

A Sparse Object Coding Scheme in Area V4

Eric T. Carlson,^{1,3,4} Russell J. Rasquinha,^{1,3,4}
Kechen Zhang,¹ and Charles E. Connor^{2,3,*}

¹Department of Biomedical Engineering

²Solomon H. Snyder Department of Neuroscience
Johns Hopkins University School of Medicine, Baltimore,
MD 21205, USA

³Zanvyl Krieger Mind/Brain Institute, Johns Hopkins
University, Baltimore, MD 21218, USA

Summary

Sparse coding has long been recognized as a primary goal of image transformation in the visual system [1–4]. Sparse coding in early visual cortex is achieved by abstracting local oriented spatial frequencies [5] and by excitatory/inhibitory surround modulation [6]. Object responses are thought to be sparse at subsequent processing stages [7, 8], but neural mechanisms for higher-level sparsification are not known. Here, convergent results from macaque area V4 neural recording and simulated V4 populations trained on natural object contours suggest that sparse coding is achieved in midlevel visual cortex by emphasizing representation of acute convex and concave curvature. We studied 165 V4 neurons with a random, adaptive stimulus strategy to minimize bias and explore an unlimited range of contour shapes. V4 responses were strongly weighted toward contours containing acute convex or concave curvature. In contrast, the tuning distribution in nonsparse simulated V4 populations was strongly weighted toward low curvature. But as sparseness constraints increased, the simulated tuning distribution shifted progressively toward more acute convex and concave curvature, matching the neural recording results. These findings indicate a sparse object coding scheme in midlevel visual cortex based on uncommon but diagnostic regions of acute contour curvature.

Results

We recorded the responses of isolated V4 neurons from three rhesus macaque monkeys (*Macaca mulatta*) performing a visual fixation task. During the fixation period, five randomly selected stimuli were flashed in the receptive field (RF) for 750 ms each (with a 250 ms interstimulus interval). The stimuli were presented in a color roughly optimized for the neuron and animated in a small-diameter circular motion pattern against a textured background to establish figure/ground organization.

We used an evolutionary stimulus strategy to minimize bias in the experimental design and maximize the effective stimulus space being explored. This strategy is based on initially random, abstract stimuli evolving probabilistically in response to neural activity, so as to concentrate sampling in relevant regions of shape space. Although our ultimate goal was to

demonstrate a relationship of V4 tuning to natural objects, the use of natural objects themselves as stimuli could have biased the results in that direction. Instead, we sought to show that even abstract stimulus responses are related to the constraints imposed by natural object coding.

For each stimulus lineage, a first generation of random stimuli was created by probabilistic placement of control points defining concatenated Bezier splines, with constraints against collision and looping (Figure 1A; see also Movie S1 available online). These spline-based contours could form either closed figures or partial boundaries that extended beyond the RF (Figure 1B). Subsequent stimulus generations included partially morphed descendants of previous stimuli in addition to newly generated random stimuli (Figure S1). The probability of a given stimulus producing descendants was related to the neural response it evoked (see Supplemental Experimental Procedures). The partial morphing procedure provides a way of breaking up effective stimuli into components and recombining them with other elements to discover which shape parameters are critical for evoking responses.

For the example cell in Figure 1, two independent lineages, starting from different first generations (Figure 1B), ran in parallel. After seven generations, high-response stimuli in both lineages converged toward the same configuration: a convex projection next to a concave indentation, both oriented toward the left/upper left (Figure 1C). Other neurons converged on different contour configurations (Figure S2A). Convergence of multiple lineages is not essential to the analysis of general response bias presented here, so our sample includes 65 neurons studied with only one lineage.

V4 Neural Responses Are Biased toward Acute Contour Curvature

Previous studies have established that V4 neurons are sensitive to both orientation and curvature (the derivative of orientation) [9–14]. Here, we characterized each V4 neuron's response bias in the curvature/orientation domain with a spike-weighted matrix (Figure 2A). To construct this matrix, we determined for each stimulus which bins were occupied by at least one point along its contour, and we summed the stimulus's response rate into those bins. After repeating this for all stimuli, we normalized each sampled bin by the number of samples, interpolated a surface across these bins, and smoothed with a Gaussian function. We then used a smoothed threshold function to reveal the regions consistently associated with high responses (see Supplemental Experimental Procedures for further details). Fitted tuning models, with position dimensions and nonlinear terms, are much more accurate for predicting responses [12, 13, 15], but the spike-weighted matrix is more inclusive, avoids the assumptions and instabilities of model fitting, and is thus better suited for our purpose here of measuring overall population bias. For the example cell, high responses were associated with convex curvature with an orientation range centered near 160° and concave curvature with an orientation range centered near 180° (Figure 2A). This pattern corresponds to the convex/concave configuration visible in the highest-response stimuli (Figure 1C).

*Correspondence: connor@jhu.edu

⁴These authors contributed equally to this work

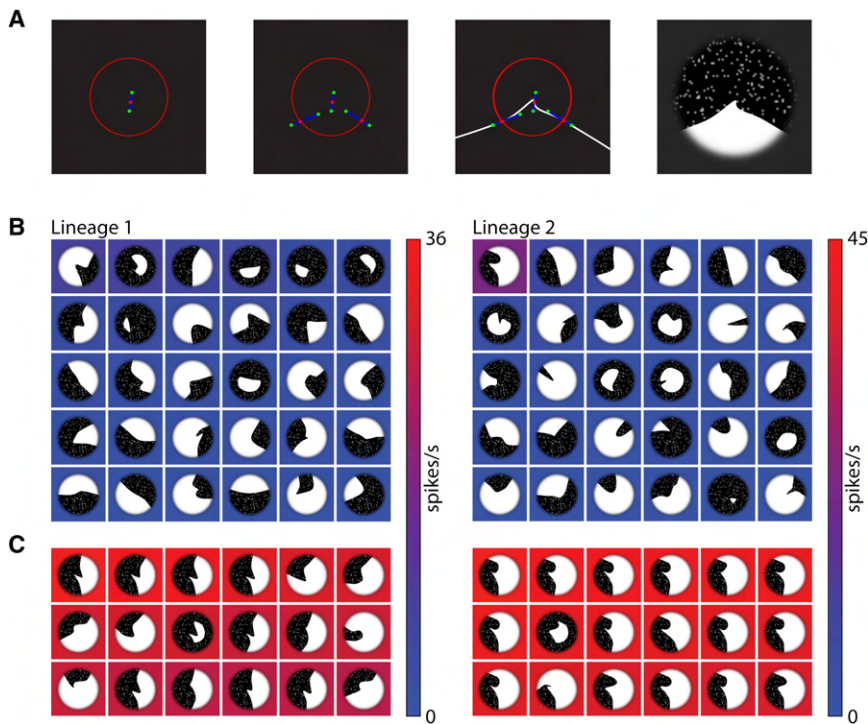


Figure 1. Evolutionary Sampling Method

(A) Stimulus construction based on concatenated Bezier splines. Any Bezier spline segment is defined by two terminal points at either end and two intermediate points that determine the path between the endpoints. Completely random placement of these four control points can produce loops and cusps, so our method for control point placement was probabilistic but constrained. Each of our contours was composed of either two or three (equal probability) Bezier spline segments joined end to end. The example illustrated here is based on two spline segments. The first step in the construction of this stimulus is random placement of a single terminal point (center red dot) within 0.8 receptive field (RF) radii of the RF center. This point was the endpoint at which the two spline segments were joined. Next, two intermediate control points (green dots) were placed on either side of this center, with the constraint that these three points must be collinear (as shown by the blue line; this ensures orientation continuity at the join) and all within 0.8 RF radii of the RF center. Next, the two terminal points (red dots) defining the other ends of the two spline segments were randomly positioned along the RF boundary (red circle). Then, two intermediate control points (green dots) were placed around each of these terminal points, again with the constraint of collinearity. Finally, the contour

was continued with straight lines beyond the RF boundary. (For 10% of randomly generated stimuli, the contour was instead completed within the RF by adding circular segments along the 0.8 RF radius boundary.) If the resulting contour configuration contained loops or cusps, or transgressed the RF boundary, it was “straightened” until these anomalies disappeared. The straightening process involved gradually morphing the flawed configuration toward a straight line in 10% increments. The morphing procedure was terminated at the first step at which no anomalies occurred. This produced a contour as close as possible to the random specification but without anomalies.

