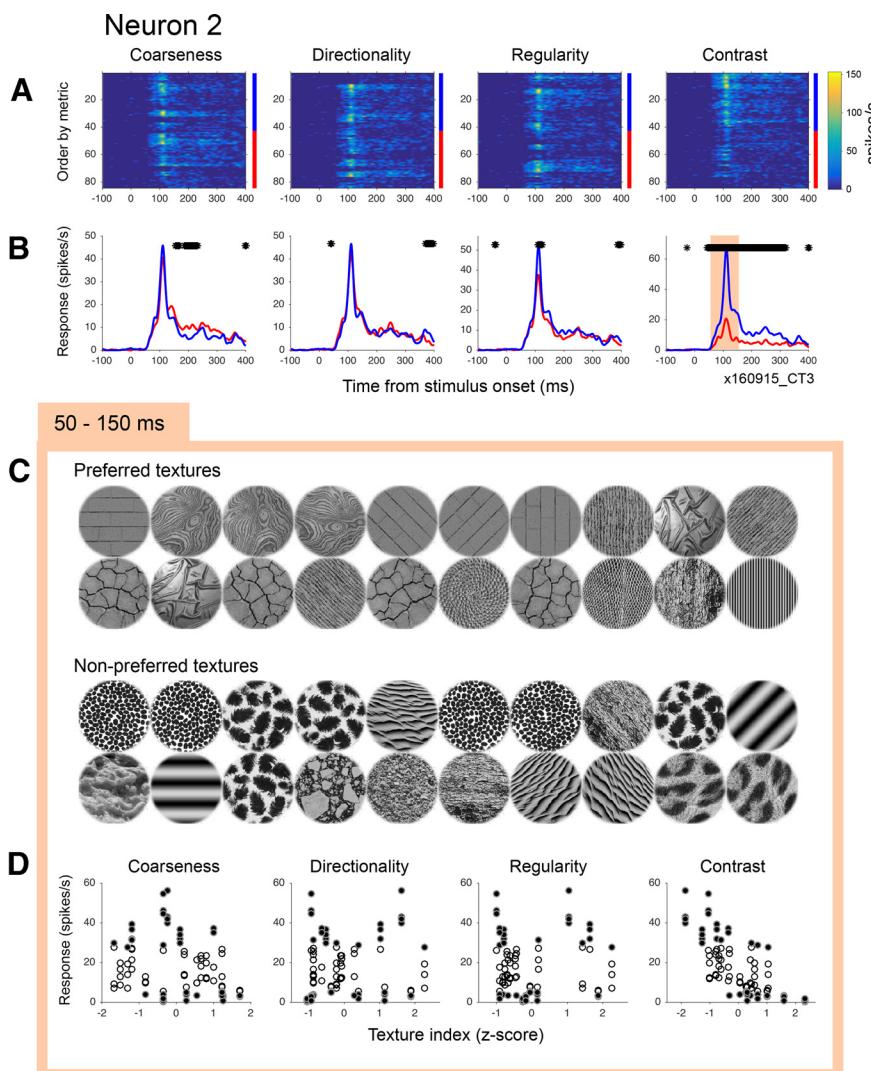


Figure 3. Example neuron selective for texture contrast. Low-contrast textures evoked stronger responses from Neuron 2 compared with high-contrast textures (**A,B**, rightmost panels). Statistically significant difference between responses to low- and high-contrast textures emerged 43 ms after stimulus onset. A mild preference for coarse textures is notable later in the response. All conventions are as in Figure 2. **C**, The 20 most and least preferred textures are shown based on the number of spikes during the 50–150 ms window after stimulus onset (shading in **B**). **D**, Scatter plots represent neuronal responses to all texture stimuli during the 50–150 ms window as a function of each texture index. Filled symbols represent the 20 preferred and nonpreferred textures shown in **C**.

directionality, regularity, and contrast) that arise from the study of human texture perception. Having previously demonstrated that the response amplitude of many V4 neurons is well modulated by at least one of these texture attributes (Kim et al., 2019), here we focus on response dynamics and describe the relative contributions of these attributes to the development of texture selectivity over time in V4.

Time course of selectivity: representative examples

To determine whether a subset of texture attributes were especially strong or early modulators of V4 responses, we rank-ordered textures based on each attribute (see Fig. 1) and divided them into two halves along each axis to facilitate a comparative analysis. Figure 2A shows the responses of example Neuron 1 to textures rank-ordered by each of the texture attributes (four columns). The averages for the top and bottom halves of the ranking are plotted in Figure 2B (red traces represent high-rank average). This neuron responded more strongly to coarse than to

fine textures. Stepwise regression analysis (see Materials and Methods) revealed a statistically significant influence of coarseness on the neuronal responses (standardized coefficient for coarseness: 0.78, $p < 0.01$ for 50–150 ms duration; also see Fig. 2D). This preference is reflected in the average PSTHs for coarser (red) and finer (blue) stimuli (Fig. 2B), which reveal that a strong and sustained difference between responses emerged ~ 50 ms after stimulus onset. The other texture attributes (directionality, regularity, and contrast) showed minimal influence on the responses of Neuron 1 (Fig. 2A,B,D, right three panels). Figure 2C shows the most preferred and nonpreferred textures based on the activity during 50–150 ms after stimulus epoch when there was a sustained difference between the responses to coarse and fine stimuli. It is noteworthy that the nonpreferred textures all appear fine compared with the preferred stimuli, consistent with results presented above.

The responses of Neuron 2 were strongly modulated by luminance contrast within the texture patch (Fig. 3): results of stepwise regression analysis based on activity in the 50–400 ms epoch revealed a statistically significant influence of contrast (standardized coefficient for contrast: -0.72 , $p < 0.01$) but not of the other three attributes (compare panels in Fig. 3D). Responses indeed decreased with increasing contrast (Fig. 3A,B,D, rightmost column), and there was a strong and sustained difference in responses between the lower-contrast stimuli (blue curve) and the higher-contrast stimuli (red). Here again, the difference emerged early (~ 50 ms after stimulus onset) and was sustained for >200 ms; it appears that the difference was present from the onset of the visual response. The 20 least preferred texture patches (Fig. 3C, bottom) varied in terms of their coarseness, directionality, and regularity, but they were all high contrast, unlike the preferred textures (Fig. 3C, top).

