Selectivity and tolerance for visual texture in macaque V2

Corey M. Ziemba[a,1,2], Jeremy Freeman[a,b,1], J. Anthony Movshon[a], and Eero P. Simoncelli[a,c]

[a]Center for Neural Science, New York University, New York, NY 10003; [b]Howard Hughes Medical Institute, Janelia Research Campus, Ashburn, VA 20147; and [c]Howard Hughes Medical Institute, New York University, New York, NY 10003

Edited by Wilson S. Geisler, The University of Texas at Austin, Austin, TX, and approved April 13, 2016 (received for review June 4, 2015)

As information propagates along the ventral visual hierarchy, neuronal responses become both more specific for particular image features and more tolerant of image transformations that preserve those features. Here, we present evidence that neurons in area V2 are selective for local statistics that occur in natural visual textures, and tolerant of manipulations that preserve these statistics. Texture stimuli were generated by sampling from a statistical model, with parameters chosen to match the parameters of a set of visually distinct natural texture images. Stimuli generated with the same statistics are perceptually similar to each other despite differences, arising from the sampling process, in the precise spatial location of features. We assessed the accuracy with which these textures could be classified based on the responses of V1 and V2 neurons recorded individually in anesthetized macaque monkeys. We also assessed the accuracy with which particular samples could be identified, relative to other statistically matched samples. For populations of up to 100 cells, V1 neurons supported better performance in the sample identification task, whereas V2 neurons exhibited better performance in texture classification. Relative to V1, the responses of V2 show greater selectivity and tolerance for the representation of texture statistics.

vision | primary visual cortex | macaque | texture perception | extrastriate visual cortex

Visual perception in primates arises from the responses of neurons in a variety of areas within the cerebral cortex. These responses are typically characterized by measuring selectivity for specific visual attributes, such as light intensity or color, and local structural properties, such as spatial position, orientation, and spatial frequency. Stimulus selectivity, along with the complementary notion of "invariance" or "tolerance" to irrelevant variation, provides a de facto language for describing the functional roles and relationships of neurons in visual areas. For example, simple cells in the primary visual cortex, area V1, are selective for orientation (1) and spatial frequency (2–4). Complex cells exhibit similar selectivity, but are also more tolerant to changes in spatial position (1, 5, 6). Component cells in area MT (or V5) exhibit selectivity for orientation and speed, but (relative to their V1 inputs) are more tolerant of changes in location and spatial frequency, whereas MT pattern cells are tolerant to changes in orientation (and, more generally, spatial structure) (7).

Neurons in the inferotemporal visual cortex (IT) are selective for visual images of particular objects, but are tolerant to identity-preserving transformations, such as translation, rotation, or background context (8, 9). This tolerance increases from area V4 to IT (10), suggesting that an increase in selectivity is balanced by an increase in tolerance, preserving overall response levels and their distribution across neurons (11). However, the selectivity and tolerance of visual representations in midventral areas, particularly area V2, have been more difficult to establish because we lack knowledge of the relevant visual attributes. V2 neurons receive much of their afferent drive from V1, have receptive fields that are roughly twice the size of the receptive fields in V1, and exhibit similar selectivity for orientation and spatial frequency (12, 13). Indeed, the responses of V2 neurons to many forms of artificial stimuli, including gratings, curves, and texture-defined patterns, are only modestly different from the responses of neurons in V1 (14–17).

Recent work suggests that local statistical measurements that capture the appearance of visual textures might provide a feature space for characterizing the responses of V2 neurons (18–20). Sensitivity to multipoint correlations in arrays of binary (black and white) pixels first arises in V2 (20), and is strongest for those correlations that are most informative about binarized natural images (21) and most perceptually salient (22). This sensitivity to higher order correlations is also present for more naturalistic stimuli. Images of natural visual texture evoke correlated responses in rectified V1-like filters tuned for differing orientation, scale, and position (23). V2 neurons are well driven by synthetic texture stimuli containing these naturally occurring correlations, and less so by texture stimuli that lack them (19). Moreover, the performance of human observers in detecting these correlations is predicted by the differential increase in average V2 response levels (19). All of these results provide evidence that area V2 plays a role in representing the higher order statistics of visual textures. However, the ways in which this representation supports visual tasks, such as discrimination, have yet to be explored.

Here, we provide a more direct test of the link between V2 and the representation of the higher order statistics of natural textures. We generated stimuli that are matched to the statistics of naturally occurring homogeneous texture images. These stimuli are perceptually similar to one another, and similar to the original texture image, despite marked differences in the position and detailed arrangement of their local features (23–25). This property can be used to generate pronounced distortions in peripheral viewing that remain imperceptible so long as the distortions preserve texture statistics over spatial regions the size of V2 receptive fields (18). If V2 is encoding these local statistics,

Significance

The brain generates increasingly complex representations of the visual world to recognize objects, to form new memories, and to organize visual behavior. Relatively simple signals in the retina are transformed through a cascade of neural computations into highly complex responses in visual cortical areas deep in the temporal lobe. The representations of visual signals in areas that lie in the middle of this cascade remain poorly understood, yet they are critical to understanding how the cascade operates. Here, we demonstrate changes in the representation of visual information from area V1 to V2, and show how these changes extract and represent information about the local statistical features of visual images.

Author contributions: C.M.Z., J.F., J.A.M., and E.P.S. designed the experiments; C.M.Z. and J.F. performed the experiments and analysis; C.M.Z., J.F., J.A.M., and E.P.S. interpreted the results; and C.M.Z., J.F., J.A.M., and E.P.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

[1]C.M.Z. and J.F. contributed equally to this work.

[2]To whom correspondence should be addressed. Email: ziemba@cns.nyu.edu.

and is responsible for these perceptual phenomena, then the responses of populations of V2 neurons to statistically matched stimuli should reveal a particular form of tolerance. Specifically, populations of neurons in V2 should respond similarly to stimuli that are statistically matched, despite variation in local image detail. This kind of tolerance would complement previously reported tolerances to geometric image transformations, such as translation or rotation, found at higher levels of visual cortex (8–10).

We studied this tolerance to statistical resampling by analyzing responses of a collection of V1 and V2 neurons to images of synthetic texture, generated to match the statistics of different texture "families." V2 responses across families of statistically matched stimuli were more homogeneous than V1 responses, reflecting an increased tolerance that was only partly explained by the larger size of their receptive fields. Using a neural population decoder, we found V2 was better than V1 at discriminating between-family images matched for different statistics and worse at discriminating within-family images matched for the same statistics, a pattern of performance that broadly resembles human perceptual experience (23, 25).

Results

Generation of Naturalistic Texture Stimuli. We studied the population representation of visual information in areas V1 and V2 using naturalistic images generated from a texture model defined in terms of joint and marginal statistics of a simulated population of V1 simple and complex cells (23). These statistics include local correlations between the output of pairs of model neurons that differ in preferred spatial frequency, position, and/or orientation. Some of these correlations are second-order statistics that capture the amount of energy at specific orientations and spatial frequencies; we refer to these statistics as "spectral." Other correlations are of higher order, capturing naturalistic features beyond the power spectrum. We first computed this set of statistics for a grayscale photograph of a natural texture, and then generated synthetic texture images by starting with an image of Gaussian white noise and iteratively adjusting the pixels until the image had the same statistics (computed over the entire extent of the synthesized image) as the original photograph (23).

We refer to a set of images with identical statistics as a texture "family" (Fig. 1*A*, columns). Within a family, different white noise seeds yield different images, and we refer to all such images as "samples" from that family (Fig. 1*A*, rows). By construction, samples are identical in their model statistics, but differ in the location and arrangement of features within the image. Previous work (23, 24) and visual inspection of Fig. 1*A* reveals that samples from a given family are similar in appearance to each other, and to the original photograph from which their statistics were drawn. We recently showed that these stimuli produce enhanced responses in V2 neurons, compared with images that are matched only for their Fourier power spectra (19). This enhancement was not found in V1 neurons.

For the present study, we chose 15 original natural photographs to define 15 different texture families. These images were perceptually distinct, and human sensitivity to their higher order statistics spanned a range that was similar to the range found over a much larger set of natural photographs (19). We synthesized 15 different samples from each family, yielding 225 unique images.