Figure/ground polarity was chosen randomly, and the figure area was filled with a color roughly optimal for the neuron being studied (shown here as white). The ground area was filled with texture dots in the same color against a black background. The stimulus colors were constant within 1 RF radius and then faded gradually to the background gray color at a distance of 2 RF radii. The figure was animated with a fine circular motion throughout the presentation period, in order to promote figure ground perception based on the figure moving over the background texture. The orientation of the stimulus remained constant while its center point moved along a circular path with diameter $0.06 \times \text{RF radius}$. This small circular motion was effective for conveying figure/ground organization without substantially changing the position of the stimulus. The motion period was 1 cycle/s, and direction (clockwise versus counterclockwise) and starting point along the circle were chosen randomly on each presentation. The construction sequence and circular movement are shown in [Movie S1](#).

(B) First stimulus generations for lineage 1 (left) and lineage 2 (right) for an example V4 neuron. The aperture in which each stimulus is shown corresponds to the approximate RF boundary. The color of the surrounding square indicates the average response rate, indexed by the scale bar at right. Response rates were calculated by summing spikes over the stimulus presentation period and dividing by the 750 ms stimulus duration. Stimuli are rank ordered by response rate from top left to bottom right.

(C) Highest-response stimuli across seven generations for both lineages.

To estimate V4 response bias at the population level, we averaged spike-weighted matrices across the 165 neurons in our sample. The resulting population matrix (Figure 2B) is clearly anisotropic in the curvature dimension (Kolmogorov-Smirnov, $p < 0.05$), with a strong bias toward acute convex and concave curvature. The correlation between response energy and curvature magnitude was highly significant ($r = 0.98$, $p < 0.001$). Similar results were obtained with a range of threshold functions.

The bias toward acute curvature was robust across independent lineages. For 100 neurons studied with two lineages, one lineage each was used to construct the population matrix shown in Figure 2C. The other lineages for each neuron were used to construct Figure 2D. These two matrices, based on independent data sets, are both strongly biased toward acute curvature. Random sampling of lineages from these 100 neurons was used to construct a 95% confidence interval for the curvature tuning function (Figure 2E). The histogram of correlations between lineages for these neurons in the curvature/orientation domain (Figure 2F, blue bars) had a median

value of 0.37. In many cases, lower correlation was due to partial rotation invariance of V4 responses, which meant that different lineages could evolve stimuli with similar shape but different orientations. Thus, correlations in the curvature domain alone (Figure 2F, red bars) were higher, with a median value of 0.65. The acute curvature bias was also consistent when analysis was restricted to neurons with stronger correlation between lineages and higher maximum responses (Figure S2C).

The acute curvature bias was not inherent to the evolutionary sampling method. We performed simulated experiments on 100 hypothetical neurons emitting random response rates. The average maximum stimulus curvature in generation six was 1.3% lower than in generation one. The average mean curvature was 0.04% higher. These negligible differences could not have produced the strong tuning biases observed here.

Acute Curvature Bias Is Related to Sparse Object Coding

We hypothesized that the V4 acute curvature bias reflects constraints associated with object coding. We therefore

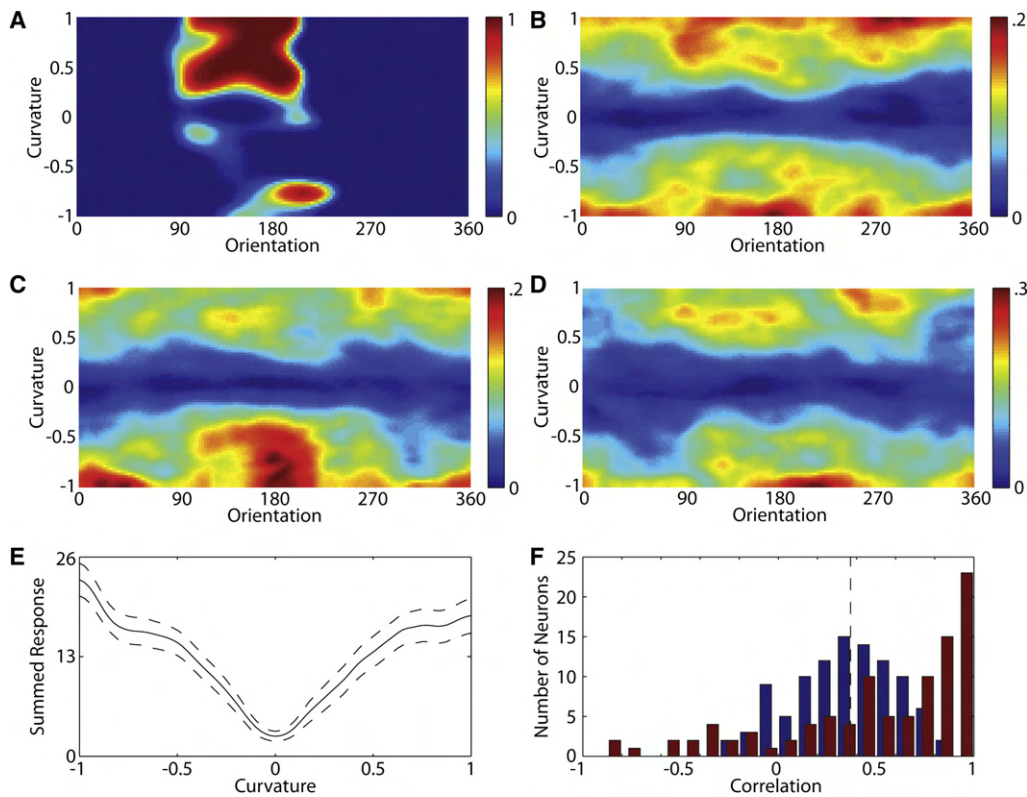


Figure 2. V4 Neural Recording Experiment Results

(A) Spike-weighted response matrix for the Figure 1 example neuron, based on both lineages.
 (B) Average spike-weighted matrix for 165 V4 neurons.
 (C) Average spike-weighted matrix based on the first lineage for each of 100 neurons tested with two independent lineages.
 (D) Average spike-weighted matrix based on the second lineage for each of the same 100 neurons.
 (E) Average curvature response strength, derived by collapsing the average spike-weighted response matrix for 100 neurons with two independent lineages. The dashed line shows the 90% confidence interval for the curvature response strength function. The confidence interval was calculated using a bootstrap procedure. For each iteration of the bootstrap, one of two lineages was randomly selected for each neuron, and the curvature response function was recalculated based on the 100 selected lineages. The dashed lines show the 5th and 95th percentile points in the resulting distribution of curvature response values at each point.
 (F) Distribution of correlations between separate spike-weighted response matrices for 100 neurons with two independent lineages. The distribution of correlations for the complete response matrices is shown by the blue bars. The median correlation (dashed vertical line) was 0.37. (The correlation for the Figure 1 example cell was 0.57.) In many cases, correlation was low because of differences in the orientation dimension, reflecting partial rotation invariance in V4. Therefore, the response matrices were collapsed onto the curvature dimension and correlations were recalculated (red bars). This distribution has a strong peak near 1 and a median value of 0.65. Thus, curvature tuning was robust across lineages.

attempted to explain the bias by simulating V4 population coding of natural objects. We found that object discrimination training alone did not produce the observed distribution. However, simulations with an additional sparse coding constraint were biased toward acute curvature in the same way as V4 neural responses.