In striking contrast to the above examples, encoding of directionality in the responses of Neuron 3 emerged later, ~ 100 ms after stimulus onset (Fig. 4). Initially, during the first 100 ms after stimulus onset, responses to directional and nondirectional textures were quite similar (compare red and blue curves, Fig. 4B, second column), but responses diverged significantly at ~ 100 ms with stronger responses evoked by nondirectional textures. There were significant transient modulations by the other texture attributes (Fig. 4B, asterisks), but these were considerably weaker than the sustained modulation by directionality. Results from stepwise regression based on activity in the 50–400 ms time window revealed a statistically significant effect of directionality but not the other three factors (standardized coefficient for directionality: -0.69 , $p < 0.01$; Fig. 4D). It is noteworthy that the influence of directionality is unaffected by the orientation of the texture: all four orientations of a texture (i.e., vertically aligned

data points sharing the same directionality index) evoked similar responses. Figure 4C shows the 20 most preferred and least preferred textures based on activity in the 100–400 ms after stimulus onset. The non-preferred textures were all directional but, importantly, not orientation-specific.

Finally, Neuron 4 showed strong response modulation for multiple texture features (Fig. 5). Stepwise regression revealed a statistically significant dependence on coarseness and also on regularity (standardized coefficients: 0.44, −0.63, respectively, $p < 0.01$). Coarse textures generated stronger responses early, during the 50–150 ms interval after stimulus onset (Fig. 5A,B,D, leftmost column), consistent with example Neuron 1, whereas irregular textures were associated with stronger activity later in the response period (>100 ms after stimulus onset, Fig. 5A,B,F, third column). Because directionality and regularity are correlated in our stimulus set (see Materials and Methods), a strong difference in responses between nondirectional and directional textures is also evident (Fig. 5A,B, second column). Because the dynamics differ across multiple encoded attributes for this neuron, the texture preference changes over time. The most and least preferred textures based on early activity (i.e., 50–150 ms poststimulus time) reveal discrimination between coarse and fine textures and also high and low contrast (Fig. 5C). However, during a later epoch (i.e., 100–400 ms poststimulus time), the distinction lies predominantly along the regular/directional versus irregular/nondirectional axis (Fig. 5E), with a strong preference for irregularity. Nonpreferred textures based on activity in the later epoch included both coarse (low SF) and fine (high SF) gratings and high- and low-contrast stimuli (Fig. 5E).

Population results: relative contribution of different attributes

To characterize population trends in terms of strength of modulation, we used stepwise linear regression to relate responses to each of the four texture attributes. This method allowed us to identify a minimal number of factors that influence neuronal responses, given the mild correlations between independent variables present in our stimulus ensemble (see Materials and Methods). Across 127 neurons, 108 (85%; 50 from Monkey 1 and 58 from Monkey 2) were associated with a statistically significant regression fit (Fig. 6A, shaded regions). Of these, roughly equal numbers were modulated by one or two factors (43 and 42, respectively; Fig. 6B) and a smaller number by 3 factors (23). Among the four texture attributes, coarseness, which is related to SF content, was the most frequent modulator (71), and contrast was the second most frequent modulator (54) (Fig. 6C,E).

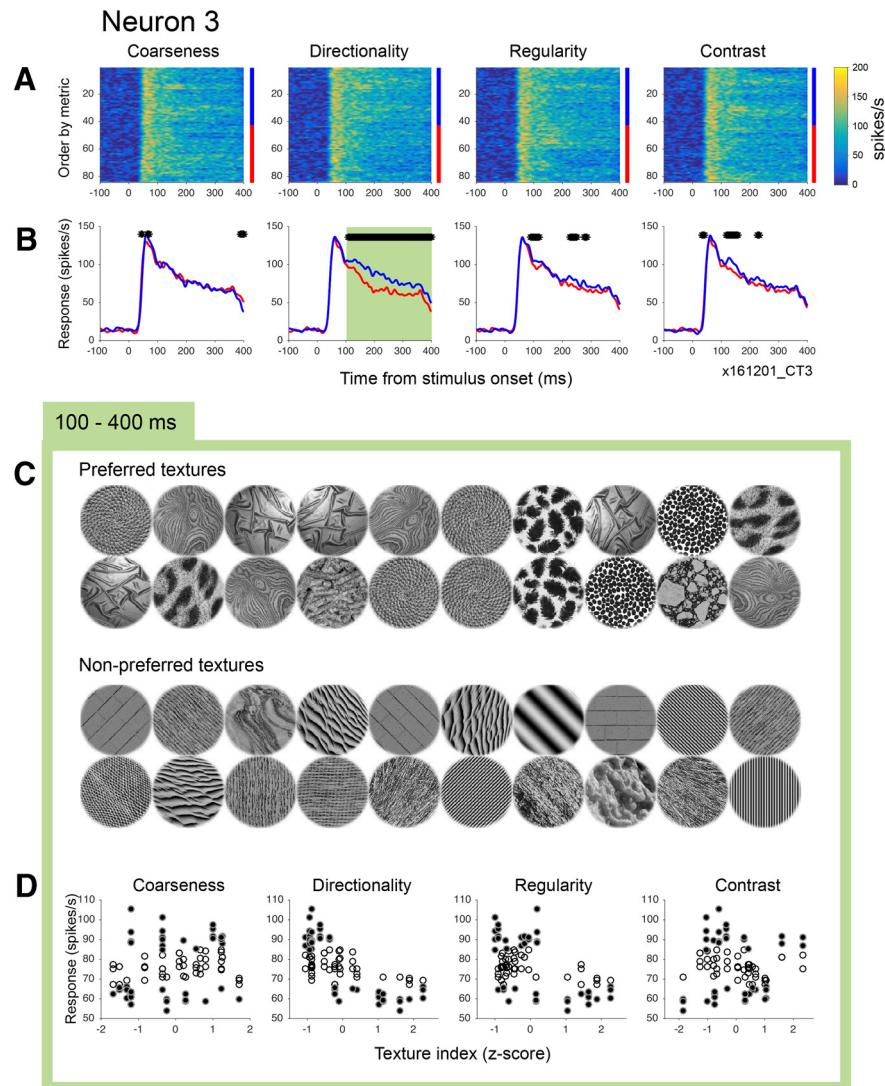


Figure 4. Example neuron selective for nondirectional textures. **A, B**, Textures lacking directional information evoked stronger responses from Neuron 3. Statistically significant difference between responses to nondirectional and directional textures emerged 109 ms after stimulus onset. **C**, The 20 most and least preferred textures are shown based on the number of spikes during the 100–400 ms window after stimulus onset (shading in **B**). All conventions are as in Figure 2. **D**, Scatter plots represent neuronal responses to all texture stimuli during the 100–400 ms window as a function of each texture index. Filled symbols represent the 20 preferred and nonpreferred textures shown in **C**.