Single Neuron Responses to Naturalistic Texture Stimuli. We recorded the spiking activity of 102 V1 and 103 V2 neurons in 13 anesthetized macaque monkeys to these texture stimuli. We presented the stimuli within a 4° aperture centered on the receptive field of each recorded neuron. Each of the 225 different stimuli appeared 20 times in pseudorandom order and was displayed for 100 ms, separated by 100 ms of uniform gray at the mean luminance. The same stimulus sequence was presented to each neuron. We have previously published a comparison of these responses to the responses

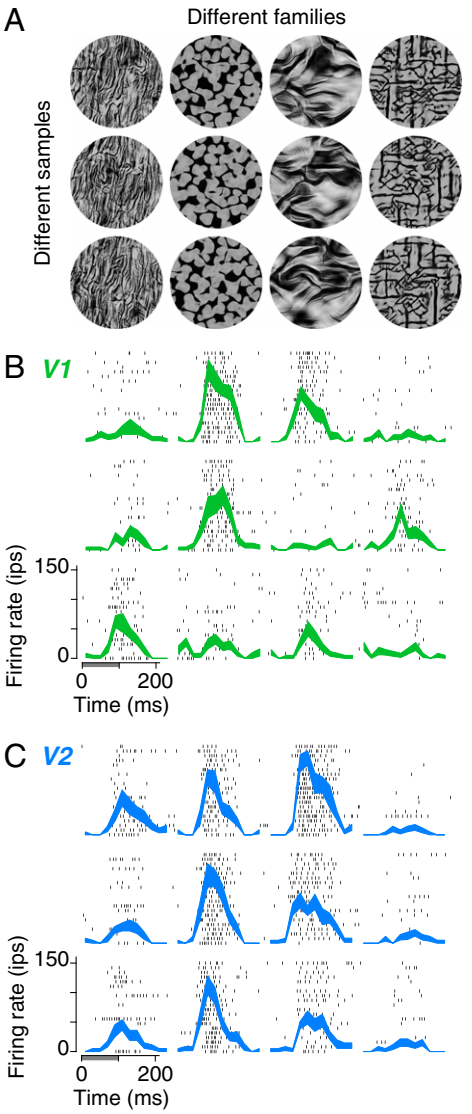

Fig. 1. Examples of texture stimuli and responses of V1 and V2 neurons. (*A*) Naturalistic textures. Each column contains three different samples from each of four texture families. The samples within each family are statistically matched, but differ in detail because the synthesis procedure is initialized with independent images of Gaussian white noise. (*B*) Raster plots and mean firing rates for an example V1 neuron, responding to textures in *A*. The gray bar indicates presentation of the stimulus (first 100 ms), and each row of black ticks represents the timing of spikes on a single presentation of the stimulus. The thickness of the lines indicates SEM across 20 repetitions of each of the images in *A*. (*C*) Same as in *B*, for an example V2 neuron.

obtained from spectrally matched (phase-scrambled) noise stimuli (19). Here, we present a new analysis of these data, which seeks to determine the relative selectivity and tolerance of V1 and V2 neurons for the different texture families and the image samples drawn from those families, respectively.

Texture stimuli elicited selective responses in most V1 and V2 neurons (Fig. 1 *B* and *C*). Neurons in both V1 and V2 displayed a characteristic firing rate for each image, with some variability across presentations. For most texture families, firing rates of V1 neurons were highly variable across the samples (Fig. 1*B*). In contrast, V2 neurons exhibited similar firing rates across samples, as well as more consistent differences in average firing rate across families (Fig. 1*C*); that is, V2 neurons appeared to be more tolerant to the variations in image detail that occur across samples

within a texture family and more selective for the statistical parameters that define the family.

To quantify this observation, we used a nested ANOVA to partition the total variance in firing rate for each neuron into three components representing variation across families (Fig. 1, columns), across samples within a family (Fig. 1, rows), and across repeated presentations of each sample (residual spiking variability across rows of each raster in Fig. 1 *B* and *C*). We first note that a smaller portion of V2 response variance was explained by the stimulus, compared with V1 (Fig. 2 *A* and *B*, *Insets*), consistent with previous reports (26). The reduction in explainable variance in V2 was accompanied by a reduction in the population average firing rate compared with V1 [8.3 impulses per second (ips) in V2 compared with 13.6 ips in V1], and may reflect a greater effect of anesthesia in V2.

Although variance across samples dominated the responses of most V1 neurons (Fig. 2*A*), many V2 neurons exhibited as much or more variance across families (Fig. 2*B*). However, the absolute levels of variance across and within families are affected by our particular choice of texture stimuli. To eliminate the influence of the stimulus ensemble, we compared the ratio of variance across and within families for neurons in V1 and V2 (Fig. 2 *C* and *D*). This ratio is similar to the *F*-statistic from our ANOVA analysis, with a large value indicating high tolerance to the statistical variation of samples within families for our stimulus set. We found a significantly larger value of the variance ratio in our population of V2 neurons compared with V1 (Fig. 2 *C* and *D*; $P < 0.001$, *t* test on the log variance ratio). Twenty-nine percent of neurons in V2 were more variable in their firing rate across vs. within families compared with 16% of V1 neurons. These data indicate that on the whole, the V2 population exhibited more stable responses across samples within a family.

Analyzing the Influence of Receptive Field Properties on Tolerance. We wondered whether this difference in tolerance was a consequence of well-known differences in receptive field properties between V1 and V2. For example, V2 contains a larger proportion of neurons that can be classified as complex [as opposed to simple (1, 13)], and the receptive fields of V2 neurons at a given eccentricity are about twice as large as the receptive fields in V1 (12, 27). Both of these properties would be expected to contribute to the variance ratio. Specifically, simple cells are sensitive to phase and should exhibit more response variation than complex cells across samples. Similarly, neurons with small receptive fields have a more limited area over which to compute statistics; thus their responses are expected to fluctuate with changes in local statistics across samples (note that the statistics of sample images within a family are identical only when measured across the entire image).

To examine these and other effects on the variance ratio, we measured responses of a subset of our V1 and V2 populations to drifting sinusoidal gratings, and used these measured responses to quantify 10 conventional receptive field properties. We then used a stepwise regression separately in both areas to determine which of these properties might explain the across-to-within-family variance ratios (*Methods*). Altogether, receptive field properties accounted for only a limited amount of diversity of the variance ratios in both areas (Fig. 3*I*; V1, $R^2 = 0.28$; V2, $R^2 = 0.42$). This result was not due to data insufficiency in our estimation of the variance ratio, because one-half of our data could predict the other accurately (V1, $R^2 = 0.89 \pm 0.02$; V2, $R^2 = 0.86 \pm 0.02$; mean and SD of bootstrapped distribution) (*Methods*). As expected, we found that size and the spatial phase sensitivity of receptive fields were significantly correlated with the variance ratio, and this relationship held for both V1 and V2 (Fig. 3 *A*–*D*). For V1 neurons, no other properties were significantly correlated (Fig. 3 *E* and *G*).