Each simulation comprised 100 model neurons with V4-like Gaussian tuning for contour curvature, orientation, and object-relative position. We have previously shown that this type of model is successful in explaining neural response variance across large, diverse stimulus sets and supports reconstruction of complex stimuli from neural responses [12, 13, 15]. Here, we constructed a population of these previously validated models and allowed only their mean tuning positions (in the curvature, orientation, and position dimensions) to vary during training. Figure 3A shows three example tuning profiles in red, green, and blue. The curvature/orientation domain is analogous to that in Figure 2. Relative position is specified as angle with respect to object center, with 0 representing contours on the

right side of the object, 90 representing contours at the top, etc. For tractability, tuning for retinotopic position (i.e., RF position) was ignored. Instead, each model neuron effectively stood for a subpopulation of V4 neurons, with different RF locations tiling retinotopic space, but with equivalent tuning for curvature, orientation, and object-relative position.

A set of 10,966 natural object photographs from the Hemera Photo-Objects database was used for training and testing. Each object boundary was fit with a continuous contour (Figure S3) and numerically described as a set of sample points specified by curvature, orientation, and angular relative position (Supplemental Experimental Procedures). Each model neuron responded according to how close any of these contour fragments fell to its curvature/orientation/position tuning peak. For example, the table stimulus (Figure 3B) would evoke a strong response from all three example neurons, based on its correspondingly colored contour fragments. Other objects would evoke different response patterns depending on their constituent contour fragments.

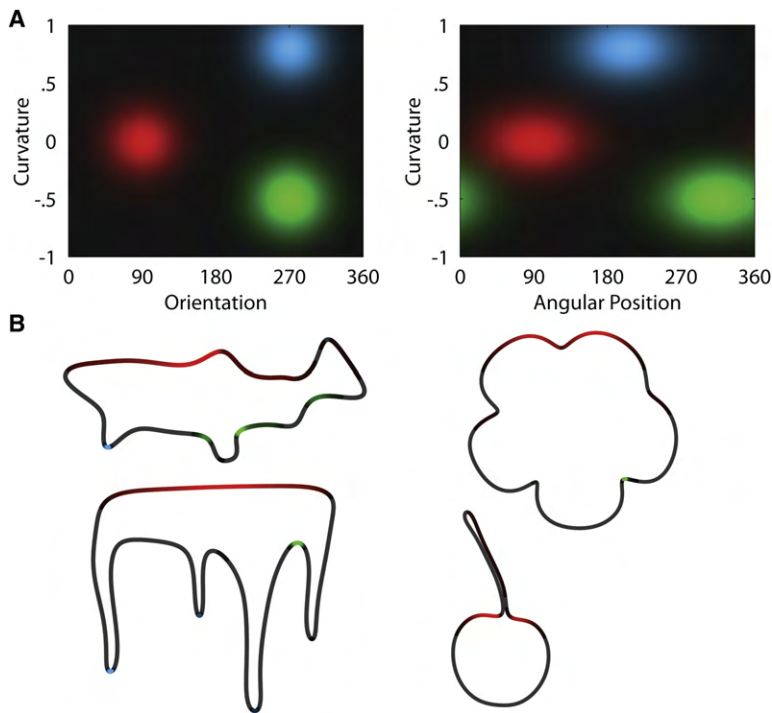


Figure 3. Simulation of V4 Object Discrimination
(A) Tuning functions for three example model V4 neurons (red, green, and blue), plotted in the curvature/orientation domain (left) and the relative position domain (right). Standard deviations in each dimension were constrained to a constant value. (B) Example object contours. Object contour regions with curvature/orientation/position values close to the example neuron peaks are tinted the corresponding colors.

The training objective was to minimize estimated pairwise object discrimination error by an ideal observer. Pairwise error was defined as the Gaussian error function (erfc) of interobject distance in the 100-dimensional neural response space. Error was minimized over tuning parameters by constrained optimization in MATLAB. For the sake of tractability, discrimination error was estimated for a subsample of 100,000 randomly selected object pairs. A new subsample was selected every 300 iterations, over the course of 7,500 iterations in total (see [Supplemental Experimental Procedures](#) for further details). We ran a total of five simulations, starting from different sets of randomly generated tuning parameters.

For the resulting 500 model neurons, identification accuracy (based on maximum likelihood) for five sets of 100 objects not used in training averaged 100%. The curvature/orientation tuning peaks for these 500 model neurons are shown in [Figure 4A](#) (circles). Tuning peaks are broadly distributed across the orientation dimension but clustered in the curvature dimension near flat and shallow curvatures. This pattern is the opposite of the V4 neural response distribution, which is replotted for comparison in [Figure 4](#) (colored surfaces). This suggests that the V4 neural response distribution is not optimized for object discrimination alone.

Simulated tuning distributions became similar to the V4 neural response distribution when sparseness was added as an additional constraint. In these simulations, the error function to be minimized was a weighted sum of discrimination error and a standard index of response density (RD) [4, 6], measured across the model neuron population for each object and then averaged across objects:

$$RD = \left\langle \frac{\langle x_{i,j} \rangle_i^2}{\langle x_{i,j}^2 \rangle_i} \right\rangle_j,$$

where $x_{i,j}$ is the response of the j th model unit to the i th object. RD is inversely related to sparseness and has a maximum value

of 1. Minimizing RD, as defined here, increased population sparseness, i.e., sparseness of the expected distribution of neural responses to a given image [16]. The simulated tuning distributions in [Figures 4B–4D](#) are based on increasing RD weights in the error function. With no sparseness constraint ([Figure 4A](#)), average RD across images was 0.80; with the strongest weighting ([Figure 4D](#)), average RD fell to 0.11. As RD fell and sparseness increased, simulated tuning progressively concentrated toward acute convex and concave curvature, matching the observed V4 response distribution. The strongest correlation with the observed V4 response distribution was obtained with an RD value of 0.11 ($r = 0.55$). Identification accuracy remained high at all three sparseness levels (100%, 100%, and 97%, respectively).

These modeling results indicate that the observed curvature bias in area V4 would produce much sparser population responses to objects while retaining sufficient information for object discrimination. The observed V4 distribution is very different from the distribution expected on the basis of object discrimination alone. The V4 distribution is strongly biased toward acute curvature, whereas the expected distribution is strongly biased toward flat contours. This difference has a strong effect on sparseness of population responses to objects. The acute curvature bias, as found in V4, produces an average RD value in the range of 0.1, whereas the flat contour bias in the expected distribution produces an average RD value in the range of 0.8. Thus, the V4 coding scheme appears markedly sparse relative to what would be expected if object discriminability were the only constraint on intermediate visual processing.

Discussion

Our results provide the first description of a sparse coding scheme in area V4, a major intermediate stage in ventral pathway visual cortex. Sparse coding is considered an important goal of sensory transformation because it increases representational capacity [4] and reduces metabolic energy requirements [17]. It is reasonable to speculate that the V4 coding scheme evolved in response to these adaptive advantages, though sparseness is not the only constraint that might produce an acute curvature bias. At the modeling level, we found that the acute curvature bias was not produced by minimizing various rearrangements of the terms in the RD expression ([Figure S4A](#)), but there must be many mathematical constraints that would produce a similar bias. On the evolutionary level, there could be many other advantages to selective representation of acute curvature, perhaps relating to higher ecological relevance for object parts with acute curvature. Thus, there may be some other constraint that drove