Together, these two factors modulated responses across 39 neurons either alone or in combination with each other. The two other factors, directionality and regularity, modulated responses of 16 neurons either alone or in combination with each other. However, a large proportion of neurons (53 of 108) were modulated by a combination of the more common, earlier modulators (coarseness and/or contrast), and the less common, delayed modulators (directionality and/or regularity). We did not observe any associations between the overall responsiveness of neurons and specific texture features (one-way ANOVA, $p = 0.87$). The average firing rate (Hz, mean \pm SE, 50–400 ms poststimulus time) of neurons that showed significant regression weights for coarseness ($N = 71$) was 15.6 ± 2.1 . And those values for the other three texture features were 17.3 ± 3.9 (directionality, $N = 31$), 18.6 ± 2.9 (regularity, $N = 40$), and 16.7 ± 2.3 (contrast, $N = 54$).

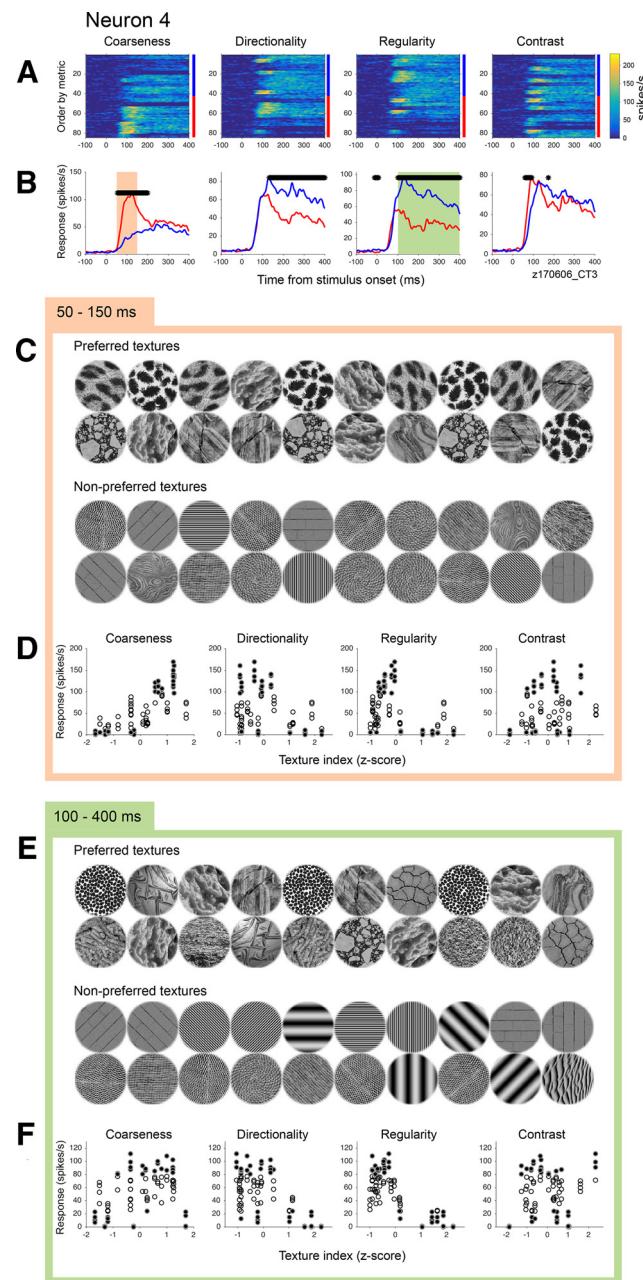


Figure 5. Example neurons selective for multiple texture features. **A, B**, Coarse and irregular textures evoke strong responses from Neuron 4. Statistically significant difference between responses to coarse and fine textures emerged 52 ms after stimulus onset, whereas that between regular and irregular textures emerged later, at 100 ms. Statistically significant difference in response between directional and nondirectional textures is also evident, but this can be explained by the correlation between directionality and regularity in our stimulus set (see Results). **C–F**, Scatter plots show mean neuronal responses to all texture stimuli during the 50–150 ms (**D**) or 100–400 ms (**F**) epoch after stimulus onset as a function of each texture index. The 20 most and least preferred textures based on early activity (**C**) and later activity (**E**) are shown. All conventions are as in Figure 2.

The overall quality of the stepwise regression model was assessed using Pearson's correlation coefficient, r . The median r value for the 108 significant stepwise regression fits was 0.50 (0.45 for Monkey 1, 0.52 for Monkey 2; Fig. 6D), indicating that our texture model could explain on average 25% of the variance in the response. Neurons with multiple significant attributes tended to be associated with better model fits (Fig. 6A, compare gray levels for upper and lower rows); mean r values for neuronal

groups with 1, 2, or 3 significant texture attributes were 0.33, 0.53, and 0.70, respectively. We also found a statistically significant relationship between the number of significant attributes and average responsiveness of neurons (one-way ANOVA, $p = 0.02$): average firing rates (Hz, mean \pm SE, 50–400 ms poststimulus time) for neuronal groups with 0, 1, 2, or 3 significant texture attributes were 7.9 ± 3.4 , 12.1 ± 2.7 , 13.5 ± 2.6 , and 23.6 ± 4.0 , respectively. These results support the idea that neurons fit with fewer attributes may be more driven, and better fit, if stimuli varied along dimensions not considered here (e.g., color).

However, as shown in the example neurons (Figs. 2–5), different texture attributes (e.g., coarseness vs regularity) modulated spiking responses during distinct temporal windows. Therefore, not considering temporal information could result in underestimating the strength of texture selectivity (see Population results: temporal dynamics).

We also evaluated relative proportions of positive and negative weights for each texture dimension (Fig. 6E). Among neurons significantly sensitive to stimulus contrast, we observed roughly equal proportion of neurons with positive and negative weights. For the other three dimensions, however, the weights were not equally distributed. Along the coarseness axis, cells favoring coarse features (55) were observed 3 times more often than cells preferring fine features (16). Along the directionality axis, directional textures were 3 times more likely to suppress than enhance responses (24 negative vs 7 positive weights). Along the regularity axis, a greater number of neurons were enhanced rather than suppressed by regularity (25 positive vs 15 negative).