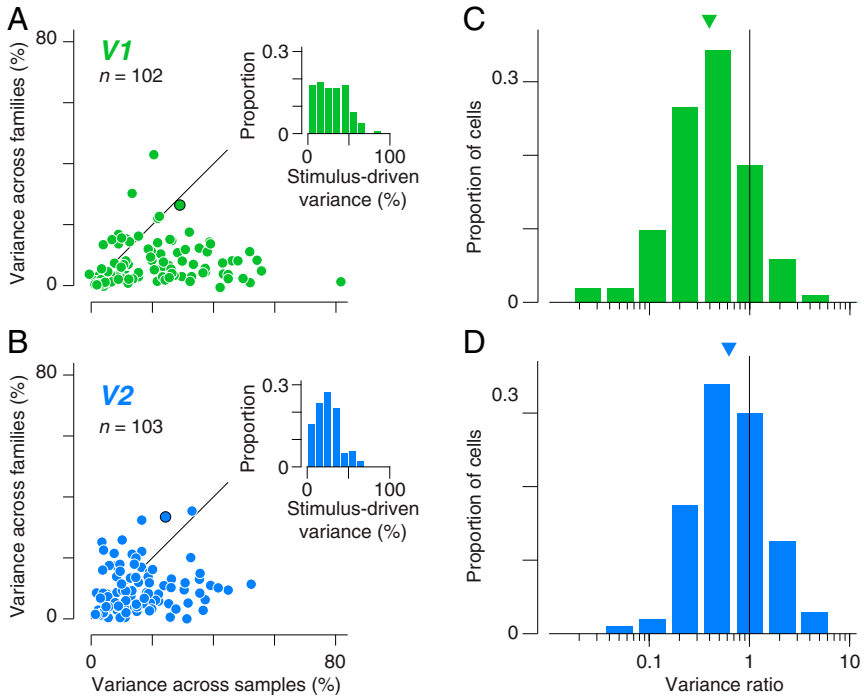

Fig. 2. Nested ANOVA analysis of single-unit responses in V1 and V2. (*A* and *B*) Response variance of single units in V1 and V2 is partitioned into a component across families, a component within families (across samples), and a residual component across stimulus repetitions (noise). The position of each point indicates, for a single neuron, the percentage of variance corresponding to the first two of these components. (*Insets*) Distribution of the sum of these first two components. Points outlined in black correspond to the example single units shown in Fig. 1. (*C* and *D*) Distributions of the ratio of across-family to across-sample variance for V1 and V2. The geometric mean variance ratio was 0.4 in V1 and 0.63 in V2 (indicated by triangles). The difference was significant ($P < 0.001$, *t* test in the log domain).

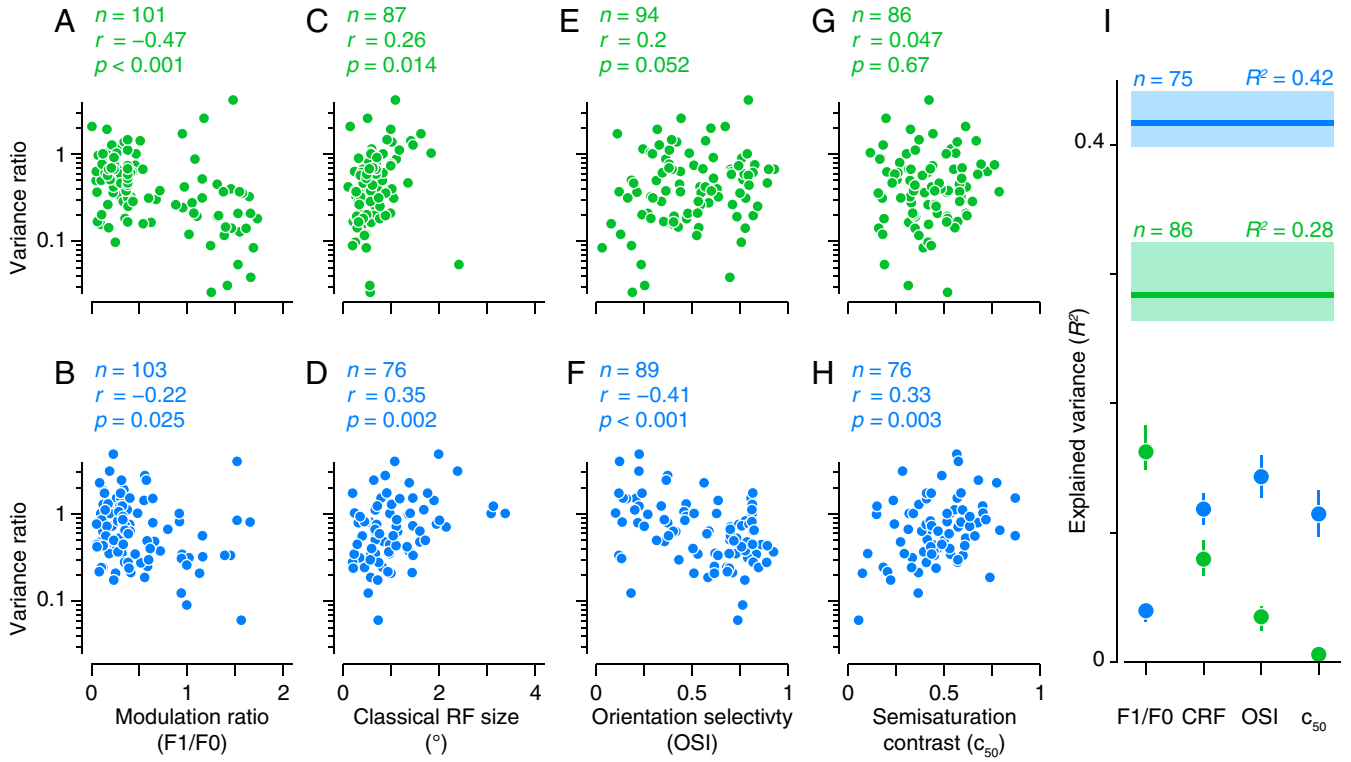

Fig. 3. Dependence of response tolerance on conventional receptive field properties. (*A–H*) Variance ratio (as in Fig. 2) plotted against receptive field properties of individual V1 (green) and V2 (blue) neurons. (*I*) Results of multiple linear regression of the variance ratio against the four receptive field properties highlighted in *A–H*. Horizontal lines show total explained variance for V1 (green) and V2 (blue). Points represent the contribution to the explained variance of different receptive field properties determined using the averaging-over-orderings technique (28). Shaded regions and error bars represent 95% confidence intervals computed using jackknife resampling.

However, in V2, orientation tuning (Fig. 3*F*) and contrast sensitivity (Fig. 3*H*) were also correlated with the variance ratio: Neurons with weaker orientation tuning and lower contrast sensitivity appeared to be more tolerant. To summarize these effects, we decomposed R^2 using the averaging-over-orderings technique (28) and examined the contribution of each property to the explained variance in V1 and V2 (Fig. 3*I*). This analysis confirmed the different pattern of contributions for the two areas. We conclude that although some of the increase in tolerance of V2 over V1 may be due to conventionally assessed differences in receptive field properties, some other factor is needed to explain fully the enhanced tolerance of V2 neurons.

Visualizing Selectivity and Tolerance of Neuronal Populations. We visualized the representation of texture stimuli within each neural population by transforming their responses from the high-dimensional response space (dimensionality = number of neurons) to a 2D space. Ideally, such a mapping would capture local and global aspects of the representation as much as possible. We used the *t*-distributed stochastic neighbor embedding (t-SNE) algorithm, which solves this problem by minimizing the difference between the high- and low-dimensional distributions of neighbor distances (29, 30). The choice of two dimensions is purely for interpretability and visualization, and is not meant to imply anything about the underlying dimensionality of representation in either area.

We normalized the firing rate of each neuron and applied t-SNE to the V1 and V2 populations separately (Fig. 4 *A* and *B*). Each of the 225 points represents population responses to a single texture sample, colored according to the family to which it belongs. Points that lie close together correspond to images that evoked similar responses from the neural population. Within V1, the groups of images from the same family generally produce scattered population responses, and the closest neighbors of most images do not correspond to samples from the same texture family (Fig. 4*A*). When applied to V2, the visualization reveals that population responses often cluster by texture family (Fig. 4*B*), with all of the samples from several families tightly grouped.

