

# Adaptive shape processing in primary visual cortex

Justin N. J. McManus<sup>a</sup>, Wu Li<sup>b</sup>, and Charles D. Gilbert<sup>a,1</sup>

<sup>a</sup>Laboratory of Neurobiology, The Rockefeller University, New York, NY 10065; and <sup>b</sup>State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing 100875 China

This contribution is part of the special series of Inaugural Articles by members of the National Academy of Sciences elected in 2006.

Contributed by Charles D. Gilbert, April 18, 2011 (sent for review March 4, 2011)

**The ability to derive meaning from complex sensory input requires the integration of information over space and time, as well as cognitive mechanisms to shape that integration. We studied these processes in the primary visual cortex (V1), where neurons are thought to integrate visual inputs along contours defined by an association field (AF). We recorded extracellularly from single cells in macaque V1 to map the AF, by using an optimization algorithm to find the contours that maximally activated individual cells. We combined the algorithm with a delayed-match-to-sample task, to test how the optimal contours might be molded by the monkey's expectation for particular cue shapes. We found that V1 neurons were selective for complex shapes, a property previously ascribed to higher cortical areas. Furthermore, the shape selectivity was reprogrammed by perceptual task: Over the whole network, the optimal modes of geometric selectivity shifted between distinct subsets of the AF, alternately representing different stimulus features known to predominate in natural scenes. Our results suggest a general model of cortical function, whereby horizontal connections provide a broad domain of potential associations, and top-down inputs dynamically gate these linkages to task switch the function of a network.**

The cortical processing of sensory information is shaped by the spatial and cognitive context surrounding the sensory stimulus. In visual perception, the very appearance of local image regions is determined not by the regions themselves, but by their relationship to the surrounding visual scene (1). In visual cortical areas with small to intermediate receptive fields (RFs), these contextual influences are manifested by the interaction between stimuli falling within the classical RF (cRF) and stimuli in the extraclassical RF surround (2–5). Likewise, the behavioral context in which a scene is viewed shapes both the observer's perception and the underlying neural responses (6, 7). It is well established that space-, feature-, and object-based attention modulate the gain of neural responses (8–15), and recent evidence suggests even more profound cognitive influences on sensory processing (2, 9, 16–18).

In the primary visual cortex (V1), contextual interactions between the cRF and the surround imbue neurons with integrative properties that likely underlie a range of important visual functions, including contour integration (2, 14, 19), image segmentation (4), perceptual fill-in (20), and perceptual learning (16). It has been proposed that much of this computational sophistication in V1 derives from the plexus of horizontal axon collaterals that run parallel to the cortical surface and that these collaterals integrate visual information over an association field (AF) of spatial interactions (21). The AF (22) is a theoretical construct that forms the basis for current neural network models of V1 (20, 23); according to the theory, V1 neurons with cocircular RFs (RFs that respond optimally to segments of common linear or circular arcs) share the strongest horizontal connections in the network (Fig. S1). This connectivity pattern would implement a basic building block of image segmentation and object recognition, by computationally linking the segments of contours and object boundaries, and it would form the basis of fundamental perceptual qualities, like the Gestalt rules of perceptual grouping. Still, the actual ge-

ometry of the lateral contextual interactions, the very feature thought to endow V1 with its integrative functions, has never been fully characterized. Moreover, previous results have shown that contextual interactions in V1 are subject to dynamic cognitive control (2, 9, 16), raising the possibility that an AF in V1 might undergo concomitant behavioral modulations.

Here, we map the AF in V1 and survey its plasticity under different behavioral states. We trained three monkeys (monkeys A, B, and C) to detect the presence of a cued contour flashed within a field of random line segments (Fig. 1). By recording the activity of single neurons during the task, we could algorithmically construct contours that maximally activated the recorded neuron under different task conditions (see *SI Discussion* for caveats). We observed dynamic patterns of visuospatial integration that match the theorized AF, conferring V1 neurons with adaptive shape selectivity and a potential role in object recognition (24, 25).

## Results

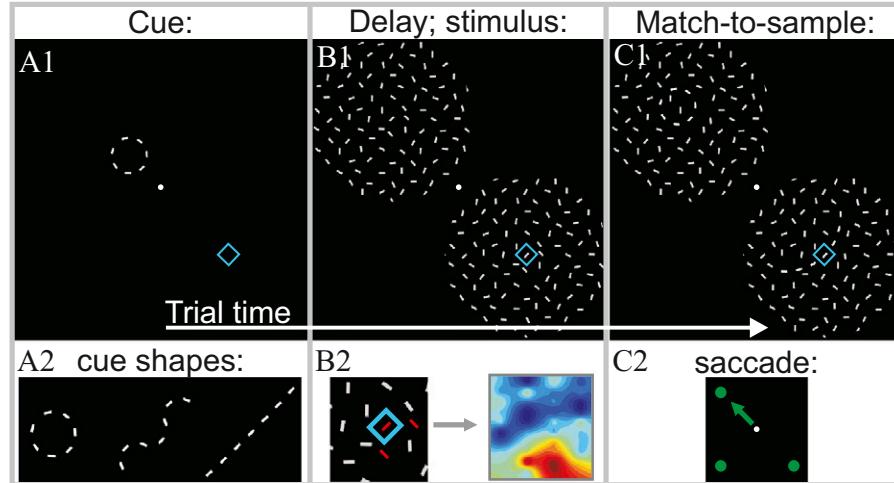
**Contour Detection Task and Stimulus Generation.** We recorded single neurons in the superficial layers of V1, from trained monkeys performing a contour detection task. Each experiment comprised hundreds of trials of the detection task, whereby a monkey determined the location of a cued contour in a delayed-match-to-sample paradigm (Fig. 1 and *SI Materials and Methods, Task Design*). During the delay period over the course of these trials, we recorded a neuron's responses to geometric stimuli, while the monkey was expecting the cued shape. Before every experiment, we selected a seven-bar contour (a closed circle, a sinusoid, or a straight line; Fig. 1A, 2) to serve as the cue in all of the ensuing trials. At the outset of each trial, the cue was presented next to the fixation point, with the central contour segment parallel to the neuron's preferred orientation (Fig. 1A, 1). A delay period followed the cue (Fig. 1B, 1), during which a stimulus contour, consisting of a number of geometrically aligned contour segments, was presented in the center of two identical fields of randomly oriented bars. The two fields were displayed in opposite quadrants of the computer monitor, with one positioned directly over the RF of the recorded neuron. In each trial, the embedded contour stimulus (Fig. 1B, 2) was chosen from a set of shapes created by a contour optimization algorithm (*SI Materials and Methods, Automated Stimulus Generation*). The neural responses to these embedded contours, each repeated over multiple trials, were used by the algorithm to progressively construct the neuron's preferred stimulus over the course of the experiment. At the end of each trial, a seven-bar contour was briefly flashed in each field, just before the fields were extinguished (Fig. 1C). The monkey's

Author contributions: J.N.J.M., W.L., and C.D.G. designed research; J.N.J.M. and C.D.G. performed research; J.N.J.M. analyzed data; and J.N.J.M., W.L., and C.D.G. wrote the paper.