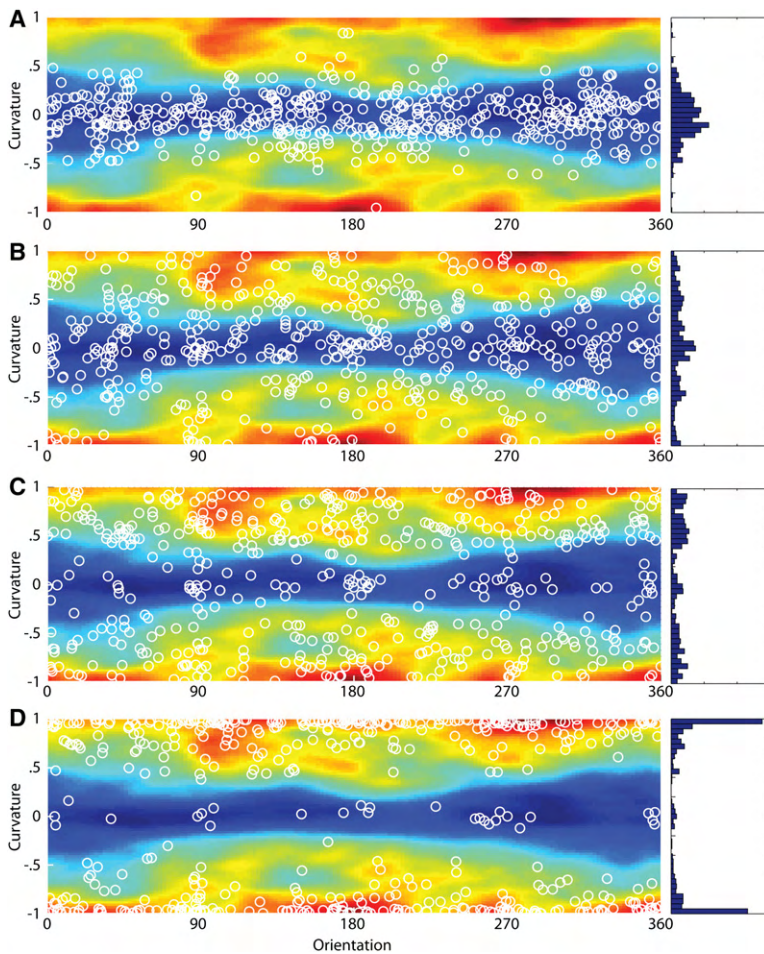


Figure 4. Simulation Population Results

In each panel, circles represent tuning function peaks in the curvature/orientation domain for 500 model neurons drawn from five simulations. These are plotted on top of the average spike-weighted V4 neural response matrix from Figure 2B for comparison. The histograms at right show the distribution of model units across the curvature dimension.

(A) Simulation results based on discrimination only, with no sparseness constraint. The response density (RD) value across both training and testing stimuli was 0.80.

(B) Lowest sparseness constraint. RD error term weight, relative to discrimination error term, was 0.01. The resulting RD value was 0.36.

(C) Intermediate sparseness constraint. Relative RD error term weight was 0.1. The resulting RD value was 0.22.

(D) Highest sparseness constraint. Relative RD error term weight was 1. The resulting RD value was 0.11.

the acute curvature bias. However, regardless of how and why it evolved, the curvature bias seems likely to produce sparser object responses in area V4, and sparseness has strong implications for computational efficiency, metabolic efficiency, and memory storage.

This conclusion derives from the assumption that our modeling results are appropriate for interpreting the observed acute curvature bias in area V4. The models were closely based on previously validated models of intermediate visual neurons [12, 13, 15], but mechanisms of intermediate and higher-level vision remain controversial, and no current model can be regarded as definitive. Area V4 neurons might operate in ways not captured by our models that affect sparseness. Moreover, we cannot make any claim, based on these analyses, about the absolute level of sparseness in V4 responses. We are only claiming that, given the tuning bias toward acute curvature, V4 responses are likely to be sparser than they would be without such a tuning bias. This seems logical, apart from any specific modeling results, given the low frequency of curved contours in relation to flat contours (Figure S4B). Neurons tuned for less common image elements are bound to respond less frequently. V4 neurons show a strong tuning bias for acute curvature. Given the relatively low frequency of acute curvature, these neurons are bound to respond more sparsely than neurons without such a bias.

In early visual cortex, sparse coding is achieved by exploiting local statistical regularities in natural images to reduce

redundancy of neural signals [1, 2]. Gabor-like RF structures in V1 reduce redundancies associated with local spatial frequency correlations in natural images [5]. Nonlinear interactions with the nonclassical surround exploit image correlations that extend beyond the classical RF [6]. Coding in early visual cortex is constrained to minimize information loss, because the rest of the brain gets most of its detailed visual information directly or indirectly from V1 [18]. Redundancy reduction based on local statistical regularities can be achieved without substantial loss of information.

Sparsification in midlevel cortex might require other mechanisms given the different constraints of intermediate shape processing. Neurons in midlevel visual cortex integrate information across larger RFs and nonclassical surrounds [18].

Statistical correlations are bound to be lower on this larger scale, because physical relatedness between object parts is statistically weaker across greater distances. Thus, redundancy reduction may not be an option for sparsification at this scale.

However, midlevel cortex is also more specialized, with less need for complete preservation of image information and more scope for emphasizing information required for specific aspects of visual perception [18]. V4, in particular, is part of the ventral pathway [18, 19], which emphasizes shape, color, and texture in the service of object perception. Given this specialization, further sparsification could be achieved by biasing representation toward image features with high object information content but lower probability of occurrence. In this way, a given object could be represented in terms of a small number of uncommon but diagnostic elements.

This alternate kind of sparse coding strategy appears to be implemented in V4 by emphasizing the representation of acute contour curvature, which is appropriately uncommon and diagnostic. Acute curvature was approximately an order of magnitude less common than flat or shallow curvature in our natural object set (Figure S4B). This reflects the fact that, on the scale of visual perception, natural objects have mostly smooth rather than highly intricate boundaries. Thus, sparse coding simulations based on primarily acute curvature tuning (Figures 4C and 4D) had low response densities (0.22 and 0.11, respectively). In contrast, nonsparse simulations based on tuning for

more common flat and/or shallow contour regions had high response densities (Figure 4A, 0.80; see also Figure S4C). At the same time, regions of acute curvature are still highly informative about object identity. In our simulations, accuracy remained high for the sparsest condition (RD = 0.11, accuracy = 97%) even when the remaining low-curvature model neurons ($-0.4 < c' < 0.4$) were removed (accuracy = 85%). Curved contour regions are also perceptually salient [20–22] and more perceptually informative than flat contours [1, 23].

Bias toward representation of uncommon features with specialized information content could be a general strategy for sparse coding in higher-level cortex. Some evidence suggests that object coding is sparse at the final stages of the ventral pathway in inferotemporal (IT) cortex and medial lobe temporal structures like the hippocampus [7, 8, 24]. Sparseness at these higher levels could be achieved by selectivity for more complex features [15, 25, 26] with even higher information content. Bias toward tuning for acute curvature, which has been demonstrated in IT [27], might also enhance response sparseness at this level. Alternatively, IT cortex might be optimized for discrimination at the expense of sparseness [28]. The combined simulation and adaptive search strategy used here might help to elucidate coding strategies in higher-level visual cortex as well as in other sensory modalities.

Supplemental Information

Supplemental Information includes four figures, Supplemental Experimental Procedures, and one movie and can be found with this article online at doi:10.1016/j.cub.2011.01.013.

Acknowledgments

We thank Zhihong Wang, William Nash, and William Quinlan for technical assistance. This work was funded by the National Eye Institute.