Decoding Neuronal Populations. The low-dimensional t-SNE visualization (Fig. 4) provides an intuition for how the representation in V2 differs from V1, which can be more precisely quantified using a neural population decoder. To this end, we analyzed the ability of V1 and V2 representations to support two different perceptual discrimination tasks. For the first task, we built a Poisson maximum likelihood decoder to discriminate between the 15 different samples within a texture family based on the responses within a neural population (*Methods* and Fig. 5*A*). Performance in both areas, averaged across all texture families, increased as the number of neurons included in the analysis increased, but V1 outperformed V2 for all population sizes (Fig. 5*B*). The representation of image content in V1 thus provides more information for discriminating between specific samples. For the second task, we built another decoder to discriminate between the 15 different texture families (*Methods* and Fig. 5*A*). We tested this decoder's ability to generalize across samples by training on a subset of samples and testing on samples not used in the training. For both V1 and V2, and for all population sizes, absolute performance on this task was worse than on the sample classification task, although the difference was much larger in V1 (Fig. 5*B*). However, in contrast to the sample classification task, V2 outperformed V1 for all population sizes. To examine whether this result could be a consequence of the differences in receptive field properties described above (Fig. 3), we excluded neurons classified as simple from both areas and selected subpopulations matched for classical receptive

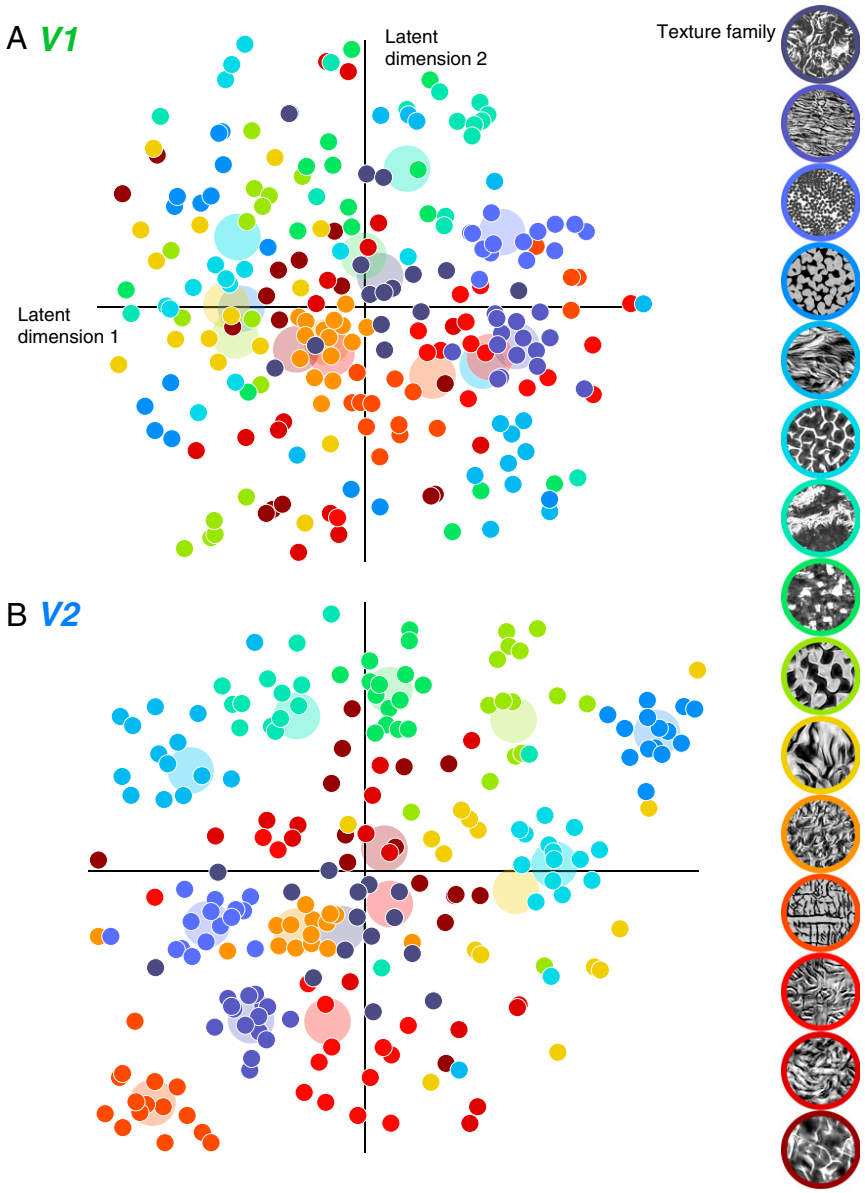

Fig. 4. Two-dimensional visualization of neural population responses in V1 and V2. (*A*) V1 population response to each visual texture stimulus, displayed in a 2D coordinate system that captures the responses of 102 V1 neurons [computed using t-SNE (30)]. Each point represents one texture image, with color indicating the texture family. The larger, desaturated disks in the background indicate the centroid of all samples within each family. (*B*) Same analysis for the responses of 103 V2 neurons.

field (CRF) size. This matching procedure had little effect on V2 performance in either task, but did reduce V1 performance on the sample task and increase V1 performance on the family task (*Methods*). However, performance in the two areas remained significantly different, suggesting more complex forms of selectivity are involved.

Comparing Selectivity of Neuronal Populations. To elucidate the V2 response properties that allow it to outperform V1 in family classification, we examined the dependence of performance on the differences in statistics between pairs of texture families. We built a Poisson maximum likelihood decoder to best discriminate between each pair of texture families (105 different comparisons). Comparing performance in V1 and V2 reveals two prominent features (Fig. 6*A*). First, performance in V1 and V2 was highly correlated across the different texture discriminations ($r = 0.82$, $P < 0.001$), suggesting that some of the features that drive performance in V1 are also responsible for performance in V2. Second, V2 neurons performed better for nearly all pairs, and this improvement was approximately independent of the performance seen in V1 (Fig. 6*A*). A straight-line fit suggests that if V1 discrimination performance were at chance, V2 performance would be 65% correct [discriminability (d′) = 0.54]. To understand this relationship, we sought to separate those stimulus properties that drive performance in both V1 and V2 from those stimulus properties that underlie the increase in performance of V2 over V1.

We chose texture families for this study that differed in their spectral content: the relative amount of energy at different orientations and spatial frequencies. V1 neurons are highly selective for spectral content (4), and this selectivity is maintained in V2 (13). We wondered whether the spectral characteristics of the stimuli could explain V1 performance. Across all 105 pairs of texture families, we measured the magnitude of the difference in spectral statistics between the two families. We then predicted

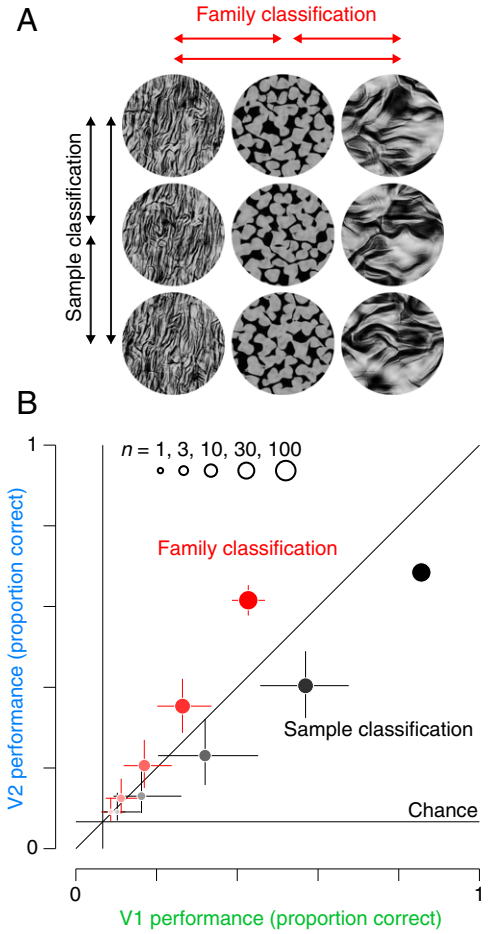

Fig. 5. Quantifying representational differences between V1 and V2. (*A*) Schematic of sample (black) and family (red) classification. For sample classification, holdout data were classified among the 15 different samples for each family. Performance for each of the families was then averaged together to get total performance. For family classification, the decoder was trained on multiple samples within each family, and then used to classify held out data into each of the 15 different families. (*B*) Comparison of proportion of correct classification of V1 and V2 populations for family classification (red) and sample classification (black). We computed performance measures for both tasks using five different population sizes, indicated by the dot size ($n = 1$, $n = 3$, $n = 10$, $n = 30$, and $n = 100$). Chance performance for both tasks was 1/15. Error bars represent 95% confidence intervals of the bootstrapped distribution over included neurons and cross-validation partitioning.