The authors declare no conflict of interest.

<sup>1</sup>To whom correspondence should be addressed. E-mail: gilbert@rockefeller.edu.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1105855108/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1105855108/-DCSupplemental).



**Fig. 1.** Trial and task design. (Upper row) The sequence of frames that composed each trial. (Lower row) Schematic depictions of the trial events. The cyan rectangle represents the cRF of the recorded neuron. (A, 1) The cuing phase of the trial, initiated when the monkey holds his gaze on the fixation point, white dot. (B, 1 and B, 2) The stimulus presentation, with a geometric stimulus (red bars in B, 2) embedded in a field of random line segments. The neural activity recorded during this trial period guided the generation of increasingly effective stimuli (B, 2; see *SI Materials and Methods, Task Design* for details). (C, 1 and C, 2) After a random duration of geometric stimulus exposure, the bar elements in each field were abruptly rearranged to form a salient contour. The monkey performed a delayed match-to-sample task by making a saccade toward the direction of the field containing the cued contour, which could be in either field, or toward an alternative direction if neither field contained the cue. (A, 2) The three contour shapes that served as potential cues for our experiments. The cue contours and the geometric stimuli were always oriented so that the central bar matched the recorded neuron's preferred orientation (PO). In this example, the recorded neuron has a PO of 45°. The central bar of the test contours was always centered in the cRF.

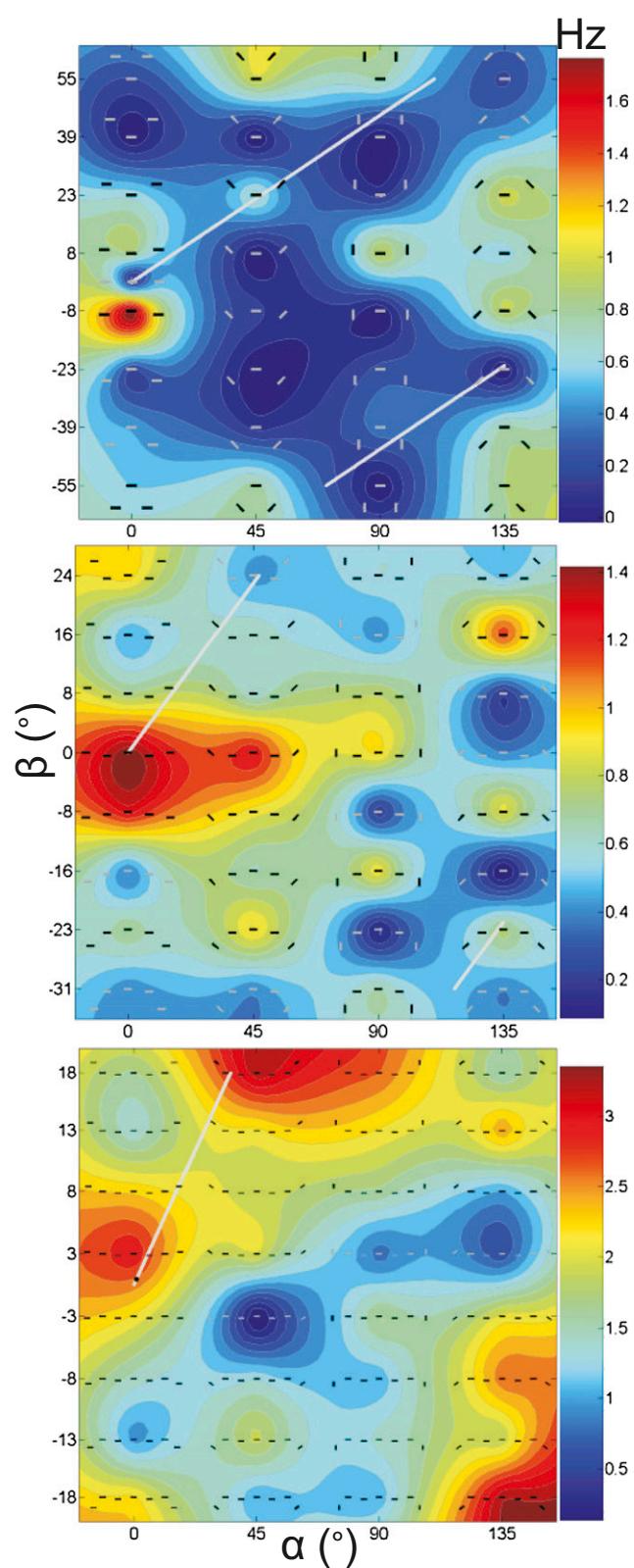
task was to signal whether either contour was the cued target, by making a saccade toward the location where it was flashed or toward a third location if both contours were distractors (open circles). The task was designed to engage top-down mechanisms during the delay period when neural responses were measured. The optimization algorithm converged on a preferred contour by sequentially testing the neural responses to stimulus sets containing successively longer contours; the neural responses to each stimulus set of a given contour length guided the generation of longer contours in the next set of test shapes.

**Geometric Tuning Surfaces.** To visualize neural responses to the sequential stimulus sets generated by the optimization routine, we constructed 3D “tuning surfaces” (the higher-dimensional analog of tuning curves; see Fig. S2 and S3 and *SI Materials and Methods, Three-Dimensional Tuning Surfaces* for details of their construction). For each neuron, we generated a series of surfaces, rendered as heat maps, where each map describes the neural responses to all of the stimuli of the same length created during a phase of the optimization algorithm (e.g., Figs. 2–4). The coordinates in a heat map define a continuous domain of stimulus geometries; the discrete points that we tested within this domain are indicated either with depictions of the corresponding stimuli or, where space is limited, with dots. Theoretical responses to stimuli between these tested coordinates were computed via scattered interpolation (Hardy’s multiquadratics) (26). The angular ( $\alpha, \beta$ ) coordinates describe the orientation ( $\alpha$ ) and position ( $\beta$ ) of the outermost pair of contour bars in each stimulus. In addition, the gray lines drawn over the surfaces are the stimuli whose outermost contour bars are cocircular with the central bar in the RF (for a description of cocircularity, see Fig. S1). In many of our experiments, we compared the five-bar stimuli that arose during the optimization program with a predetermined stimulus set, the five-bar versions of the target/distracter contours. Whenever they were included, the neural responses to these predefined stimuli are indicated by the correspondingly colored shapes below the five-bar tuning surface (e.g., Fig. 3 and Figs. S3 and S4).