Population results: temporal dynamics

To examine how texture feature selectivity evolves dynamically, for each neuron in our population, we fit the stepwise regression model across time (from -100 to 500 ms relative to the stimulus onset) for a 100 ms window sliding in 10 ms increments.

Figure 7 illustrates the evolution of selectivity for texture dimensions across time in single neurons. Selectivity for specific texture dimensions was often transient in individual neurons, but we found remarkable consistency in the polarity of regression weights across time. For example, looking down the column for coarseness in Figure 7B, regression weights often remained positive (red) or negative (blue) over the entire period of significant modulation, with only a few (6 of 127) neurons showing significant sign changes during the 50–400 ms poststimulus time. We did not find any differences across the four attributes in this regard. However, with respect to the latency of onset of selectivity, we found consistent differences between coarseness and contrast versus directionality and regularity. For coarseness and contrast, the strongest regression weights (positive or negative) emerged soon after stimulus onset, within the first 100 ms (Fig. 7B, first and fourth columns). In contrast, the strongest selectivity for directionality and regularity (Fig. 7B, second and third columns) emerged later: regression weights were either undefined or relatively weak during the 0–100 ms after stimulus onset and became stronger in the later period (>150 ms after stimulus onset).

This difference between dimensions in the dynamics of the onset and progression of selectivity is captured by simply computing the variance of regression weights, Var_{wt} , across neurons as a function of time quantified for each texture dimension (Fig. 7C). When selectivity along a specific dimension has not emerged across the population, regression weights are expected to be small and similar across neurons. When selectivity for a

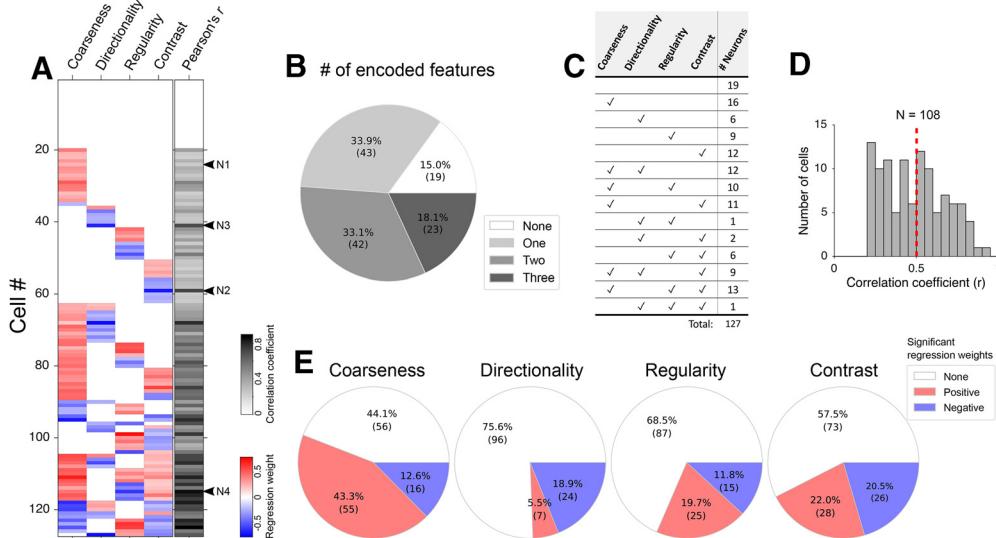


Figure 6. Population results: relative contribution of different texture attributes. **A**, Fitted weights for the four texture attributes based on the stepwise regression analysis for individual neurons are shown (see Materials and Methods). Models were based on neuronal responses in the 50–400 ms window after stimulus onset. Red and blue represent positive and negative weights, respectively (see color bar). Grayscale represents goodness of fit in terms of the correlation coefficient (Pearson's r) between the observed and predicted responses for each neuron. Neurons (N1–N4) corresponding to examples in Figure 2–5 are identified. **B**, Pie plot represents relative proportions (and numbers) of neurons encoding single or multiple texture features based on the number of coefficients deemed significant in the stepwise regression model. **C**, Tabulation of the frequency of the different combinations of feature dimensions that provided the best fit for neuronal responses based on the regression models. **D**, Distribution of goodness-of-fit values quantified by the correlation coefficient (r) between the observed and predicted data values from the 108 neurons with a statistically significant stepwise regression fit. Red dashed line indicates the median value, 0.50. **E**, The relative proportions of positive and negative weights for each texture attribute.

specific dimension emerges in the V4 population, with different neurons exhibiting different preferences, regression weights across neurons will diverge. High values of Var_{wt} are an indicator of this diversity and thus of the emergence of selectivity. The coarseness feature showed the biggest and fastest increase in Var_{wt} as a function of time, peaking (>0.1) at 100 ms after stimulus onset and then declining to a sustained level beyond 200 ms. Both regularity and directionality showed a slower increase in Var_{wt} , reaching a peak level at ~ 150 ms that was sustained until after stimulus offset at 300 ms. Finally, the Var_{wt} for contrast also increased rapidly, reaching a peak before 100 ms; but compared with coarseness, Var_{wt} was weaker in magnitude and less transient. The earlier emergence of coarseness and contrast and the later emergence of directionality and regularity can be easily visualized when peak Var_{wt} is normalized across the four dimensions (Fig. 7D).

These distinct temporal dynamics observed for different texture features were further confirmed by comparing the histograms of significantly modulated neurons across time for each of the four texture dimensions (Fig. 8). For each texture attribute, we counted the number of neurons that exhibited a statistically significant difference between the average response evoked by the two halves of the stimulus set along each axis (coarser vs finer stimuli, regular vs irregular stimuli, etc.; see Fig. 1). The number of neurons that show differential responses for coarseness increases rapidly beginning soon after stimulus onset and peaks at ~ 93 ms, where 55 of 127 units were significantly modulated for coarseness. After the peak, the frequency decreased markedly again and remained constant after 150 ms. Therefore, the dynamics of coarse versus fine processing could be characterized by a fast-transient period followed by a low sustained component.

On the other hand, the time courses of selectivity for directionality and regularity (Fig. 8B,C) showed slower and delayed

rising and falling phases compared with coarseness. The maximum frequencies were reached at 141 ms (directionality, 35 of 127 neurons) and 234 ms (regularity, 40 of 127 neurons), respectively. After the peaks, the frequencies did not decrease substantially while the stimulus was being presented (0–300 ms).