V1 discrimination performance from the statistical differences, over all pairs (*Methods* and Fig. 6*B*). The spectral differences predicted V1 performance well ($r = 0.7$, $P < 0.001$), and the same model also provided a good prediction for V2 performance ($r = 0.59$, $P < 0.001$). Reoptimizing the weights to predict V2 responses barely improved the correlation ($r = 0.6$, $P < 0.001$), consistent with the notion that the spectral information represented in V2 is directly inherited from V1. However, the spectral statistics captured little of the difference in performance between V1 and V2 ($r = 0.22$, $P < 0.05$).

These analyses suggest that the superior performance of V2 must be due to the higher order (i.e., beyond second order) correlations present in the texture model. To test this theory, we extracted the parameters that capture higher order statistics through correlations of filter response magnitudes across position, frequency, and orientation, and projected out the portion captured by the spectral statistics. We then predicted the difference in V1 and V2 performance (Fig. 6*C*). Differences in the higher order statistics, in contrast to spectral statistics, provided a good prediction for the V1/V2 performance difference ($r = 0.61$, $P < 0.001$).

In summary, V1 discrimination performance was well captured by the spectral statistics of naturalistic textures. This same set of statistics captured a significant portion of V2 discrimination performance, but most of the superiority of V2 over V1 comes from higher order statistics.

Discussion

Our results support the hypothesis that populations of V2 neurons represent statistics of the activity of local ensembles of V1 neurons, which capture the appearance of naturally occurring textures. Using a set of stimuli for which these statistics are tightly controlled, we showed that, relative to neurons in V1, V2 neurons exhibit increased selectivity for these statistics, accompanied by an increased tolerance for randomized image variations that do not affect these statistics. This "tolerance to statistical resampling" complements the more widely discussed visual invariances to geometric distortions (e.g., translation, rotation, dilation) (8, 10) or changes in the intensity, color, or position of a light source (9, 31).

Our results also help to integrate and interpret other findings. The selectivity of V2 neurons for many artificial stimuli, including gratings, angles, curves, anomalous contours, and texture-defined patterns, is nearly the same as the selectivity of V1 neurons (14–17, 32–35). This result would be expected if V2 neurons are selective for a broad set of V1 response statistics and not for a small subset of specialized combinations of V1 inputs, as assumed by these approaches. On the other hand, the tolerance of V2 cells identified here does seem consistent with the previously identified behaviors of "complex unoriented" V2 cells (36), which are selective for patches of light of a particular size but tolerant to changes in position over a much larger region. Such a property may explain why orientation selectivity so strongly predicted tolerance in V2 but less so in V1. This relationship might also reflect greater heterogeneity of orientation tuning within V2 receptive fields (16), providing a substrate for computing local orientation statistics.

Our results complement recent work demonstrating V2 selectivity for third- and fourth-order pixel statistics. Yu et al. (20) examined responses of V1 and V2 neurons to binary images synthesized with controlled pixel statistics up to fourth order, and found that neuronal selectivity for multipoint (i.e., third and fourth order) correlations is infrequent in V1 but common in V2. The strength of this work derives from the well-defined stimulus ensemble, which covers the full set of statistics up to fourth order, and allows a thorough assessment of the selectivity for individual statistics in the responses of single neurons. On the other hand, the restriction to statistics of a particular order, although mathematically natural, is not necessarily aligned with the restrictions imposed by the computational capabilities of biological visual systems, and this may explain why selectivity of V2 neurons for these statistics is only modestly greater than selectivity of V1 neurons. The stimuli in our experiments are constrained by statistics that are defined in terms of an idealized response model for a V1 population. Although they also constrain multipoint pixel statistics, they do not isolate them in pure form, and they span too large a space to allow a thorough experimental characterization of selectivity in individual cells. On the other hand, they represent quantities that may be more directly related to the construction of V2 responses from V1 afferents, and they allow direct synthesis of stimuli bearing strong perceptual resemblance to their ecological counterparts (18, 23, 24, 37).

The particular statistics we matched to create our texture families are surely not represented fully and only in V2, and this may explain why the reported difference in selectivity and tolerance between V1 and V2, although robust, is not qualitative. In particular, these statistics include both the local correlation of oriented linear filter responses (equivalent to a partial representation of

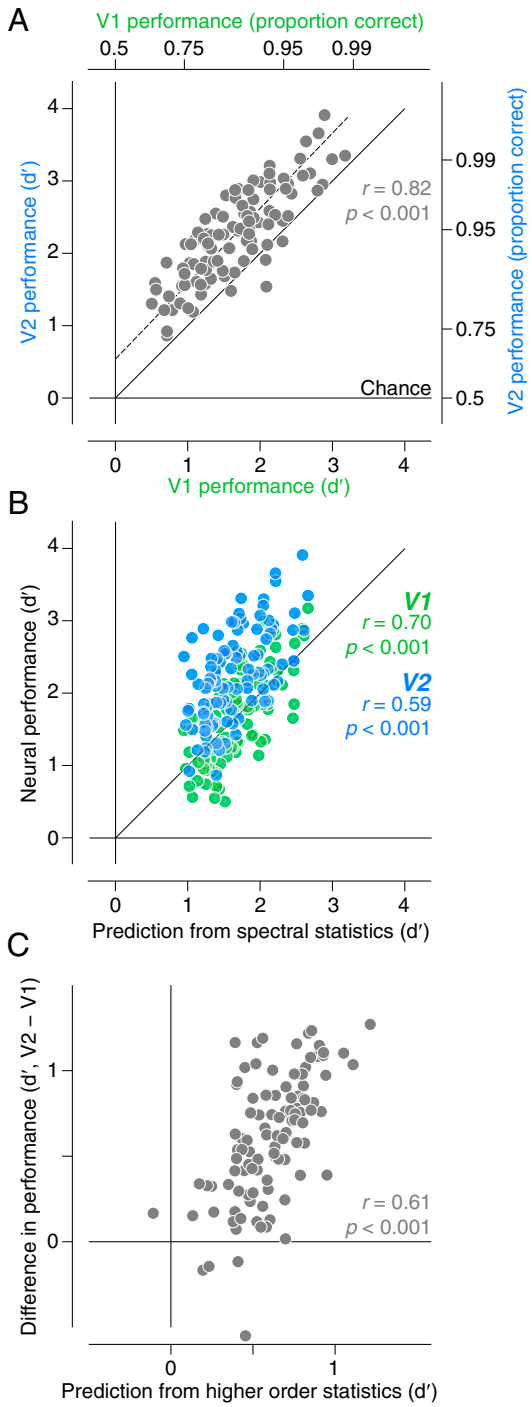

Fig. 6. Texture discrimination performance of neural populations. (*A*) Comparison of V1 and V2 performance on pairwise texture discrimination. Performance values were plotted on coordinates that varied linearly in discriminability (d′). The right and top axes indicate corresponding values of performance expressed as the proportion correct. Each point represents one of 105 pairwise comparisons among the 15 texture families. The dashed line indicates the best fit using total least squares. (*B*) Comparison of V1 and V2 performance with the performance of a model capturing spectral statistics. The magnitude of difference in spectral statistics for each texture family pair was weighted to account best for the performance of V1. Both V1 performance and V2 performance were plotted against this spectral prediction. (*C*) Comparison of the difference in V1 and V2 performance with the strength of higher order correlation differences. The magnitude of difference in higher order correlations for each texture family pair was weighted to predict best the difference in V1 and V2 performance.

average spectral power across the image) as well as pairwise correlations between the magnitudes of responses of oriented filters tuned to different orientations, spatial frequencies, and positions. We created different families from the statistics extracted from 15 original photographs, which differed in both the spectral and higher order statistics. We found that the spectral differences between different families accounted for a substantial portion of the discrimination performance of both V1 and V2 populations. However, V2 nearly always outperformed V1, and this superiority was well predicted by the differences in higher order statistics. This finding suggests that an artificial stimulus set in which families differ only in higher order statistics would better differentiate the discrimination performance of V1 and V2.