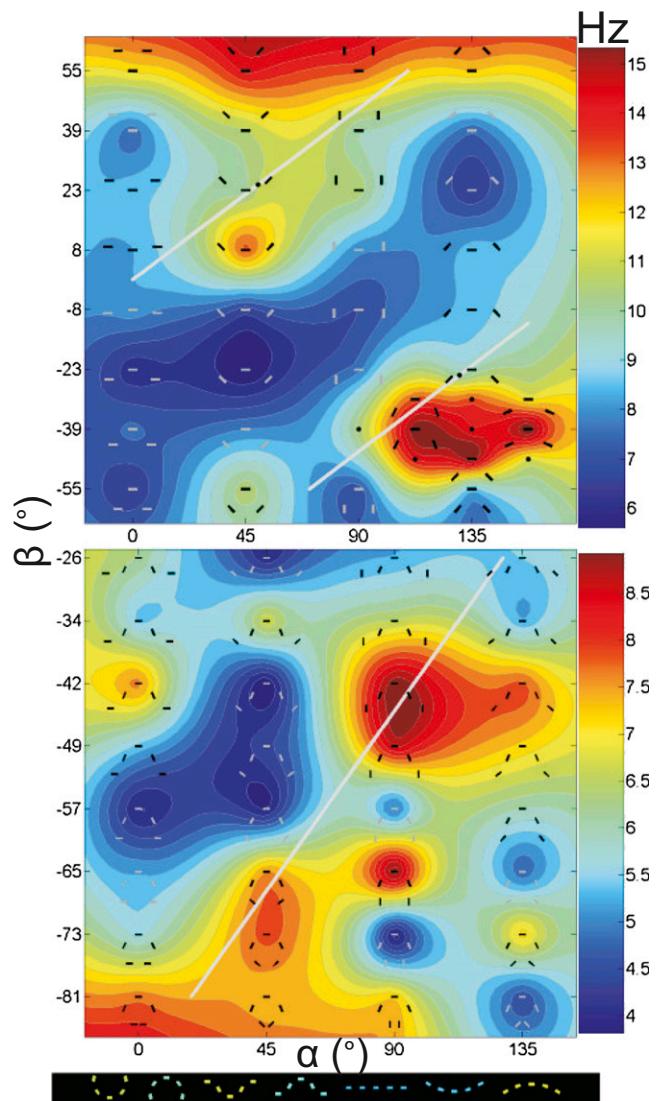
All V1 neurons showed selectivity for complex shapes, which was seen as a relief from inhibition (2) imposed by the random background fields. The tuning surfaces in Figs. 2–4, with additional examples in Figs. S3 and S4, describe the three major patterns of selectivity we encountered. We measured this neuronal shape selectivity while monkeys were cued to detect the line (Fig. 2), circle (Fig. 3), or wave (Fig. 4) shapes. Eye tracking analyses (*SI Results, Eye Tracking Analyses*) showed that the shape selectivity was not an artifact of eye movements (Fig. S9) and that eye movements did not cause systematic changes in neural firing rates (Fig. S10).

Fig. 2 shows the responses from a neuron, recorded from monkey A during the line detection task, to the three-, five-, and seven-bar stimulus sets that were tailored to the cell’s activity. This cell was profoundly inhibited by the random field, but it retained sharp geometric selectivity for stimuli near the coordinate (0°, 0°), corresponding to contours with iso-oriented, coaligned bars. After the optimization routine settled on the most facilitatory three-bar stimulus at coordinate (0°, -8°), two additional bars were appended to its ends in various configurations, creating a set of five-bar stimuli used to search for the optimal extension of the contour. The peak at collinearity on the five-bar surface was the optimum configuration of bars at the ends of these contours, which extended well beyond the cRF. The local maximum at (0°, 3°) on the seven-bar surface demonstrates selectivity for contour components several RF diameters away from the RF center. For stimuli beyond five bars in length, however, neural responses to differences at the contour ends were often less selective, so we did not attempt to construct contours longer than seven bars.

A fundamentally different mode of geometric selectivity was obtained from a neuron recorded while cuing the monkey to detect the circular target (Fig. 3). In its three-bar tuning surface, the cell expressed two broad regions of robust activation, corresponding to different sets of circular contours. This structure was mirrored in the neuron’s five-bar surface, which exhibited two peaks along the cocircularity line. Both peaks comprised stimuli whose outermost contour bars were cocircular to the bar in the RF. Close to one