Finally, the time course of significant contrast selectivity (Fig. 8D) resembled that of coarseness selectivity. Its rising phase was very steep and as early as that for coarseness; however, its falling phase was more gradual.

In summary, V4 neurons revealed distinct temporal dynamics of texture selectivity that reflect the weighted combinations of multiple encoded features. Selectivity for coarseness and contrast arose earlier and tended to decay earlier across the population, whereas selectivity for directionality and regularity arose later, more gradually, and tended to decay less while the stimulus was ongoing.

Time course of population-based texture classification

Finally, we examined whether texture categories (e.g., coarse vs fine, directional vs nondirectional) could be decoded from V4 population activity, and we quantified how decoding accuracy improves with population size. To do this, we simulated population responses by randomly sampling tens of V4 neurons (from 10 to 120, in steps of 10) and fed these texture responses into SVM classifiers with binary categorical classes assigned for each of four texture dimensions (see Materials and Methods). We found that the proportion correct of the classifier for each texture attribute increased gradually with population size and appeared to reach an asymptote for population sizes on the order of 100 neurons (Fig. 9A). The decoding accuracies for regularity and coarseness (yellow and blue lines) were higher than those for directionality and contrast (red and purple), consistent with the ordering of variance values in Figure 7C.

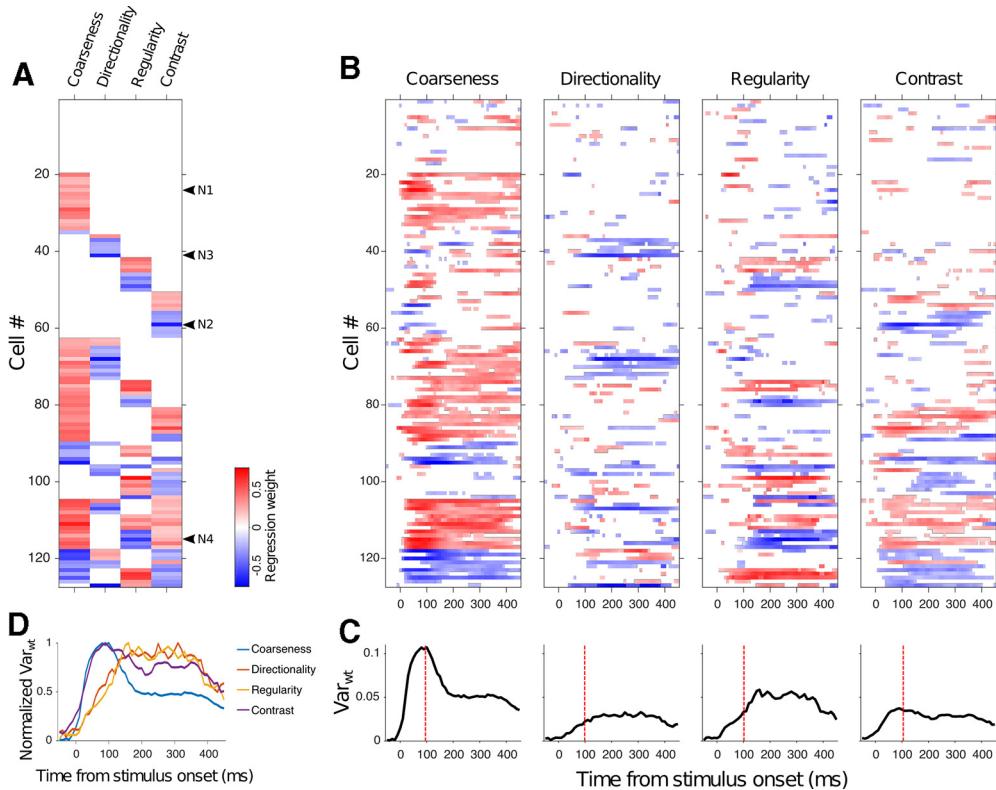


Figure 7. Population results: temporal dynamics of the stepwise regression fit. **A**, The fitted weights for the four attributes based on the stepwise regression model fit for individual neurons. Models were based on neuronal responses in the 50–400 ms window after stimulus onset (same as in Fig. 6A). **B**, Weights as a function of time for the four texture attributes based on the stepwise regression fit. Models were based on neuronal responses within a 100 ms sliding window from −100 to 500 ms sliding in 10 ms increments. Red and blue represent positive and negative weights, respectively. **C**, Variance in texture weights across neurons (Var_{wt}) is shown as a function of time for each of the four texture dimensions. High values of Var_{wt} are an indicator of the emergence of selectivity. **D**, Var_{wt} , normalized by the maximum across texture attributes, is plotted to facilitate direct comparison of the temporal dynamics. Var_{wt} for coarseness and contrast rose rapidly and reached the maximum values no later than 100 ms after stimulus onset, whereas those for directionality and regularity evolved more slowly and reached the maximum values at ∼150 ms after stimulus onset.

To examine how population decoding performance evolves over time for each texture dimension, we trained SVM classifiers with populations of 60 cells based on neuronal responses averaged within a 30 ms sliding window. The results were averaged across 100 simulation runs (Fig. 9B). The performance trajectory for coarseness rises and reaches high levels (>0.7) fastest compared with the other three dimensions. Consistent with Figure 9A, peak decoding performance for regularity and coarseness was higher than the other two dimensions and the peak performances were achieved at 117 ± 4.4 (mean \pm SE) ms for coarseness, at 128 ± 3.7 ms for contrast, 164 ± 4.7 ms for regularity, and 189 ± 5.3 ms for directionality. The latency differences across features were statistically significant (Mann–Whitney U test, $p < 0.01$). But more importantly, this analysis demonstrates that texture category information can be reliably decoded within a short temporal window based on responses of a relatively small neuronal population.

Discussion

We investigated the relationship between perceptual texture attributes and the dynamics of single-neuron texture encoding in V4. We found that different texture dimensions were associated with different temporal dynamics: selectivity for coarseness and contrast emerged early, whereas the encoding of directionality and regularity was significantly delayed, emerging ∼50 ms later. Here we relate our results to previous findings on the latency of encoding of surface and form characteristics in visual cortex and discuss potential underlying mechanisms.