Received: July 22, 2010

Revised: November 29, 2010

Accepted: January 5, 2011

Published online: February 10, 2011

References

1. Attneave, F. (1954). Some informational aspects of visual perception. *Psychol. Rev.* 61, 183–193.
2. Barlow, H.B. (1959). Sensory Mechanisms, the Reduction of Redundancy, and Intelligence, NPL Symposium on the Mechanization of Thought Process 10 (London: H.M. Stationery Office), pp. 535–539.
3. Field, D.J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *J. Opt. Soc. Am. A* 4, 2379–2394.
4. Treves, A., and Rolls, E.T. (1991). What determines the capacity of autoassociative memories in the brain? *Network* 2, 371–397.
5. Olshausen, B.A., and Field, D.J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature* 381, 607–609.
6. Vinje, W.E., and Gallant, J.L. (2000). Sparse coding and decorrelation in primary visual cortex during natural vision. *Science* 287, 1273–1276.
7. Young, M.P., and Yamane, S. (1992). Sparse population coding of faces in the inferotemporal cortex. *Science* 256, 1327–1331.
8. Quiroga, R.Q., Kreiman, G., Koch, C., and Fried, I. (2008). Sparse but not ‘grandmother-cell’ coding in the medial temporal lobe. *Trends Cogn. Sci.* 12, 87–91.
9. Desimone, R., and Schein, S.J. (1987). Visual properties of neurons in area V4 of the macaque: Sensitivity to stimulus form. *J. Neurophysiol.* 57, 835–868.
10. Gallant, J.L., Braun, J., and Van Essen, D.C. (1993). Selectivity for polar, hyperbolic, and Cartesian gratings in macaque visual cortex. *Science* 259, 100–103.
11. Pasupathy, A., and Connor, C.E. (1999). Responses to contour features in macaque area V4. *J. Neurophysiol.* 82, 2490–2502.
12. Pasupathy, A., and Connor, C.E. (2001). Shape representation in area V4: Position-specific tuning for boundary conformation. *J. Neurophysiol.* 86, 2505–2519.
13. Pasupathy, A., and Connor, C.E. (2002). Population coding of shape in area V4. *Nat. Neurosci.* 5, 1332–1338.
14. Wilkinson, F., James, T.W., Wilson, H.R., Gati, J.S., Menon, R.S., and Goodale, M.A. (2000). An fMRI study of the selective activation of human extrastriate form vision areas by radial and concentric gratings. *Curr. Biol.* 10, 1455–1458.
15. Brincat, S.L., and Connor, C.E. (2004). Underlying principles of visual shape selectivity in posterior inferotemporal cortex. *Nat. Neurosci.* 7, 880–886.
16. Willmore, B., and Tolhurst, D.J. (2001). Characterizing the sparseness of neural codes. *Network* 12, 255–270.
17. Levy, W.B., and Baxter, R.A. (1996). Energy efficient neural codes. *Neural Comput.* 8, 531–543.
18. Felleman, D.J., and Van Essen, D.C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* 1, 1–47.
19. Ungerleider, L.G., and Mishkin, M. (1982). Two cortical visual systems. In *Analysis of Visual Behavior*, D.J. Ingle, M.A. Goodale, and R.J.W. Mansfield, eds. (Cambridge, MA: MIT Press), pp. 549–586.
20. Treisman, A., and Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychol. Rev.* 95, 15–48.
21. Wolfe, J.M., Yee, A., and Friedman-Hill, S.R. (1992). Curvature is a basic feature for visual search tasks. *Perception* 21, 465–480.
22. Wilson, H.R., Wilkinson, F., and Asaad, W. (1997). Concentric orientation summations in human form vision. *Vision Res.* 37, 2325–2330.
23. Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychol. Rev.* 94, 115–147.
24. Baddeley, R., Abbott, L.F., Booth, M.C.A., Sengpiel, F., Freeman, T., Wakeman, E.A., and Rolls, E.T. (1997). Responses of neurons in primary and inferior temporal visual cortices to natural scenes. *Proc. Biol. Sci.* 264, 1775–1783.
25. Kobatake, E., and Tanaka, K. (1994). Neuronal selectivities to complex objects features in the ventral visual pathway of the macaque cerebral cortex. *J. Neurophysiol.* 71, 856–867.
26. Yamane, Y., Carlson, E.T., Bowman, K.C., Wang, Z., and Connor, C.E. (2008). A neural code for three-dimensional object shape in macaque inferotemporal cortex. *Nat. Neurosci.* 11, 1352–1360.
27. Kayaert, G., Biederman, I., Op de Beeck, H.P., and Vogels, R. (2005). Tuning for shape dimensions in macaque inferior temporal cortex. *Eur. J. Neurosci.* 22, 212–224.
28. Rolls, E.T., and Tovee, M.J. (1995). Sparseness of the neuronal representation of stimuli in the primate temporal visual cortex. *J. Neurophysiol.* 73, 713–726.

Current Biology, Volume 21

Supplemental Information

A Sparse Object Coding Scheme in Area V4

Eric T. Carlson, Russell J. Rasquinha, Kechen Zhang, and Charles E. Connor

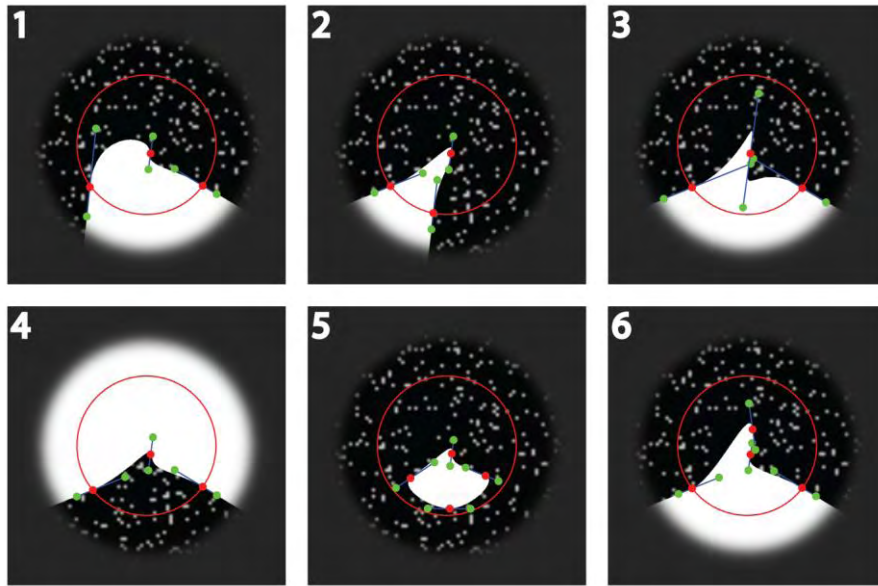


Figure S1, Related to Figure 1. Partial Morphing of Contour Stimuli

Partially morphed descendants of higher response stimuli were designed to preserve some parameters while changing others in order to discriminate which parameters were critical for neural responses. Since the partially morphed control point configuration could produce loops, cusps, or RF boundary transgressions, the same straightening procedure described in the Figure 1 legend was used to cure these flaws. The example stimuli in this figure are all morphed descendants of the ancestor stimulus in Figure 1. The following morphing procedures occurred with the indicated probabilities:

- Randomly change the positions of the intermediate control points (green dots) surrounding a central join between spline segments (red dot) (13.6%). This changes local contour orientation and curvature. The same collinearity constraint applies. (See example 1.)
- Randomly change the positions of a subset of control points (27.2%). (See example 2.)
- Randomly change curvature throughout the stimulus, by changing how close the intermediate points (green dots) are to the terminal points (red dots) (13.6%). Closer intermediate points produces sharper curvature near the terminal points. (See example 3.)
- Invert figure polarity (so that figure becomes ground and vice versa) (2.7%). (See example 4.)
- Convert a figure that extends beyond the RF border to a figure that completes within the RF (2.7%). (See example 5.)
- Convert a randomly chosen Bezier spline segment into two segments by adding the necessary randomly-placed control points. This increases the complexity of the stimulus. (6.8%). (See example 6.)
- Randomly replace two Bezier segments with one, by removing the appropriate control points (13.6%). This has the effect of simplifying contour shape.
- Morph the contour in the direction of a straight line by a random amount (13.6%).
- Randomly rotate the stimulus about the RF center (3.4%).
- Convert the filled figure to a thin line with width equal to 0.5° (2.7%).