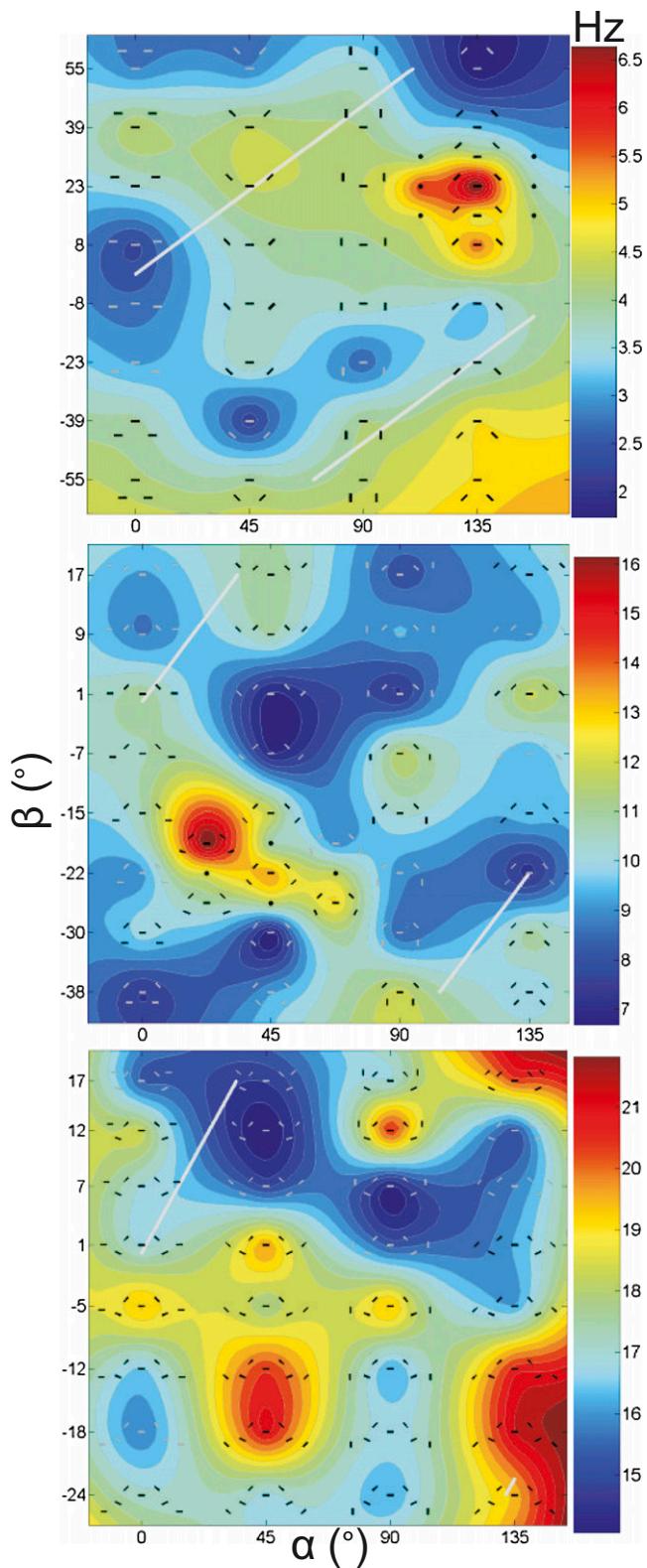


**Fig. 2.** Shape selectivity for a single neuron recorded during the line detection task. (*Top, Middle, and Bottom*) Responses of the cell (from monkey A) to the three-, five-, and seven-bar stimuli that were generated by the optimization routine. The neuron was recorded during the delay period (Fig. 1) after the monkey had been cued to detect the linear target; the cell displayed a corresponding preference for linear geometries that was apparent for three- and five-bar stimuli.



**Fig. 3.** Circular optima under the circle detection task. The three-bar (*Top*) and five-bar (*Bottom*) tuning surfaces of one neuron were recorded while monkey B was engaged in the circle detection task. The cell expressed clear selectivity for distinct subsets of circular geometries. The contours at the *Bottom* are the five-bar versions of the target and distracter contours; their colors indicate the corresponding neural responses to these stimuli.

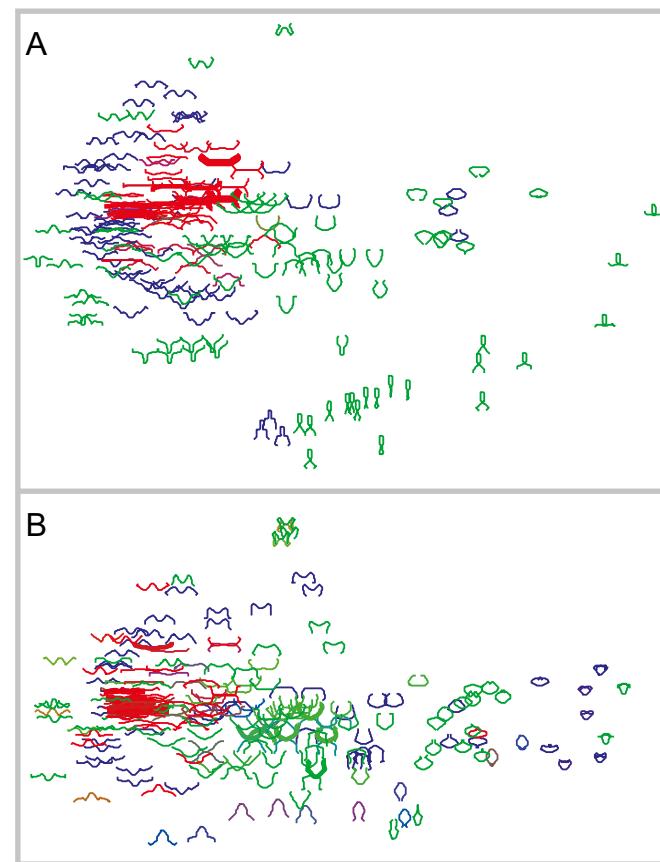
peak, all of the contour bars in each stimulus [e.g.,  $(45^\circ, -65^\circ)$ ;  $(45^\circ, -73^\circ)$ ] were approximately tangent to a single circle; but close to the other peak, the outermost contour bars lay along a broader, open circular arc than the innermost three bars [e.g.,  $(90^\circ, -42^\circ)$ ;  $(90^\circ, -49^\circ)$ ]. The horizontal panel (Fig. 3, *Bottom*) demonstrates that the cell strongly preferred the contours generated by the optimization routine, compared with the distracter- and target-like stimuli. Three characteristic features of our data are evidenced here. First, neuronal response optima were seen as one or more islands of activation—regions of the stimulus space representing contours with a particular spatial scale, radius of curvature, and orientation. Second, neurons were often selective for a contour of a particular orientation and only weakly responsive to the same contour rotated by  $180^\circ$ . Third, most cells were not maximally activated by any of the target or distracter contours the animals were trained to detect; rather, they preferred similar but not necessarily identical stimuli (Figs. S3 and S4).



**Fig. 4.** Heat maps with wave-like optima under the sinusoid detection task. The stimulus evolution and neural responses for a cell were recorded while monkey A performed the sinusoid detection task. The geometric optima expressed by this cell are complex configurations of line segments, whose positions and orientations suggest contours with reversing directions of curvature.

The final class of geometric selectivity we observed is characterized by the responses of a cell recorded during the sinusoid detection task (Fig. 4). Selective for perceptually “wave-like” shapes, the cell has maxima on its tuning surfaces that are far from either cocircularity line. Despite the “rippled” or disjointed shape of this cell’s optimum contours, its selectivity is as sharp as the line-selective neuron in Fig. 2, recorded under the line detection task. These data highlight the more basic observation, seen throughout our data, that the smoothest contours did not always elicit the strongest responses.