Cortical mechanisms for computing perceptual dimensions of texture

Among texture attributes, coarseness and contrast encoding may directly derive from SF and contrast encoding in upstream areas. Numerous studies document encoding of contrast in LGN and contrast and SF in V1 (Schiller et al., 1976; De Valois et al., 1982; Shapley and Lennie, 1985), with selectivity for these surface characteristics emerging within tens of milliseconds of response onset (Bredfeldt and Ringach, 2002; Mazer et al., 2002). In particular, low contrast and high SF are associated with longer latencies and slower integration times in V1 (e.g., Gawne, 2000; Bair and Movshon, 2004). This differential latency could explain the early, rapid emergence of V4 selectivity for contrast and coarseness, and its frequency, compared with directionality and regularity.

Computing directionality and regularity, on the other hand, may rely on extracting higher-order image statistics. The computation of directionality, reflecting the dominance of a restricted range of orientation, requires a comparison across orientation channels and may involve integration over SFs and positions. Similarly, regularity reflects repetitive patterns, and its calculation may depend on the simultaneous activation of multiple neurons that share similar SF and/or orientation tuning but have disparate RF locations. These computations may involve more elaborate processing that could require more time. In this context, our results are consistent with prior studies in V2 demonstrating delayed emergence of selectivity for higher-order texture statistics (Freeman et al., 2013; Okazawa et al., 2017), the delayed emergence of selectivity for stimulus symmetry (or regularity) versus stimulus

contrast (Norcia et al., 2002; Jacobsen and Höfel, 2003; Kohler et al., 2018), and the delayed emergence of pattern versus component direction selectivity in MT neurons (Smith et al., 2005).

Recent findings from fMRI studies demonstrate that rotational symmetries are encoded in V3 and V4, but not earlier (Kohler et al., 2016; Audurier et al., 2021). Thus, while selectivity for higher-order image statistics may begin to arise upstream of V4 (e.g., V2), the encoding of regularity may rely on computations in V3, V4, and beyond.

Overall, the temporal organization of texture processing revealed by our results presents a novel framework for parameterizing texture in terms of when different aspects are processed in midlevel visual cortex based on features encoded in upstream stages. The V4 dynamics characterized here can also help interpret temporal evolution of texture perception that could be revealed in future studies. Our results also support the temporal primacy of form processing over texture. V4 selectivity for directionality and regularity is considerably delayed compared with the emergence of shape selectivity demonstrated in prior studies (Bushnell et al., 2011; Kim et al., 2019). Selectivity for coarseness and contrast across the V4 population also appears to emerge more slowly than shape, but this would need to be directly compared in further studies. These results may provide the neurophysiological basis for psychophysical observations that rapid animal detection in natural scenes is based on form rather than surface cues (Elder and Velislavljević, 2009). Our results are also consistent with models of coarse-to-fine processing that postulate faster processing of information that carries the “gist” of the scene and slower processing of spatial detail (Oliva, 2005; Allen and Freeman, 2006; Hegdé, 2008).

Across the population of V4 neurons, we found a preferential encoding of coarse, nondirectional textures (Fig. 6). Our coarseness metric directly relates to SF, and past studies report preferred SFs of <1 cycle/° in V4 (Lu et al., 2018). In our study, the median coarseness corresponded to 2–3 cycles/°, and this could explain the preference for coarser textures in our stimulus set. It is conceivable that directional textures are more likely to suppress than facilitate neuronal activity because multiple parallel contours may cause iso-orientation surround suppression in V1 (Knierim and Van Essen, 1992; Bair et al., 2003). Such iso-orientation surround suppression is hypothesized to enhance the representation of object boundaries by suppressing the encoding of uniform texture (i.e., de-texturization) (Gheorghiu et al., 2014). Because our stimulus set had a sizable positive correlation between directionality and regularity (i.e., directional textures tended to be regular), partitioning the influence of these two parameters is not entirely feasible. Taking these two dimensions together, there were roughly similar levels of encoding of nondirectional/irregular and

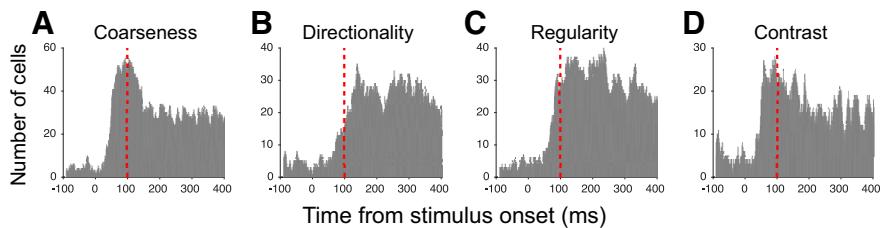


Figure 8. Temporal dynamics: the frequency of significantly modulated neurons. **A–D**, The number of significantly modulated neurons as a function of time for each texture dimension. Red lines at 100 ms after the stimulus onset are included to facilitate comparison across panels.

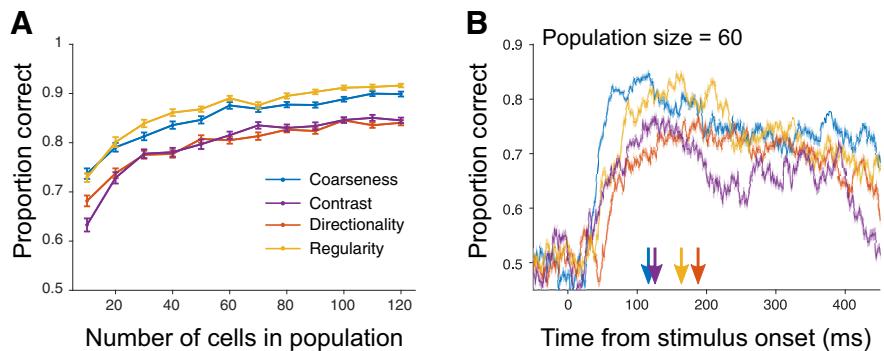


Figure 9. Decoding of texture category from population responses. **A**, Decoding accuracy of SVM classifiers for each of the four texture dimensions is plotted as a function of population size. At each size, neurons were sampled with replacement to generate a simulated subpopulation, and an SVM was trained to assign each texture to one of two categories along each texture dimension (e.g., coarse vs fine) based on spiking responses during the 50–400 ms window after stimulus onset. The simulation was repeated 100 times, and the average cross-validation scores were obtained for each texture dimension. Error bars indicate SEM. **B**, Time course of SVM classifier performance was quantified using a sliding window (bin width: 30 ms, step size: 1 ms) for a population size of 60 neurons. The simulation was repeated 100 times, and the average cross-validation scores were obtained for each texture dimension. Shaded area represents ± 1 SEM. Arrows indicate the peak proportion correct for each curve.

directional/regular textures. We also found similar levels of preference for high- and low-contrast textures, which is surprising given the stronger, faster V1 responses to higher contrast stimuli (Carandini and Heeger, 1994; Reich et al., 2001). But importantly, our contrast metric measures the variance and skewness in texture pixel values rather than luminance contrast against a background.