How do V2 neurons achieve higher classification and discrimination performance than their V1 inputs? There are two possible answers: reducing variability in the representation of individual families or increasing the mean separation in the representations of different families. The first of these possibilities can be achieved by combining many V1 inputs so as to average out their trial-by-trial variability. Larger receptive fields may be an indication of such a construction: Indeed, larger receptive fields are associated with higher variance ratios and better family classification performance. However, when we matched receptive field sizes between the two areas, V2 still performed better in family classification. Thus, we posit that V2 neurons are also taking advantage of the second option, transforming their V1 inputs to make family differences more explicit in their average responses. This transformation amounts to "untangling" the representation of visual features that were not directly decodable in the V1 representation (38). Specifically, V1 neurons do not appear to signal the presence of higher order correlations with a consistent change in firing rate, whereas V2 neurons do (19, 20). As a result, larger differences in higher order correlations between families explain a significant portion of the increased discrimination performance in V2 (Fig. 6*C*).

Perceptually, invariances related to statistical resampling were originally proposed by Julesz (39) as a testable prediction of statistical texture models, and have been used to test and refine such models in both vision (21–24) and audition (40). Theories regarding the statistical summary nature of "crowding" effects in peripheral vision (41–44) have also been tested for such perceptual invariances (18, 37), and are consistent with the representation of texture statistics in area V2. Although our analysis of V2 responses is qualitatively consistent with these perceptual observations, the connection is difficult to test quantitatively. In particular, the statistics in our texture stimuli were computed by averaging over the full stimulus aperture, which was held to a fixed size of 4° for all cells to allow a reasonable interpretation of population decoding. This size was generally larger than the receptive fields of the neurons (Fig. 3 *C* and *D*). Thus, most neurons saw only a portion of the stimuli, over which the statistics would not have been fully matched. Finally, recall that the transformation from V1 to V2 is part of a cascade, and it may well be that perception relies more on downstream areas, such as V4, where neurons may show even more selectivity and tolerance for the statistics we used (31, 45).

The visual world is often described in terms of forms or "things" made up of lines, edges, contours, and boundaries, and these symbolic descriptions have played a dominant role in developing theories for both biological and machine representations of visual information. However, textures and "stuff" (46) are ubiquitous in natural visual scenes, and are not easily captured with edge or contour descriptions. The results presented here suggest that V2 neurons combine V1 afferents to represent perceptually relevant statistical features of visual textures. It is currently unknown whether this statistical description of the visual world is also sufficient to account for perception of visual form. Recent work suggests that textural statistics, such as used here, can account for aspects of peripheral vision that are not exclusive to the perception

of texture (18, 37). Additionally, recent successes in machine recognition of complex objects using multistage neural networks call into question the need for explicit boundary, contour, or edge information in high-level vision. Indeed, the model responses at different stages of these neural networks have provided a good basis for accounting for neural responses in both midlevel and late stages of visual cortex (47, 48), and attempts to visualize the selectivities of model units at midlevel stages have often revealed texture-like visual structure (49). Thus, the two-stage representation we describe here may provide a foundation for the representation of the more complex and structured signals found in images of objects or of entire visual scenes (41).

Methods

Physiology.

Recording. The data analyzed here were also used in a previous article (19), and the full methods are provided there. In brief, we recorded from 13 anesthetized, paralyzed, adult macaque monkeys (two *Macaca nemestrina* and 11 *Macaca cynomolgus*). Our standard methods for surgical preparation have been documented in detail previously (50). We maintained anesthesia with infusion of sufentanil citrate (6–30 $\mu g \cdot kg^{-1} \cdot h^{-1}$) and paralysis with infusion of vecuronium bromide (Norcuron; 0.1 $mg \cdot kg^{-1} \cdot h^{-1}$) in isotonic dextrose-Normosol solution. All experimental procedures were conducted in compliance with the NIH's *Guide for the Care and Use of Laboratory Animals* (51) and with the approval of the New York University Animal Welfare Committee. We made a craniotomy and durotomy centered ∼2–4 mm posterior to the lunate sulcus and 10–16 mm lateral, and recorded single-unit activity using quartz-platinum-tungsten microelectrodes (Thomas Recording). We distinguished V2 from V1 on the basis of depth from the cortical surface and receptive field location.

Stimulus generation. We generated stimuli using the texture analysis-synthesis procedure described by Portilla and Simoncelli (23) (software and examples are available at www.cns.nyu.edu/~lcv/texture/). Fifteen different grayscale photographs (320 × 320 pixels) of visual texture served as the prototypes for each "texture family." In brief, each image was decomposed with a multiscale multiorientation bank of filters with four orientations and four spatial scales, designed so as to tile the Fourier domain (52). For each filter, we computed the linear response and the local magnitude response (square root of sum of squared responses of the filter and its Hilbert transform), roughly analogous to the responses of V1 simple and complex cells. We then computed pairwise products across filter responses at different positions (within each orientation and scale and across a 7 × 7 neighborhood) for both sets of responses, and (for the magnitudes only) across different orientations and scales. We also included products of linear filter responses with phase-doubled responses at the next coarsest scale. All of these pairwise products were averaged across the spatial extent of the image, yielding correlations. The correlations of the linear responses are second-order statistics, in that they represent the averages of quadratic functions of pixel values. The correlations of magnitudes (and phase-doubled responses) are of higher order, due to the additional nonlinearities in the magnitude (phase-doubling) computation. We additionally computed the average magnitude within each frequency band and the marginal pixel statistics (skew and kurtosis). For each family, we synthesized 15 samples by initializing 15 different images with Gaussian white noise and adjusting each until it matched the model parameters computed on the corresponding original image (23).

Stimulus presentation. We presented visual stimuli on a gamma-corrected cathode ray tube monitor (Eizo T966; mean luminance of 33 cd/m^2) at a resolution of 1,280 × 960, with a refresh rate of 120 Hz. Stimuli were presented using Expo software on an Apple Macintosh computer. For each isolated unit, we first determined its ocular dominance and occluded the nonpreferred eye. We used drifting sinusoidal gratings to characterize the basic receptive field properties of each unit, including tuning for orientation and direction, spatial and temporal frequency, size, and contrast. We then presented the texture stimuli. We used a set of 15 texture families and generated 15 samples for each texture family for a total of 225 images. Another 225 images of phase-randomized noise were also included, but not analyzed further here. We presented the images in pseudorandom order for 100 ms each, separated by 100 ms of mean luminance. Each image was presented 20 times. Images were presented to every unit at the same scale and at a size of 4° within a raised cosine aperture. We chose a 4° aperture to be larger than all of the receptive fields at the eccentricities from which we typically record. Nearly all recorded units had receptive fields smaller than 4°, and the majority were less than 2°.

Analysis.

ANOVA. For all quantitative analyses, we averaged spike counts within a 100-ms time window aligned to the response onset of each single unit. Response onset was determined by inspection as the first time point eliciting a response above baseline; results were nearly identical when using a quantitative criterion based on the SD of the response. We first applied a Freeman–Tukey variance-stabilizing transformation (53) to the spike counts for each neuron ($z = \sqrt{x} + \sqrt{x+1}$). This preprocessing step transforms the roughly Poisson-distributed spike counts to be more Gaussian, removing dependencies between the mean and variance. We then performed a nested ANOVA analysis to partition the total variance into the portions arising across families, across samples within a family, and across repetitions of the same stimulus. The ANOVA generates an *F*-statistic that captures the ratio of variances between each hierarchical level. For the vast majority of neurons, the *F*-statistic was significant for ratios of variance across repetitions and across samples (101 of 102 in V1, 103 of 103 in V2), as well as for ratios of variance across samples and across families (91 of 102 in V1, 97 of 103 in V2). We chose to perform further analysis using the ratio between partitioned variance, but all results were qualitatively similar when using the *F*-statistic itself. To obtain the variance ratio, we divided the percent variance across families by the percent variance across samples. To avoid outlying values when either variance was very low, we stabilized the ratio by adding 2% variance to both the numerator and denominator. We tested how reliable our estimates of the variance ratio were by splitting the 20 repetitions for each condition in half and performing the ANOVA analysis separately on both halves of the data for each neuron. We repeated this process 10,000 times with different partitions of the original repetitions and asked how well our estimate on half of the data could predict the other half.