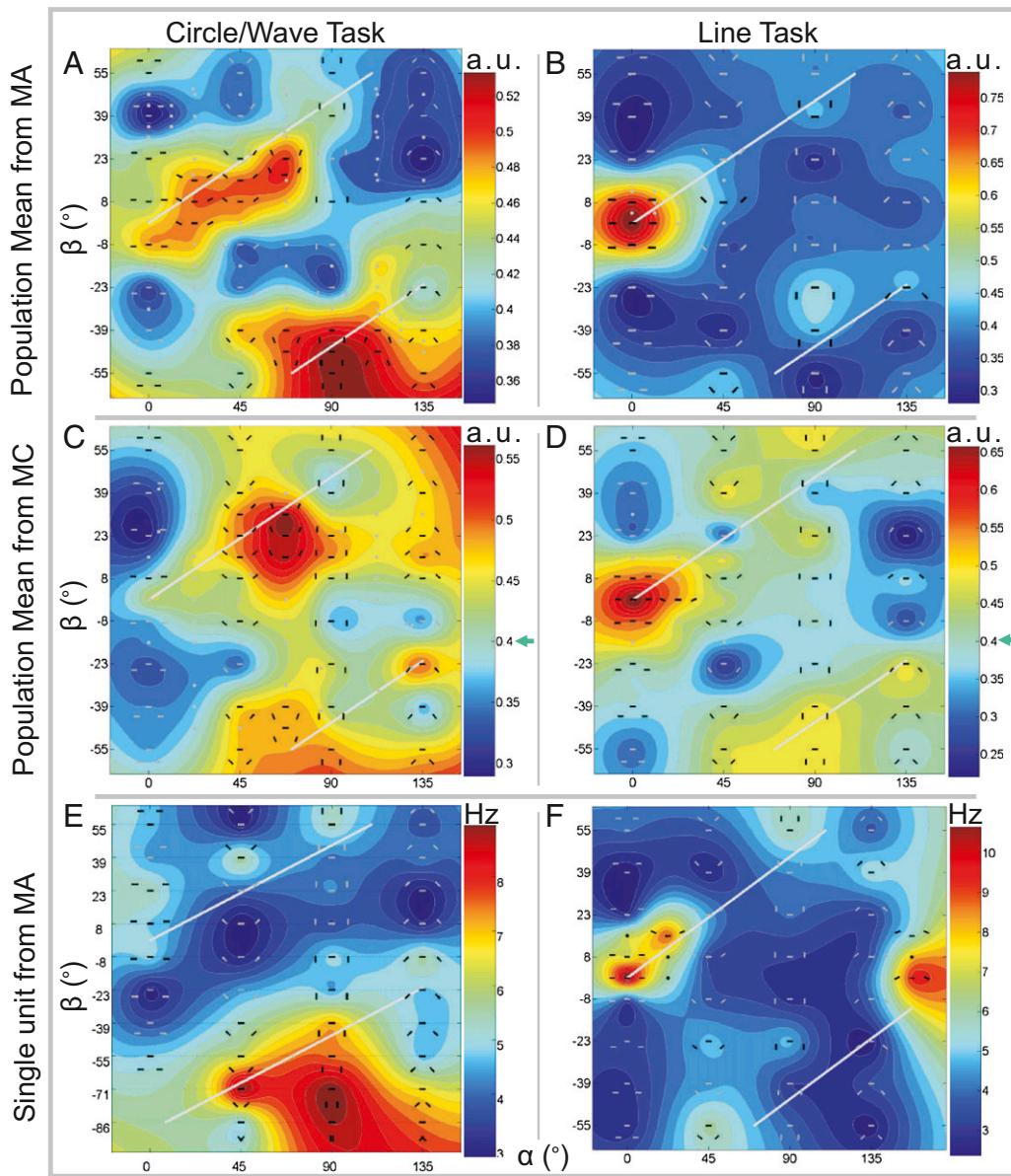
**Shape Selectivity and Task Dependence Over the Population.** The most facilitatory five-bar stimuli for all neurons in monkeys A and C, showing the full repertoire of preferred stimuli over the population and their relationship to the cue, are plotted in Fig. 5. Here, the stimuli are represented as continuous curves by directly connecting the ends of adjacent contour bars in the original stimuli. We used a multidimensional scaling (MDS) approach known as Sammon’s mapping to project the five-bar stimuli, which exist in a high-dimensional space, onto the 2D



**Fig. 5.** The repertoire of task-dependent shape selectivity over the neuronal population. (A and B) The preferred five-bar stimuli generated by all of the neurons recorded from monkeys A (A) and C (B), pooled together. To construct these plots, the firing rate of each cell was linearly mapped into the interval [0, 1] (0, weakest response; 1, response to optimum contour). All stimuli from each cell that elicited a response  $\geq 0.7$  are displayed; stimuli that were derived from more than one neuron are shown in boldface type. The position of each stimulus is determined by Sammon’s mapping (*SI Materials and Methods*) and its color indicates the task condition under which it was generated (red, line detection task; green, circle task; blue, sinusoid task). The “paperclip” shapes on the far right of the stimulus space were created by the optimization routine when the optimal three-bar shape was already a closed contour.

plane (*SI Materials and Methods, Five-Bar Population Analysis with Sammon's Mapping*). Similar stimuli are mapped to nearby points on the plane, whereas dissimilar stimuli are mapped to distant locations. Fig. 5 demonstrates that V1 neurons can be selective for a diverse range of contours, spanning the whole continuum from lines to circular arcs to undulating contours. The range and complexity of stimulus selectivity seen here reveal a previously undiscovered level of sophistication in V1 shape processing. We additionally color coded the stimuli to indicate the relationship between this stimulus selectivity and perceptual task: the red, green, and blue stimuli elicited strong neural responses under the line, circle, and wave detection tasks, respectively. The clustering of stimuli with the same color

indicates that neurons tended to guide the optimization algorithm toward similar contours under the same task condition. For both monkeys shown here, neurons preferred near-collinear, elongated contours during the line detection task (shown in red). Conversely, neurons steered the optimization toward curved contours during the circle and wave detection tasks (shown in green and blue, respectively). The distribution of preferred stimulus shapes obtained from monkey B (Fig. S5) closely resembles the distribution from monkey C (Fig. 5B), except that contours of the same color do not cluster. Neurons from monkey B also underwent profound changes in their geometric tuning as a function of perceptual task (see next paragraph), but those changes were apparent over the averaged



**Fig. 6.** Neuronal heat maps are reshaped by task. (A–D) The mean three-bar heat maps averaged over all of the neurons recorded from monkeys A and C during the circle and wave detection tasks (A and C) and during the line detection task (B and D). Because each map is combined from many cells, the set of all stimuli generated by the whole group of cells is depicted over the data. The 32 stimuli from the initial stimulus set are drawn, as are the stimuli overlying the highest regions of the mean response (other stimuli are indicated with gray points). Note the diversity of refinement stimuli created by cells under the circle/wave task, compared with the line task. Arrows on the color scale in C and D show response ( $R = 0.40$ ) to a one-bar geometric "contour" (one bar in the RF, embedded in the random field). Sample sizes: (A)  $n = 17$ ; (B)  $n = 36$ ; (C)  $n = 25$ ; (D)  $n = 37$ . (E and F) The three-bar heat maps for a single neuron from monkey A under the circle (E) and line (F) tasks.

population response, rather than in the optimum shapes preferred by individual neurons.