Limitations and future studies

Our stimuli were limited in number (being only part of a larger shape and texture study) (Kim et al., 2019), and in the independence of dimensions. Future experiments with larger sets of artificial textures could provide greater independent control of dimensions, especially regularity and directionality.

Our regression model using four key perceptual parameters explained $\sim 25\%$ of response variance. Thus, other dimensions of texture or nontextural factors (e.g., orientation of the texture elements) no doubt influence V4 responses and should be investigated in future studies. Nevertheless, our four-parameter model achieved levels of explained variance similar to the 29-parameter texture model of Okazawa et al. (2015) and the contour-based shape models of V4 with similarly few parameters (Pasupathy and Connor, 2001).

Although we found a clear relationship between perceptual dimensions of texture and V4 responses, we have not examined color, which greatly influences perception. Past studies have reported that chromatic features are processed more slowly than achromatic ones (Kelly, 1983; Burr et al., 1998; Cottaris and De

Valois, 1998). Area V4 contains a class of cells that respond most strongly to stimuli defined by chromatic contrast alone (equiluminant to their background), but these responses emerge later than those to high luminance contrast stimuli (Bushnell et al., 2011), similar to texture selectivity. A broader analysis of dynamic processing across color, texture, and form could reveal how the encoding of visual scenes with a rich array of attributes unfolds in midlevel stages.

Our experiments were conducted in fixating animals and do not indicate whether the perceptual dimensions were relevant to the monkey nor whether the dynamics of texture perception correlate with those of V4 encoding. Studies that integrate monkey perceptual judgments (e.g., texture discrimination along the various texture dimensions) with neurophysiology are needed to address these questions.

The diverse response dynamics of texture-selective neurons suggest that various cortical mechanisms may be involved in texture processing, and the influence of these various mechanisms may be integrated with different weights across neurons. An important goal for future study will be to characterize these distinct response dynamics as a function of cortical layer. In primate visual cortex, feedforward inputs are thought to terminate primarily in layer 4, and feedback connections outside of layer 4 (Felleman and Van Essen, 1991; Callaway, 2004). Superficial and deep layers, considered as the output stage to higher-order areas and feedback projections, respectively, also have slowly conducting long-range horizontal intracortical connections (Yoshioka et al., 1992; Markov et al., 2014) interconnecting functionally similar clusters of cells (Grinvald et al., 1994; Bosking et al., 1997). If delayed processing of directionality and regularity is mediated by slow horizontal connections, neurons tuned to these higher-order features may be rare in layer 4.

Modeling can also be used to provide insight into potential mechanisms of texture processing. Deep convolutional neural networks (DCNNs) trained for image-based object recognition include units that can capture and predict response properties that gradually change along the visual hierarchy. For example, DCNN units in intermediate (but not early) layers exhibit qualities found in V2/V4 neurons, such as tuning for boundary curvature, translation invariance, and sensitivity to naturalistic textures (Pospisil et al., 2018; Laskar et al., 2020). However, the DCNNs lack dynamics; thus, temporal aspects of neuronal responses were ignored. Future studies of spatiotemporal encoding within dynamic DCNNs trained on video sequences may address this limitation and provide a new tool for exploring the relationship between dynamics and visual encoding.

In conclusion, our results show that V4 neuronal response to naturalistic textures can be modeled by a weighted linear combination of multiple perceptual dimensions. Response modulation induced by directionality and regularity was significantly delayed relative to that induced by coarseness and contrast. This supports the hypothesis that the processing of higher-order texture attributes may be mediated by slower, non-feedforward connections. Future studies will reveal whether texture selectivity differs depending on anatomic factors, such as cortical layers, columns, etc.