Regression. Basic receptive field properties for each neuron (e.g., receptive field size, contrast response function) were determined offline by using maximum likelihood estimation to fit an appropriate parametric form to each tuning function. These fits were only obtainable for a subset of neurons (84% in V1, 73% in V2) due to incomplete characterization arising from time constraints during the experiment. We first asked how well we could predict the log variance ratio in each area using a large number of receptive field properties [preferred spatial frequency, spatial frequency bandwidth, orientation selectivity, CRF size, contrast exponent, semisaturation contrast (c_{50}), maximum firing rate, surround suppression index, modulation ratio (F1/F0), and texture modulation index (19)]. We used the log variance ratio because the ratios were approximately normally distributed in the log domain. We used a stepwise linear model to estimate which receptive field properties added to the goodness of fit. For V1, only receptive field size and modulation ratio were included in the model. For V2, receptive field size and modulation ratio were included, along with orientation selectivity and c_{50}. CRF size was defined as the SD of the center in a ratio of Gaussians model. The modulation ratio was computed from responses to the 1-s presentation of an optimal grating and represents the ratio between the first harmonic and mean of the average response. The orientation selectivity index (OSI) was computed as the circular variance of the baseline-subtracted firing rates to each orientation, so that OSI = 0 indicated no selectivity and OSI = 1 indicated sharp tuning for orientation. The c_{50} represents the contrast level that evoked half of the maximum firing rate in a Naka–Rushton fit to the responses to a grating of varying contrast. To examine how each of these predictors contributed to the variance ratio, we used an averaging-over-orderings (19, 28) technique to estimate variance explained by each receptive field property. This technique allowed us to assess the relative importance of each predictor in each area. We computed error bars for the contribution of each receptive field property and the overall explained variance using a jackknife procedure. We reapplied the averaging-over-orderings procedure to the dataset with one neuron left out and computed 95% confidence intervals over the distribution of all partial datasets.

t-SNE visualization. To visualize the structure of the data we used a method for dimensionality reduction known as *t*-distributed stochastic neighbor embedding (t-SNE) (30), a variant of the technique originally developed by Hinton and Roweis (29). This method attempts to minimize the divergence between the distributions of neighbor probability in the high-dimensional space and low-dimensional space. The input to the algorithm was a set of 225 data vectors, each of which collected the firing rates of all neurons in an area to a stimulus. We also normalized the data so that, for each neuron, responses to the 225 images had a mean of 0 and SD of 1. In executing the t-SNE analysis, we chose an initial dimensionality of 90 and a perplexity value of 30.

Classification decoding. We used a simple Poisson decoder to classify samples or families into one of 15 different categories. On each iteration, we randomly selected a number of units from our recorded population. Because our units

were recorded sequentially, we randomized the order of repetitions for each cell. To compute performance in the sample classification task, we estimated the mean spike counts of each neuron for each of the 15 samples within each family by computing the sample average over 10 of the 20 repetitions. For the held-out 10 repetitions of each sample, we computed which of the 15 samples was most likely to have produced the population response, assuming independent Poisson variability under the estimated mean spike counts. We computed the average performance (% correct) over all samples and families, and repeated this process 10,000 times to get a performance for each population size. To compute performance in the family classification task, we estimated the average spike counts for each family over 8 of the 15 different samples and for all repetitions. For each of the repetitions of the held-out seven samples, we computed which of the 15 families was most likely to have produced the population response. We computed the average performance over all repetitions and repeated this process 10,000 times to get a performance for each population size. We computed performance measures for both tasks using population sizes of 1, 3, 10, 30, and 100 neurons. Results were similar using several alternative decoding methods, including a linear classifier and a mixture-of-Poissons model. The potential advantage of a more sophisticated mixture-of-Poissons model was negated by the larger parameter space and insufficiency of data. We also performed family classification by training on a subset of repetitions over all samples and found increased performance in both V1 and V2, although V2 still outperformed V1.

Matched subpopulation decoding. To examine the effect of receptive field properties that differ sharply between V1 and V2 on decoding, we excluded neurons with a modulation ratio greater than 0.8 and extracted 40-neuron subpopulations in each area that were matched for the mean and variance of CRF size (mean CRF in both V1 and V2 = 0.73 ± 0.02°). We decoded our CRF-matched, complex cell subpopulations and compared performance with the performance achieved by 40 neuron subpopulations sampled randomly from the full population of both areas (mean CRF in V1 = 0.62 ± 0.05°, mean CRF in V2 = 1.1 ± 0.09°). In the sample classification task, V1 performance was significantly reduced by drawing matched subpopulations (65–55%), and there was no effect on V2 performance (which remained at 46%). V1 performed significantly better than V2 in sample classification for both unmatched ($P < 0.005$, bootstrap test resampling neurons and cross-validation partitioning) and matched ($P < 0.01$) subpopulations. In the family task, V1 performance was increased by drawing matched subpopulations (30–35%) and V2 performance was only slightly decreased (41–40%). V2 performed significantly better than V1 in family classification for both unmatched ($P < 0.05$) and matched ($P < 0.05$) subpopulations.

Discrimination decoding and prediction. We used the same decoding procedure for family classification but performed discrimination between all pairs of texture families, yielding 105 pairwise comparisons. All discrimination decoding was performed using 100 units and was repeated 10,000 times to get a performance value. We transformed the measured performance values for V1 and V2 into units of discriminability (d′) and performed total least squares regression to get a linear fit to the V1 and V2 data. We then isolated two subsets of parameters from the full set contained in the texture model used to generate our stimuli. The first consisted of the correlations of linear filter responses at nearby locations, which represent second-order pixel statistics and are most intuitively described as representing a portion of the power spectrum (as such, we refer to them as spectral). We also gathered a set of higher order statistics, consisting of correlations of magnitudes at neighboring locations, orientations, and scales, and correlations of phase-adjusted filter responses at adjacent scales (23).

To summarize the family discrimination capability of each group of statistics, we computed a matrix whose columns contained the absolute value of the difference between those statistics for each pair of texture families (105 columns, one for each pair of families). For the spectral statistics (matrix size = 125 × 105), we reduced the dimensionality (number of rows) of this matrix using principal components analysis (PCA). We found that four components captured 70% of the variance, and standard regression analysis revealed that both V1 and V2 performance was well predicted by a weighted sum of these components (Fig. 6*B*). To examine the relationship between higher order statistics and neural performance, we first removed the effects of the spectral statistics. We adjusted each of the rows of the higher order difference matrix (matrix size = 552 × 105) by projecting out the four dimensions spanned by the rows of the PCA-reduced spectral difference matrix. We then reduced the dimensionality (number of rows) of this matrix using PCA, retaining those components needed to capture at least 70% of the variance (in this case, 10 components). Regression analysis revealed that a weighted sum of these components provided a good prediction for the difference in performance between V2 and V1 (Fig. 6*C*).

ACKNOWLEDGMENTS. We thank R. L. T. Goris for useful discussions and members of the Movshon Laboratory for help with physiological experiments. This work was supported by NIH Grant EY22428, the Howard Hughes Medical Institute, and National Science Foundation Graduate Research fellowships (to C.M.Z. and J.F.).