To determine how perceptual task might reshape neural tuning surfaces and how those changes might sum up over groups of neurons, we analyzed the population responses to three-bar stimuli. Because the optimization program was always seeded with the same set of three-bar stimuli (*SI Materials and Methods, Automated Stimulus Generation*), the neural responses to this region of the stimulus space can be directly compared across cells and task conditions. We mapped the height of each three-bar tuning surface into the interval [0, 1], to normalize for differences in neuronal firing rates. We then grouped the surfaces that were constructed while the monkeys performed each of the three detection tasks and averaged the data over each group, to obtain three population tuning surfaces. We performed permutation tests between these mean tuning surfaces to determine whether any differences between them were statistically significant (*SI Materials and Methods, Permutation Test*). For all three monkeys, the differences between the average three-bar tuning surfaces under the circle and wave tasks were not statistically significant ( $P \geq 0.2$ ). We therefore pooled the data from those two task conditions and compared the mean of the merged data with the mean surface under the line detection task. The averaged tuning surfaces from monkeys A and C under the line and circle/wave detection tasks are presented in Fig. 6 *A–D*. (The corresponding data for monkey B are shown in Fig. S6.) Over the neuronal populations from all three monkeys, the cumulative response depended as much on the animals' cognitive state as on the actual stimulus geometry. Under the line detection task, the network expressed a narrow selectivity for the linear geometry; under the circle and wave tasks, the network adopted an entirely different mode of selectivity, for circular shapes. The difference between the mean tuning surfaces under the line and circle/wave tasks was statistically significant (monkey A, total number of surfaces,  $n = 53$ ,  $P = 4 \times 10^{-5}$ ; monkey B,  $n = 63$ ,  $P = 0.007$ ; and monkey C,  $n = 62$ ,  $P = 0.003$ ).

Because the construction of five- or seven-bar optimum contours required thousands of trials and hours of recording time, it was not possible to run the optimization algorithm to completion under different task conditions for the same cell. We did, however, generate three-bar tuning surfaces for a subset of cells that were each recorded under two different task conditions. Consistent with the results obtained from averaging the data over the population, we found that the same neurons were able to dynamically change their tuning surfaces according to the cued shape. We recorded from 13 neurons (7 from monkey A and 6 from monkey B) while the animal carried out two separate blocks of several hundred trials each. We switched the cue between the two blocks, and the neural responses under the two task conditions gave rise to two distinct tuning surfaces. Fig. 6 *E* and *F* illustrates the results from one such experiment, in which monkey A was cued to detect first the circle (Fig. 6*E*) and then the line (Fig. 6*F*) while we recorded from the same neuron. Coincident with the change in the cued target, the AF underwent a dramatic shift, from a cocircular to a collinear pattern of facilitation. The direction of this shift in geometric tuning was consistent across the subpopulation of 13 neurons that were recorded under both the line and the circle/wave detection task ( $P = 0.013$ ; permutation test). Fig. S7 plots the average heat map across these 13 cells, obtained under both sets of task conditions. Under the line detection task, the cumulative neural response was dominated by a sharp peak of collinear facilitation. Under the circle and wave tasks, a local maximum near collinearity remained in the population response, but the magnitude of this peak was lower, and the cells' geometric facilitation became distributed along the cocircularity lines. The shift in the three-bar tuning surfaces for the same neurons therefore paralleled the

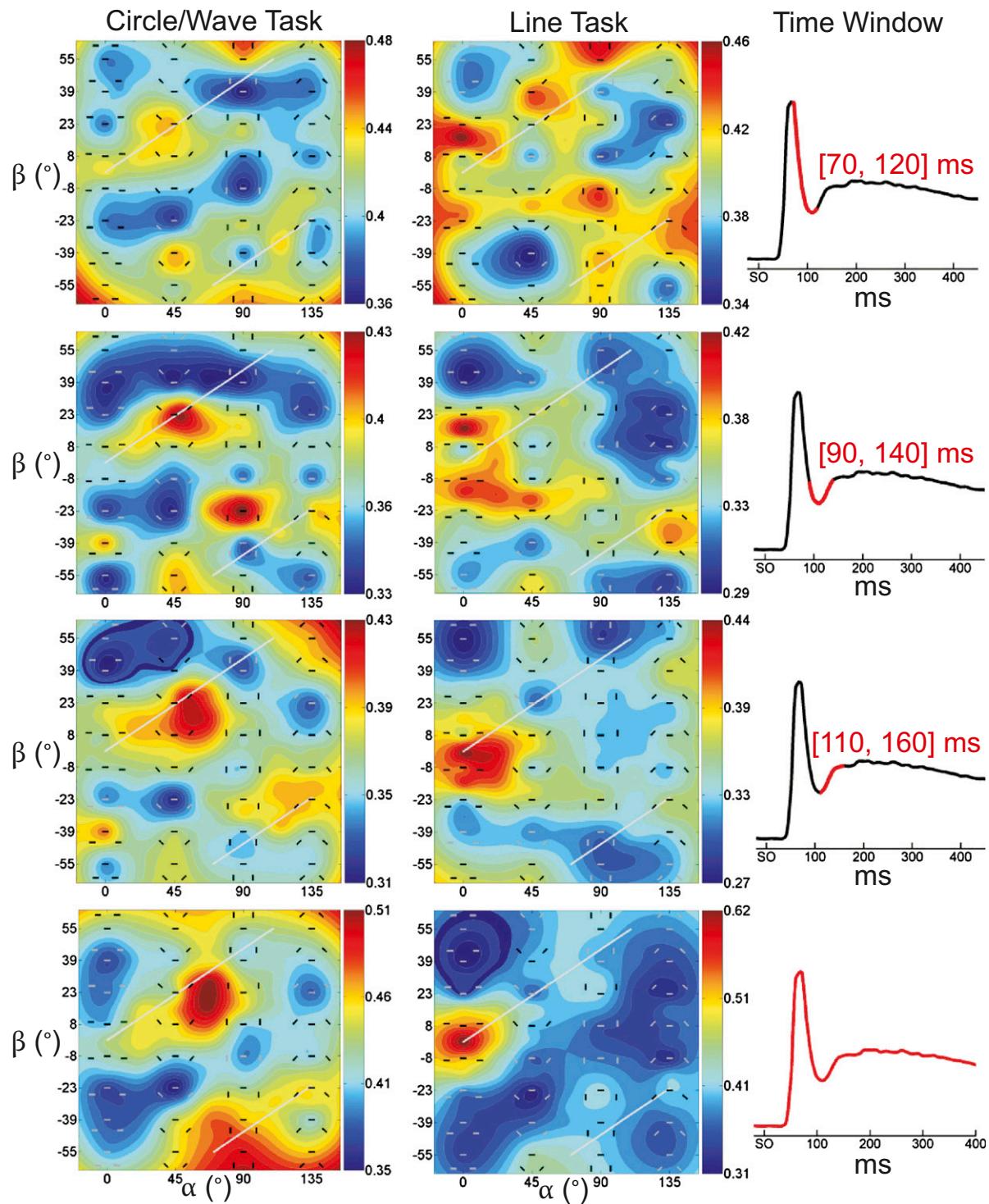
shift seen between neurons, when different task conditions were used during recordings from different cells.