References

- Allen EA, Freeman RD (2006) Dynamic spatial processing originates in early visual pathways. *J Neurosci* 26:11763–11774.
- Audurier P, Héja-Brichard Y, De Castro V, Kohler PJ, Norcia AM, Durand JB, Cottereau BR (2021) Symmetry processing in the macaque visual cortex. *Cereb Cortex*. Advance online publication. Retrieved Oct 6, 2021. doi: 10.1093/cercor/bhab358.
- Bair W, Movshon JA (2004) Adaptive temporal integration of motion in direction-selective neurons in macaque visual cortex. *J Neurosci* 24:7305–7323.
- 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.
- Bosking WH, Zhang Y, Schofield B, Fitzpatrick D (1997) Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *J Neurosci* 17:2112–2127.
- Bredfeldt CE, Ringach DL (2002) Dynamics of spatial frequency tuning in macaque V1. *J Neurosci* 22:1976–1984.
- Brodatz P (1966) Textures: a photographic album for artists and designers. Mineola, NY: Dover.
- Burr DC, Fiorentini A, Morrone C (1998) Reaction time to motion onset of luminance and chromatic gratings is determined by perceived speed. *Vision Res* 38:3681–3690.
- Bushnell BN, Harding PJ, Kosai Y, Bair W, Pasupathy A (2011) Equiluminance cells in visual cortical area V4. *J Neurosci* 31:12398–12412.
- Callaway EM (2004) Feedforward, feedback and inhibitory connections in primate visual cortex. *Neural Netw* 17:625–632.
- Carandini M, Heeger DJ (1994) Summation and division by neurons in primate visual cortex. *Science* 264:1333–1336.
- Cottaris NP, De Valois RL (1998) Temporal dynamics of chromatic tuning in macaque primary visual cortex. *Nature* 395:896–900.
- De Valois RL, Albrecht DG, Thorell LG (1982) Spatial frequency selectivity of cells in macaque visual cortex. *Vision Res* 22:545–559.
- Elder JH, Velisavljević L (2009) Cue dynamics underlying rapid detection of animals in natural scenes. *J Vis* 9:7–20.
- El-Shamayleh Y, Movshon JA (2011) Neuronal responses to texture-defined form in macaque visual area V2. *J Neurosci* 31:8543–8555.
- Felleman DJ, Van Essen DC (1991) Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex* 1:1–47.
- Freeman J, Ziembra CM, Heeger DJ, Simoncelli EP, Movshon JA (2013) A functional and perceptual signature of the second visual area in primates. *Nat Neurosci* 16:974–981.
- Gattass R, Sousa AP, Gross CG (1988) Visuotopic organization and extent of V3 and V4 of the macaque. *J Neurosci* 8:1831–1845.
- Gawne TJ (2000) The simultaneous coding of orientation and contrast in the responses of V1 complex cells. *Exp Brain Res* 133:293–302.
- Gheorghiu E, Kingdom FA, Petkov N (2014) Contextual modulation as de-texturizer. *Vision Res* 104:12–23.
- Grinvald A, Lieke EE, Frostig RD, Hildesheim R (1994) Cortical point-spread function and long-range lateral interactions revealed by real-time optical imaging of macaque monkey primary visual cortex. *J Neurosci* 14:2545–2568.
- Hegdé J (2008) Time course of visual perception: coarse-to-fine processing and beyond. *Prog Neurobiol* 84:405–439.
- Jacobsen T, Höfel L (2003) Descriptive and evaluative judgment processes: behavioral and electrophysiological indices of processing symmetry and aesthetics. *Cogn Affect Behav Neurosci* 3:289–299.
- Kastner S, De Weerd P, Ungerleider LG (2000) Texture segregation in the human visual cortex: a functional MRI study. *J Neurophysiol* 83:2453–2457.
- Kelly DH (1983) Spatiotemporal variation of chromatic and achromatic contrast thresholds. *J Opt Soc Am* 73:742–750.
- Kim T, Bair W, Pasupathy A (2019) Neural coding for shape and texture in macaque area V4. *J Neurosci* 39:4760–4774.
- 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.
- Kohler PJ, Clarke A, Yakovleva A, Liu Y, Norcia AM (2016) Representation of maximally regular textures in human visual cortex. *J Neurosci* 36:714–729.
- Kohler PJ, Cottereau BR, Norcia AM (2018) Dynamics of perceptual decisions about symmetry in visual cortex. *Neuroimage* 167:316–330.
- Lamme VA (1995) The neurophysiology of figure-ground segregation in primary visual cortex. *J Neurosci* 15:1605–1615.
- Laskar MN, Giraldo LG, Schwartz O (2020) Deep neural networks capture texture sensitivity in V2. *J Vis* 20:1–23.

- Liu F, Picard RW (1996) Periodicity, directionality, and randomness: Wold features for image modeling and retrieval. *IEEE Trans Pattern Anal Machine Intell* 18:722–733.
- Lu Y, Yin J, Chen Z, Gong H, Liu Y, Qian L, Li X, Liu R, Andolina IM, Wang W (2018) Revealing detail along the visual hierarchy: neural clustering preserves acuity from V1 to V4. *Neuron* 98:417–428.
- Markov NT, Vezoli J, Chameau P, Falchier A, Quilodran R, Huissoon C, Lamy C, Misery P, Giroud P, Ullman S, Barone P, Dehay C, Knoblauch K, Kennedy H (2014) Anatomy of hierarchy: feedforward and feedback pathways in macaque visual cortex. *J Comp Neurol* 522:225–259.
- Mazer JA, Vinje WE, McDermott J, Schiller PH, Gallant JL (2002) Spatial frequency and orientation tuning dynamics in area V1. *Proc Natl Acad Sci USA* 99:1645–1650.
- Merigan WH (2000) Cortical area V4 is critical for certain texture discriminations, but this effect is not dependent on attention. *Vis Neurosci* 17:949–958.
- Norcia AM, Candy TR, Pettet MW, Vildavski VY, Tyler CW (2002) Temporal dynamics of the human response to symmetry. *J Vis* 2:132–139.
- O'Brien RM (2007) A caution regarding rules of thumb for variance inflation factors. *Qual Quant* 41:673–690.
- Okazawa G, Tajima S, Komatsu H (2015) Image statistics underlying natural texture selectivity of neurons in macaque V4. *Proc Natl Acad Sci USA* 112:E351–E360.
- Okazawa G, Tajima S, Komatsu H (2017) Gradual development of visual texture-selective properties between macaque areas V2 and V4. *Cereb Cortex* 27:4867–4880.
- Oliva A (2005) Gist of the scene. In: *Neurobiology of attention*, pp 251–256.
- Pasupathy A, Connor CE (2001) Shape representation in area V4: position-specific tuning for boundary conformation. *J Neurophysiol* 86:2505–2519.
- Portilla J, Simoncelli EP (2000) Parametric texture model based on joint statistics of complex wavelet coefficients. *Int J Comput Vis* 12:1338–1351.
- Pospisil DA, Pasupathy A, Bair W (2018) ‘Artiphysiology’ reveals V4-like shape tuning in a deep network trained for image classification. *Elife* 7:e38242.
- Rao AR, Lohse GL (1996) Towards a texture naming system: identifying relevant dimensions of texture. *Vision Res* 36:1649–1669.
- Reich DS, Mechler F, Victor JD (2001) Temporal coding of contrast in primary visual cortex: when, what, and why. *J Neurophysiol* 85:1039–1050.
- Schiller PH, Finlay BL, Volman SF (1976) Quantitative studies of single cell properties in monkey striate cortex: III. Spatial frequency. *J Neurophysiol* 39:1334–1351.
- Shapley R, Lennie P (1985) Spatial frequency analysis in the visual system. *Annu Rev Neurosci* 8:547–581.
- Smith MA, Majaj NJ, Movshon JA (2005) Dynamics of motion signaling by neurons in macaque area MT. *Nat Neurosci* 8:220–228.
- Tamura H, Mori S, Yamawaki T (1978) Textural features corresponding to visual perception. *IEEE Trans Syst Man Cybern* 8:460–473.
- Yoshioka T, Levitt JB, Lund JS (1992) Intrinsic lattice connections of macaque monkey visual cortical area V4. *J Neurosci* 12:2785–2802.
- Zamarashkina P, Popovkina D, Pasupathy A (2020) Timing of response onset and offset in macaque V4: stimulus and task dependence. *J Neurophysiol* 123:2311–2325.
- Ziemba CM, Freeman J, Movshon JA, Simoncelli EP (2016) Selectivity and tolerance for visual texture in macaque V2. *Proc Natl Acad Sci USA* 113:E3140–E3149.