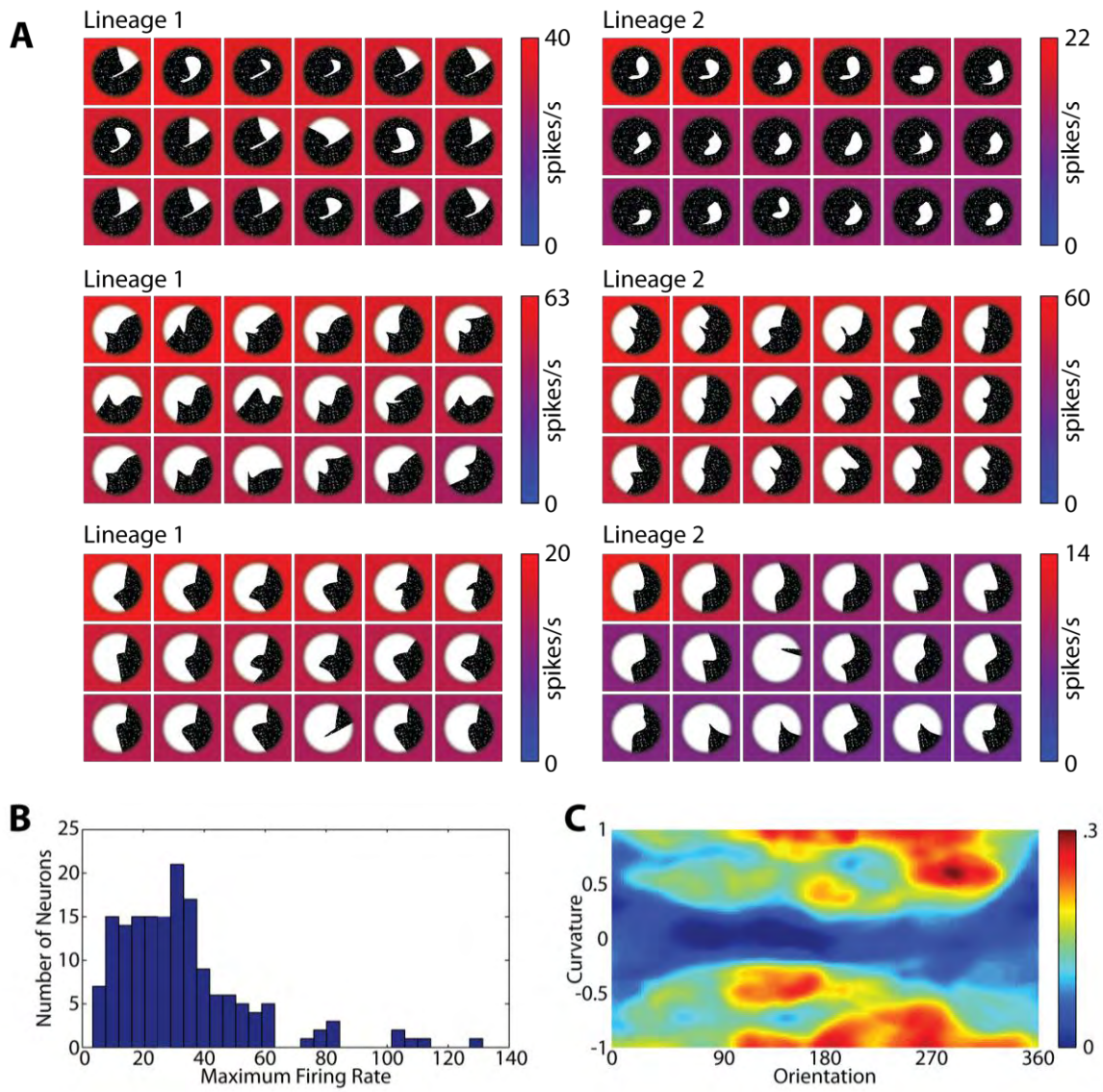


Figure S2, Related to Figure 2. Results of Evolutionary Stimulus Sampling

(A) Three example results illustrating shape convergence between two independent lineages. The correlations between separate spike-weighted response matrices for these three examples were 0.38, 0.65, and 0.39 (top to bottom). (The corresponding correlations in the curvature dimension were 0.54, 0.43, and 0.95.)

(B) Distribution of maximum spike rate responses for all neurons.

(C) A version of Figure 2B based on 28 neurons with two independent lineages and higher convergence between spike-weighted response matrices (correlation above median) and higher response rates (> 30 spikes/s).

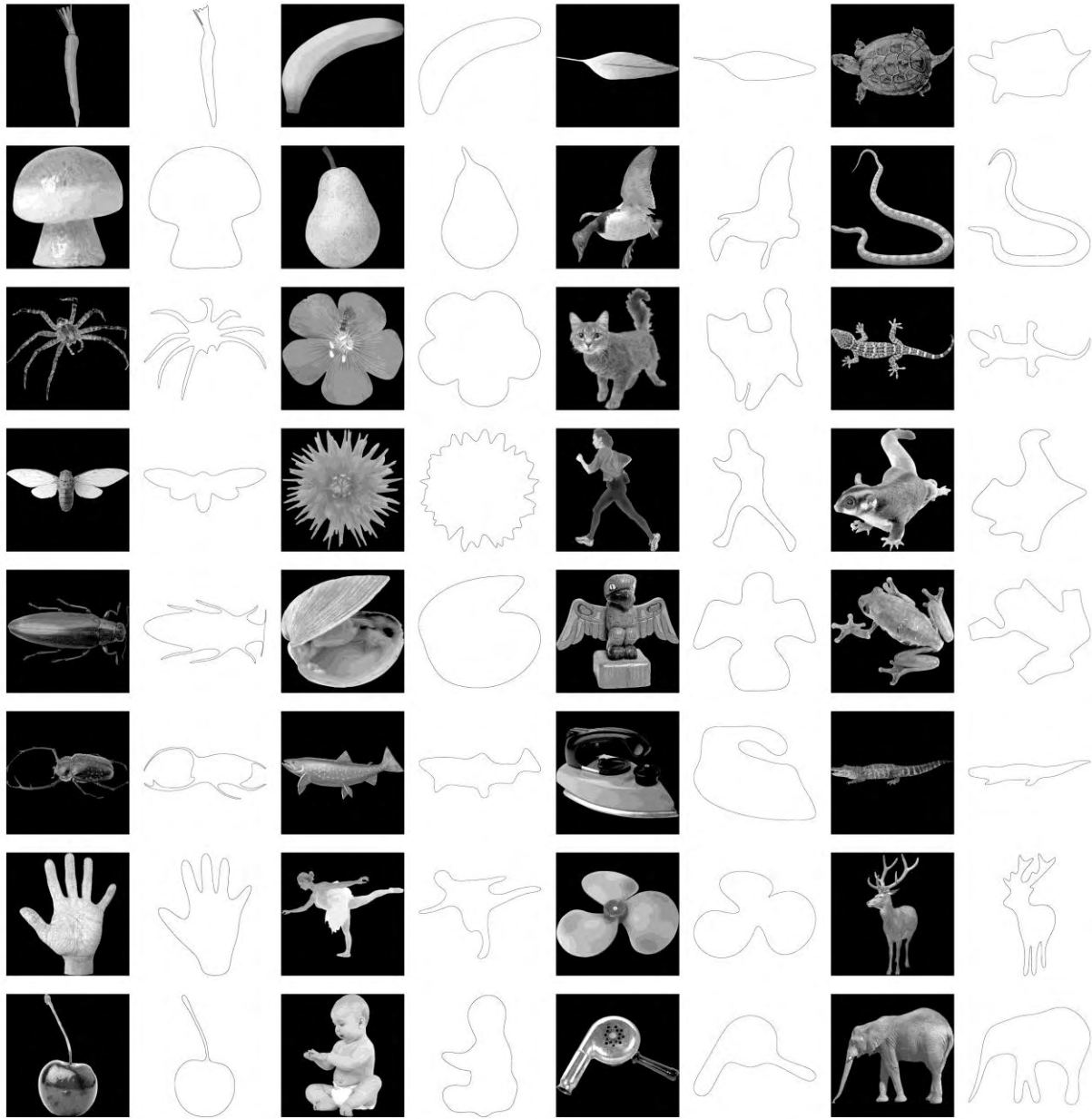


Figure S3, Related to Figure 3. Example Training Stimuli

The borders of the photographic objects were fit with spline functions using the csaps function in MATLAB.