**Time Course of Geometric Selectivity.** Finally, we analyzed the temporal dynamics of shape selectivity in V1. An ANOVA analysis (*SI Results, ANOVA Analysis*) revealed that neurons developed an initial mode of geometric selectivity—i.e., neurons began to respond differentially to at least some stimuli—at ~72 ms after stimulus onset (SO). To investigate the timing in more detail, we plotted the evolution of three-bar shape selectivity over the population, as a function of task condition. We pooled the three-bar tuning functions from all three monkeys and separated them according to task condition. Fig. 7 shows the mean heat maps recorded during either the circle or the wave task (Fig. 7, *Left*) and during the line detection task (Fig. 7, *Center*), plotted using various time windows (Fig. 7, *Right*). When the entire duration of each trial is used to construct the mean heat maps (Fig. 7, *Bottom*), the task-dependent differences between the surfaces are highly significant ( $P < 4 \times 10^{-5}$ ; permutation test). To measure the temporal onset of these differences, we moved a sliding 50-ms window across the neural response and plotted the population activity within that window under each task condition. We found distinct patterns of selectivity between the task conditions that began to develop in the window between 70 and 120 ms following SO. The shape of the tuning functions evolved over the next tens of milliseconds, reaching maturity in the time window between 110 and 160 ms. Within this window, differential peaks of collinear and cocircular facilitation were already apparent on the corresponding response surfaces ( $P = 0.036$ ; permutation test). The results suggest that geometric selectivity matures by 110 ms, when neural responses recover from an inhibitory dip in firing probability induced by the random fields (2) (Fig. 7, *Right*). We obtained precisely the same timing results from an analysis of five-bar contours (*SI Results, Five-Bar Timing Analysis* and Fig. S8), indicating that this time course is not unique to three-bar stimuli.

## Discussion

V1 has long been recognized as a geometric processor, responsible for parsing the visual scene into its component lines and edges (27). More recent evidence suggests that V1 may also merge these components into perceptually unified wholes (2) via the AF. In the standard model, the pattern of neuronal interconnections in V1 echoes the cocircular statistics of natural images (20, 28): The strongest connections in the network link neurons whose RFs fall along the same linear/circular arcs. Although the AF was originally coined as a psychophysical concept (22), our study directly shows that a network with matching properties exists in V1. Over the population, the strongest contextual interactions we observed fell on, or very near, the cocircularity line (Figs. 6 and 7). However, as an important nuance, the network in V1 never expressed the full range of lateral interactions from the AF at once (which includes all collinear and cocircular interactions). Rather, the neural population expressed subsets of these interactions—peaks of facilitation that shifted their position along the cocircularity line when the perceptual task was changed. Moreover, different cells preferred a diverse range of curved contours, including circles and sinusoids of various shapes, whereas previous results have focused on the responses to straight line geometries. The cortical strategy may be to ensure that the population activity follows a narrow pattern that is appropriate for detecting and encoding smooth contours, while maintaining a richer repertoire of shape selectivity at the level of individual neurons. This diversity of responses may be used to accommodate mechanisms of object recognition and scene segmentation (24, 25).

The prevailing view of top-down interactions emphasizes their role in gain control or attentional competition (6, 7), but



**Fig. 7.** Temporal evolution of geometric selectivity. (Left) The three-bar tuning surfaces, in different time windows, averaged from 107 different experiments from all three monkeys. The data in this column were collected during either the circle or the wave task. (Center) The three-bar heat maps, within different time windows, pooled over 72 line detection experiments from each monkey. (Right) The peri-stimulus time histogram (PSTH) obtained from pooling the spikes elicited by all three-bar stimuli from all recorded neurons in each monkey. The region of the PSTH used to construct each pair of response surfaces is highlighted in red. SO, stimulus onset. The tuning surfaces were normalized and averaged as in Fig. 6 A–D. For clarity, only the stimuli in the initial stimulus set are drawn over each surface.

they may also provide an input selection mechanism that enables cortical areas to act as adaptive processors (21). Our data demonstrate that cognitive influences can reprogram an entire network of sensory neurons (*SI Discussion, Role of Expectation*

*in Our Experiments*). This idea is a major departure from the current paradigm, which holds that top-down influences merely gate the magnitude, rather than the function, of neural responses. Mechanistically, we speculate that the adaptive

processing in V1 involves the top-down gating of horizontal connections (*SI Discussion, Mechanism of Task-Dependent Shape Selectivity in V1*), rather than traditional gain control or a simple reflection of higher sensory processing. Anatomical, physiological, and theoretical work implicate the horizontal connections as the substrate for the geometric contextual interactions (2, 19–21, 23, 29–31). The task dependency of those interactions, shown here, suggests that feedback projections may inhibit some sets of lateral interactions and/or activate others, thereby establishing different network states with different geometric optima. Furthermore, the time course of geometric facilitation is consistent with this mechanism: traditional attentional effects occur later than the modulations we observed, even in higher cortical areas (15).

Similar cortical mechanisms that impart V1 with its adaptive tuning may imbue other areas with the same flexibility, like the prefrontal cortex with its ability to load and switch between different behavioral programs. Our finding that V1 acts like an adaptive integrator suggests parallels with association cortices, which must also integrate information dynamically as a function of cognitive state.

## Materials and Methods

We performed the animal preparation procedures and many aspects of the electrophysiological recordings as previously described (2, 16). We developed

an “optimization” algorithm to adaptively measure neural responses to stimuli composed of discrete contour elements (“contour bars”; Fig. 1). Recordings were taken from macaques while they were in the delay phase of a match-to-sample task, in the midst of their expectation for a cued contour. The optimization algorithm built up the preferred shape for a recorded neuron—under a particular behavioral state—by progressively bringing contour elements into an “optimal” configuration with a fixed bar in the cRF. At the start of an experiment, during the delay phase of each trial, the algorithm tested the neural responses to a range of three-bar stimuli, pruning and redefining the stimulus set over hundreds of trials until it converged upon the neuron’s preferred three-bar stimulus. (A three-bar stimulus is a shape composed of three discrete line elements.) The algorithm then searched for optimal extensions of this three-bar stimulus, creating and pruning five-bar stimuli, and later seven-bar shapes, in the search for the neuron’s preferred contour. The optimization program changed the stimulus sets at discrete intervals during an experiment, after hundreds of test trials, but the stimulus presented within the delay portion of any trial was static. We always measured neuronal shape selectivity while monkeys were expecting a single contour from a set of three possible cue shapes (a straight line, a closed circle, or a sinusoidal wave). We assessed the cognitive influence of expectation by comparing the preferred contour shapes generated by populations of neurons while monkeys were expecting each of the cue shapes or by comparing the responses of single neurons before and after changing the expected cue contour. Details of these methods are available in *Results* and in *SI Materials and Methods*. All procedures were in accordance with the National Institutes of Health *Guide for the Care and Use of Laboratory Animals* and with the approval of the Institutional Animal Care and Use Committee at The Rockefeller University.