1. Hubel DH, Wiesel TN (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol* 160:106–154.
2. Movshon JA, Thompson ID, Tolhurst DJ (1978) Spatial summation in the receptive fields of simple cells in the cat's striate cortex. *J Physiol* 283:53–77.
3. Tolhurst DJ, Thompson ID (1981) On the variety of spatial frequency selectivities shown by neurons in area 17 of the cat. *Proc R Soc Lond B Biol Sci* 213(1191):183–199.
4. De Valois RL, Albrecht DG, Thorell LG (1982) Spatial frequency selectivity of cells in macaque visual cortex. *Vision Res* 22(5):545–559.
5. Movshon JA, Thompson ID, Tolhurst DJ (1978) Receptive field organization of complex cells in the cat's striate cortex. *J Physiol* 283:79–99.
6. Adelson EH, Bergen JR (1985) Spatiotemporal energy models for the perception of motion. *J Opt Soc Am A* 2(2):284–299.
7. Movshon JA, Adelson EH, Gizzi M, Newsome WT (1985) The analysis of moving visual patterns. *Pontificia Academica Scripta Varia* 54:117–151.
8. Zoccolan D, Kouh M, Poggio T, DiCarlo JJ (2007) Trade-off between object selectivity and tolerance in monkey inferotemporal cortex. *J Neurosci* 27(45):12292–12307.
9. Vogels R, Biederman I (2002) Effects of illumination intensity and direction on object coding in macaque inferior temporal cortex. *Cereb Cortex* 12(7):756–766.
10. Rust NC, Dicarlo JJ (2010) Selectivity and tolerance ("invariance") both increase as visual information propagates from cortical area V4 to IT. *J Neurosci* 30(39):12978–12995.
11. Rust NC, DiCarlo JJ (2012) Balanced increases in selectivity and tolerance produce constant sparseness along the ventral visual stream. *J Neurosci* 32(30):10170–10182.
12. Gattass R, Gross CG, Sandell JH (1981) Visual topography of V2 in the macaque. *J Comp Neurol* 201(4):519–539.
13. Levitt JB, Kiper DC, Movshon JA (1994) Receptive fields and functional architecture of macaque V2. *J Neurophysiol* 71(6):2517–2542.
14. Ito M, Komatsu H (2004) Representation of angles embedded within contour stimuli in area V2 of macaque monkeys. *J Neurosci* 24(13):3313–3324.
15. Hegdé J, Van Essen DC (2007) A comparative study of shape representation in macaque visual areas v2 and v4. *Cereb Cortex* 17(5):1100–1116.
16. Anzai A, Peng X, Van Essen DC (2007) Neurons in monkey visual area V2 encode combinations of orientations. *Nat Neurosci* 10(10):1313–1321.
17. El-Shamayleh Y, Movshon JA (2011) Neuronal responses to texture-defined form in macaque visual area V2. *J Neurosci* 31(23):8543–8555.
18. Freeman J, Simoncelli EP (2011) Metamers of the ventral stream. *Nat Neurosci* 14(9):1195–1201.
19. Freeman J, Ziemba CM, Heeger DJ, Simoncelli EP, Movshon JA (2013) A functional and perceptual signature of the second visual area in primates. *Nat Neurosci* 16(7):974–981.
20. Yu Y, Schmid AM, Victor JD (2015) Visual processing of informative multipoint correlations arises primarily in V2. *eLife* 4:e06604.
21. Tkačik G, Prentice JS, Victor JD, Balasubramanian V (2010) Local statistics in natural scenes predict the saliency of synthetic textures. *Proc Natl Acad Sci USA* 107(42):18149–18154.
22. Victor JD, Conte MM (2012) Local image statistics: Maximum-entropy constructions and perceptual salience. *J Opt Soc Am A Opt Image Sci Vis* 29(7):1313–1345.
23. Portilla J, Simoncelli EP (2000) A parametric texture model based on joint statistics of complex wavelet coefficients. *Int J Comput Vis* 40(1):49–71.
24. Balas BJ (2006) Texture synthesis and perception: Using computational models to study texture representations in the human visual system. *Vision Res* 46(3):299–309.
25. Ackermann JF, Landy MS (2014) Statistical templates for visual search. *J Vis* 14(3):18.
26. Goris RLT, Movshon JA, Simoncelli EP (2014) Partitioning neuronal variability. *Nat Neurosci* 17(6):858–865.
27. Shushruth S, Ichida JM, Levitt JB, Angelucci A (2009) Comparison of spatial summation properties of neurons in macaque V1 and V2. *J Neurophysiol* 102(4):2069–2083.
28. Grömping U (2007) Estimators of relative importance in linear regression based on variance decomposition. *Am Stat* 61(2):139–147.
29. Hinton GE, Roweis ST (2002) Stochastic neighbor embedding. *Adv Neural Inf Process Syst* 15:833–840.
30. Van der Maaten L, Hinton GE (2008) Visualizing data using t-SNE. *J Mach Learn Res* 9:2579–2605.
31. Arcizet F, Jouffrais C, Girard P (2008) Natural textures classification in area V4 of the macaque monkey. *Exp Brain Res* 189(1):109–120.
32. Sincich LC, Horton JC (2005) The circuitry of V1 and V2: Integration of color, form, and motion. *Annu Rev Neurosci* 28:303–326.
33. Hegdé J, Van Essen DC (2000) Selectivity for complex shapes in primate visual area V2. *J Neurosci* 20(5):RC61.
34. Lee TS, Nguyen M (2001) Dynamics of subjective contour formation in the early visual cortex. *Proc Natl Acad Sci USA* 98(4):1907–1911.
35. Mahon LE, De Valois RL (2001) Cartesian and non-Cartesian responses in LGN, V1, and V2 cells. *Vis Neurosci* 18(6):973–981.

36. Hubel DH, Livingstone MS (1985) Complex-unoriented cells in a subregion of primate area 18. *Nature* 315(6017):325–327.
37. Balas B, Nakano L, Rosenholtz R (2009) A summary-statistic representation in peripheral vision explains visual crowding. *J Vis* 9(12):13.1–18.
38. DiCarlo JJ, Cox DD (2007) Untangling invariant object recognition. *Trends Cogn Sci* 11(8):333–341.
39. Julesz B (1962) Visual pattern discrimination. *IRE Trans Inf Theory* 8(2):84–92.
40. McDermott JH, Simoncelli EP (2011) Sound texture perception via statistics of the auditory periphery: Evidence from sound synthesis. *Neuron* 71(5):926–940.
41. Lettvin JY (1976) On seeing sidelong. *The Sciences* 16(4):10–20.
42. Parkes L, Lund J, Angelucci A, Solomon JA, Morgan M (2001) Compulsory averaging of crowded orientation signals in human vision. *Nat Neurosci* 4(7):739–744.
43. Pelli DG, Palomares M, Majaj NJ (2004) Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *J Vis* 4(12):1136–1169.
44. Greenwood JA, Bex PJ, Dakin SC (2009) Positional averaging explains crowding with letter-like stimuli. *Proc Natl Acad Sci USA* 106(31):13130–13135.
45. Okazawa G, Tajima S, Komatsu H (2015) Image statistics underlying natural texture selectivity of neurons in macaque V4. *Proc Natl Acad Sci USA* 112(4):E351–E360.
46. Adelson EH (2001) On seeing stuff: The perception of materials by humans and machines. *Proc SPIE Int Soc Opt Eng* 4299:1–12.
47. Yamins DLK, et al. (2014) Performance-optimized hierarchical models predict neural responses in higher visual cortex. *Proc Natl Acad Sci USA* 111(23):8619–8624.
48. Khaligh-Razavi S-M, Kriegeskorte N (2014) Deep supervised, but not unsupervised, models may explain IT cortical representation. *PLOS Comput Biol* 10(11): e1003915.
49. Zeiler MD, Fergus R (2014) Visualizing and understanding convolutional networks. *Lecture Notes in Computer Science*, Computer Vision – ECCV 2014, eds D Fleet, T Pajdla, B Schiele, T Tuytelaars (Springer International, Cham, Switzerland), Vol 8689, pp 818–833.
50. Cavanaugh JR, Bair W, Movshon JA (2002) Nature and interaction of signals from the receptive field center and surround in macaque V1 neurons. *J Neurophysiol* 88(5): 2530–2546.
51. Committee for the Update of the Guide for the Care and Use of Laboratory Animals, Institute for Laboratory Animal Research, Division on Earth and Life Studies, National Research Council (2011) *Guide for the Care and Use of Laboratory Animals* (National Academies Press, Washington, DC), 8th Ed.
52. Simoncelli EP, Freeman WT (1995) The steerable pyramid: A flexible architecture for multi-scale derivative computation. *Proceedings of the IEEE Conference on Image Processing* 3:444–447.
53. Freeman MF, Tukey JW (1950) Transformations related to the angular and the square root. *The Annals of Mathematical Statistics* 21(4):607–611.