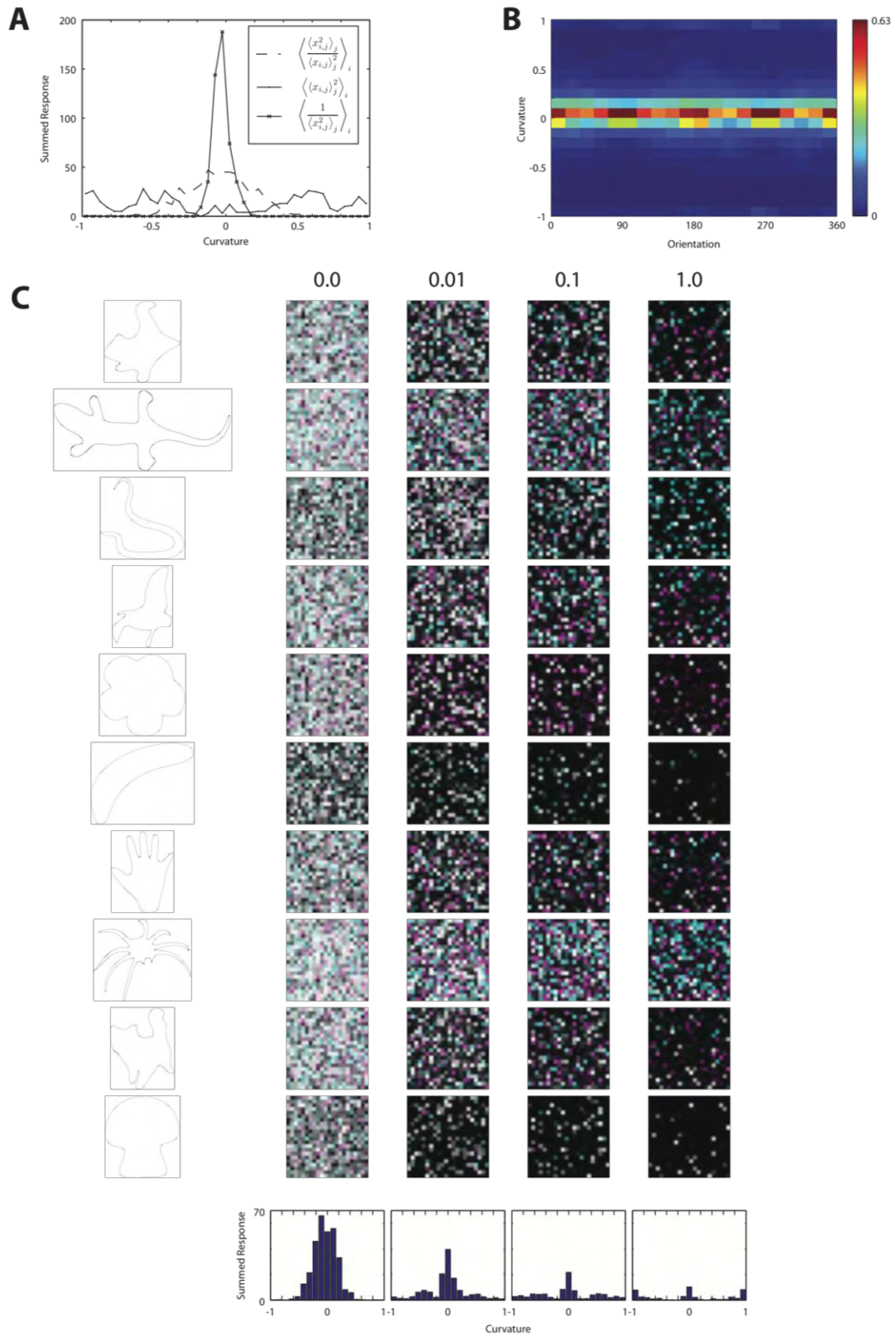


Figure S4, Related to Figure 4.

Figure S4, Related to Figure 4. Sparse Coding and Acute Curvature Tuning

(A) Distributions of curvature tuning in additional simulations in which minimization of RD was replaced with minimization of a related factor. In one case, the factor being minimized was the inverse of RD. This would tend to produce a non-sparse response distribution:

$$1/RD = \left\langle \frac{\langle x_{i,j}^2 \rangle_j}{\langle x_{i,j} \rangle_j^2} \right\rangle_i$$

In the second case, the factor being minimized was the numerator of the RD expression, i.e. expected value squared. This would tend to minimize mean response:

$$\text{numerator} = \left\langle \frac{\langle x_{i,j} \rangle_j^2}{1} \right\rangle_i$$

In the third case, the factor being minimized was the denominator of the RD expression, i.e. the inverse of expected squared value. This would tend to maximize the second moment (variance) with respect to 0:

$$\text{denominator} = \left\langle \frac{1}{\langle x_{i,j}^2 \rangle_j} \right\rangle_i$$

In each case, the weight of the alternative minimization factor was 1. None of these rearrangements of the components in the RD expression produced an acute curvature tuning bias. Thus, among the alternatives we considered, the acute curvature bias was specific to the sparseness constraint, by which mean response is minimized at the same time that second moment is maximized.

(B) Frequency distribution in the curvature/orientation domain for 10966 objects used in the V4 population simulation. Flat and shallow contour curvature is far more common than acute curvature.

(C) Changes in model neuron responses with increasing sparseness constraint. Each matrix of colored squares represents the responses of 484 neurons to an example object (shown at left of each row) under a given sparseness constraint (indicated above each column). Each cell in the matrix represents the response of one neuron. Brightness indicates response strength, and color indicates curvature tuning. White corresponds to tuning for flat contours. Tuning for increasingly acute convex curvature is indicated by increasing magenta saturation. Tuning for increasingly acute concave curvature is indicated by increasing cyan saturation. As the sparseness constraint increases (left to right), high (bright) responses become more rare, and high responses are increasingly dominated by acute curvatures (magenta and cyan). The histograms in the bottom row show the distribution of response strength across curvature, summed across 10966 stimuli, for the different sparseness conditions. As the sparseness constraint increases, total response power drops, and remaining response power is increasingly distributed toward acute convex and concave curvature. Even with the strongest sparseness constraint, there remains a response peak near 0 curvature, although the number of model units tuned in this range is small (Figure 4D). The persistence of this 0 curvature response peak reflects the high frequency of flat contours and demonstrates how strongly neurons in this tuning range contribute to response density. This explains why increasing the sparseness constraint produced progressively fewer model units in the flat curvature range.

Supplemental Experimental Procedures

V4 Neural Recording

We recorded the activity of electronically isolated neurons from the lower-field visual representation in dorsal V4 of two female rhesus monkeys (*Macaca mulatta*). Monkeys were seated in front of a computer monitor at a distance of 50 cm with the head immobilized by means of a custom-built titanium head post surgically attached to the skull with orthopedic screws and dental acrylic. Eye position was monitored with an infrared eye tracker (ISCAN). During recordings, the animals were required to maintain fixation on a 0.1° spot within a 0.75° radius for 5 s in order to obtain a juice reward. Neural activity was measured with epoxy-coated tungsten electrodes (A-M Systems, Carlsburg, Washington, USA) inserted with a custom-built microdrive assembly. All animal procedures conformed to National Institutes of Health and USDA guidelines and were approved by the Johns Hopkins University Animal Care and Use Committee.

The RF center was hand plotted using shape stimuli under the experimenter's control. The RF diameter was estimated using a formula based on data from Gatass et al., 1988 [1] showing that average V4 RF diameter equals approximately $1^\circ + 0.625 \times \text{RF eccentricity}$. The neuron was tested with seven colors adjusted to an approximate luminance of 20 cd/m² (white, red, green, blue, yellow, cyan, magenta) and black (0 cd/m²) against a background gray (2.5 cd/m²).

Adaptive Search Procedure

Each stimulus generation consisted of 30 stimuli. The first generation of stimuli was defined according to the random procedure described in Figure 1A. For some cells two separate lineages were tested concurrently, with their stimuli randomly interleaved. Stimuli were flashed in random order for a total of 5 repetitions per stimulus. The average response to each stimulus was calculated by counting spikes within the stimulus presentation window, which was 750 ms, dividing by stimulus duration, and averaging across repetitions.

Subsequent generations contained 25–75% randomly generated stimuli. The percentage of random stimuli approached 25% as maximum evoked responses approached $4 \times$ spontaneous firing rate, which was calculated by averaging across blank stimulus intervals randomly interleaved with stimulus presentations.

The remaining stimuli in subsequent generations were partially morphed descendants of previously tested ancestor stimuli (see Figure S1). The population of potential ancestors was recalculated for each successive generation based on all preceding generations. Stimuli were included in the potential ancestor pool only if they evoked an average response greater than either $2 \times$ spontaneous firing rate or $0.5 \times$ maximum evoked response, whichever threshold was higher. Probability of a given stimulus producing descendants was proportional to its average response rate minus the threshold value. The resulting probability distribution was sampled with replacement until all the required descendants for the next generation were defined.

The experiment continued as long as the neuron remained well isolated. Lineages ranged from 4 generations (120 stimuli) to 12 generations (360 stimuli). The majority of neurons in our sample (100) were studied with two lineages running in parallel.

Spike-Weighted Response Matrix Construction

Curvature of stimulus contours was originally calculated as inverse radius in units equivalent to 1 RF radius. Since the range of curvature is infinite, we squashed the curvature range with the following function:

$$c' = \frac{2.0}{1 + e^{-a \cdot c}} - 1.0$$

where c is curvature, c' is squashed curvature, and $a = 0.05$. Squashed curvature runs from -1.0 (acute concave) through 0.0 (flat) to 1.0 (acute convex).

Bin width was 2.5° in the orientation dimension and 0.025 in the squashed curvature dimension. For each stimulus, the contour was densely sampled. The average response to the stimulus was summed into each bin occupied by at least one sample point along the contour. After all stimuli were processed, the value in each bin was divided by the number of stimuli contributing to that bin's sum. Thus, each bin's

value represented the average response to stimuli containing at least one point with the corresponding curvature/orientation value pair. The surface was linearly interpolated using the TriScatteredInterp function in MATLAB and then smoothed by convolving with a Gaussian function ($\sigma_{\text{orientation}} = 20^\circ$, $\sigma_{\text{curvature}} = 0.1$), and the minimum to maximum range was normalized to run from 0 to 1. The result was thresholded with a Gaussian-smoothed ($\sigma = 0.05$) step function with the transition at 0.8.

Photographic Stimulus Processing

Object photographs were selected from the HEMERA Photo-Objects Database. Objects were constrained to have single, topologically circular boundaries (no disconnected parts, no holes).

The object boundary was extracted from the image's transparency map and smoothed with the spline fit function csaps in MATLAB, which minimizes (a) squared error between original and spline fit and (b) average squared value of the second derivative (curvature) of the spline. The parameter, p , which specifies relative weight of the two terms, was set at 10^{-6} , to remove details that would be imperceptibly small under normal viewing conditions.

For each point on the smoothed object boundary, orientation was calculated as the direction of a surface normal vector pointing away from the object interior. Angular position of each point on the smoothed object boundary was calculated with respect to object center of mass. Curvature was calculated as inverse radius, with radius units equal to $0.5 \times$ maximum length (the largest distance between any two points on the border). The curvature dimension was squashed using the same formula described above. Since the radius unit used in calculating curvature of neural recording stimuli was RF radius, the implicit assumption in comparing the two analyses was that each object's maximum length corresponded to RF diameter. This equivalence is arbitrary but within the range of natural perceptual experience. Under naturalistic conditions, object sizes would vary with respect to RF size, but replicating this variation would have made the simulations intractable. Given that curvature is squashed, the results would be robust to changes in the absolute curvature equivalence.

To make the simulations more tractable, the object boundary point values were subsampled by dividing the boundary into segments of length 0.1 RF radius, with 50% overlap between successive segments. For each segment, curvature was set to the maximum value within that segment. Orientation was set to the orientation of the sample point (within the segment) with maximum curvature. Relative angular position was set to the position of the sample point (within the segment) with maximum curvature.

Simulation Training

Each simulation comprised 100 model V4 neurons defined by Gaussian tuning in the orientation, curvature, and relative angular position dimensions. For the sake of tractability, tuning function standard deviations were fixed ($\sigma_o = 30^\circ$, $\sigma_c = 0.2$, $\sigma_p = 45^\circ$). Only the three Gaussian means (μ_c , μ_o , μ_p), were optimized during training. These values were initialized randomly. The response of a given model neuron to a given stimulus was the maximum across sample points for the following equation:

$$R(s) = \exp\left(-\frac{(c_s - \mu_c)^2}{2\sigma_c^2} - \frac{(o_s - \mu_o)^2}{2\sigma_o^2} - \frac{(p_s - \mu_p)^2}{2\sigma_p^2}\right)$$

where $R(s)$ is the response to sample point s , c_s is the curvature at sample point s , o_s is the orientation at sample point s , and p_s is the relative angular position of sample point s .

The simulations were optimized for discriminability of objects based on response pattern across the 100 model V4 neurons. Discrimination error was averaged across object pairs with an assumption of Gaussian response noise:

$$DE = \left\langle \operatorname{erfc}\left(\frac{\|x_a^r - x_b^r\|}{2\sqrt{2}\sigma_{DE}}\right) \right\rangle_{a,b}$$

where DE is discrimination error, x_a is the 100-dimensional neural response vector for object a , x_b is the response vector for object b , and σ_{DE} is the standard deviation of the Gaussian response noise function,

set at 0.2 in the simulations reported here. This formula is proportional to the Bayes error rate of an ideal observer, assuming equal prior probability for different stimulus objects and independent Gaussian response noise with equal variances for different neurons [2]. An additional error term was used to prevent similar tuning functions between neurons:

$$SE = \left\langle \operatorname{erfc} \left(\frac{\|y_p^r - y_q^r\|}{2\sqrt{2}\sigma_{SE}} \right) \right\rangle_{p,q}$$

where SE is similarity error, y_p is the response vector for the pth cell across all objects used in training, y_q is the response vector for the qth cell across all objects used in training, and σ_{SE} is the standard deviation of the Gaussian function controlling cell similarity, set at 2 in the simulations reported here. The overall error term to be minimized was:

$$E = \ln(\mathbf{DE} + SE)$$

In some simulations, RD (response density, see main text) was included in the error term to increase sparseness of neural response patterns:

$$E = \ln(\mathbf{DE} + SE + \lambda \mathbf{RD})$$

The sparseness weighting factor λ was set to 0.01, 0.1, and 1 in the three sparse simulations presented here.

The error function was minimized using a nonlinear constrained optimization algorithm (fmincon, MATLAB). For each simulation, a random set of 3000 objects were held out for testing, leaving 7966 objects for training. From these, a random set of 3000 objects was chosen, and for these objects a random set of 100,000 pairings was chosen for optimization of E. After each 300 evaluations of the error function, a new set of objects and pairings was chosen. This was repeated 25 times, for a total of 7500 evaluations. Five simulations were run for each error condition, producing a population of 500 model neurons in each case.

Supplemental References

1. Gattass, R., Sousa, A.P., and Gross, C.G. (1988). *J. Neurosci.* 8, 1831–1845.
2. Duda, R. O., Hart, P. E., and Stork, D. G. (2001). *Pattern Classification*, 2nd ed., New York: Wiley, section 3.7.1.