**ACKNOWLEDGMENTS.** We thank Valentin Piéch and Shimon Ullman for helpful discussions. This work was supported by National Institutes of Health Grant EY007968.

\*Piéch V, et al. (2009) A network model of top-down influences on local gain and contextual interactions in visual cortex. *Soc Neurosci Abstr* 701.10.

1. Albright TD, Stoner GR (2002) Contextual influences on visual processing. *Annu Rev Neurosci* 25:339–379.
2. Li W, Piéch V, Gilbert CD (2006) Contour saliency in primary visual cortex. *Neuron* 50: 951–962.
3. von der Heydt R, Peterhans E, Baumgartner G (1984) Illusory contours and cortical neuron responses. *Science* 224:1260–1262.
4. Lee TS, Mumford D, Romero R, Lamme VAF (1998) The role of the primary visual cortex in higher level vision. *Vision Res* 38:2429–2454.
5. Kapadia MK, Westheimer G, Gilbert CD (2000) Spatial distribution of contextual interactions in primary visual cortex and in visual perception. *J Neurophysiol* 84: 2048–2062.
6. Maunsell JHR, Treue S (2006) Feature-based attention in visual cortex. *Trends Neurosci* 29:317–322.
7. Reynolds JH, Chelazzi L (2004) Attentional modulation of visual processing. *Annu Rev Neurosci* 27:611–647.
8. Reynolds JH, Pasternak T, Desimone R (2000) Attention increases sensitivity of V4 neurons. *Neuron* 26:703–714.
9. Ito M, Gilbert CD (1999) Attention modulates contextual influences in the primary visual cortex of alert monkeys. *Neuron* 22:593–604.
10. McAdams CJ, Maunsell JHR (1999) Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *J Neurosci* 19:431–441.
11. Treue S, Martínez Trujillo JC (1999) Feature-based attention influences motion processing gain in macaque visual cortex. *Nature* 399:575–579.
12. Motter BC (1993) Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. *J Neurophysiol* 70:909–919.
13. Motter BC (1994) Neural correlates of attentive selection for color or luminance in extrastriate area V4. *J Neurosci* 14:2178–2189.
14. Roelfsema PR, Lamme VAF, Spekreijse H (1998) Object-based attention in the primary visual cortex of the macaque monkey. *Nature* 395:376–381.
15. Chelazzi L, Duncan J, Miller EK, Desimone R (1998) Responses of neurons in inferior temporal cortex during memory-guided visual search. *J Neurophysiol* 80:2918–2940.
16. Li W, Piéch V, Gilbert CD (2004) Perceptual learning and top-down influences in primary visual cortex. *Nat Neurosci* 7:651–657.
17. David SV, Hayden BY, Mazer JA, Gallant JL (2008) Attention to stimulus features shifts spectral tuning of V4 neurons during natural vision. *Neuron* 59:509–521.
18. Ghose GM, Bearl DW (2010) Attention directed by expectations enhances receptive fields in cortical area MT. *Vision Res* 50:441–451.
19. Chisum HJ, Mooser F, Fitzpatrick D (2003) Emergent properties of layer 2/3 neurons reflect the collinear arrangement of horizontal connections in tree shrew visual cortex. *J Neurosci* 23:2947–2960.
20. McManus JNJ, Ullman S, Gilbert CD (2008) A computational model of perceptual fill-in following retinal degeneration. *J Neurophysiol* 99:2086–2100.
21. Gilbert CD, Sigman M (2007) Brain states: Top-down influences in sensory processing. *Neuron* 54:677–696.
22. Field DJ, Hayes A, Hess RF (1993) Contour integration by the human visual system: Evidence for a local “association field”. *Vision Res* 33:173–193.
23. Li ZP (1998) A neural model of contour integration in the primary visual cortex. *Neural Comput* 10:903–940.
24. Epshtain B, Lifshitz I, Ullman S (2008) Image interpretation by a single bottom-up top-down cycle. *Proc Natl Acad Sci USA* 105:14298–14303.
25. Ullman S (2007) Object recognition and segmentation by a fragment-based hierarchy. *Trends Cogn Sci* 11:58–64.
26. Amidror I (2002) Scattered data interpolation methods for electronic imaging systems: A survey. *J Electron Imaging* 11:157–176.
27. Marr D (1982) *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information* (Freeman, San Francisco).
28. Sigman M, Cecchi GA, Gilbert CD, Magnasco MO (2001) On a common circle: Natural scenes and Gestalt rules. *Proc Natl Acad Sci USA* 98:1935–1940.
29. Bosking WH, Zhang Y, Schofield B, Fitzpatrick D (1997) Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *J Neurosci* 17: 2112–2127.
30. Stettler DD, Das A, Bennett J, Gilbert CD (2002) Lateral connectivity and contextual interactions in macaque primary visual cortex. *Neuron* 36:739–750.
31. Ben-Shahar O, Zucker S (2004) Geometrical computations explain projection patterns of long-range horizontal connections in visual cortex. *Neural Comput* 16:445–476.

# Supporting Information

McManus et al. 10.1073/pnas.1105855108

## SI Results

**ANOVA Analysis.** To determine how quickly neural responses became geometrically selective after the stimulus onset, we performed a nonparametric two-way ANOVA (Friedman's test) on a dataset that was amenable to statistical analysis. The data consisted of 22 neurons from monkey C, for which each geometric stimulus was presented 40 times. For these cells, the stimulus set consisted of the 32 three-bar stimuli normally used to seed the optimization algorithm. We performed the statistical test on a sliding 50-ms window, which we swept across the time course of the neural response, to determine when the tuning surfaces took on a statistically nonplanar shape. When the left edge of the spike-counting window was moved past 50 ms (corresponding to spikes collected 50–100 ms after stimulus onset), the *P* value for the null hypothesis (that the tuning surfaces were flat) fell below 0.1. Past 61 ms, the *P* value fell below 0.01; past 72 ms, the *P* value reached an asymptote of  $<10^{-10}$ . (When applied to the same data, the parametric two-way ANOVA yielded virtually identical results.) This analysis demonstrates that an initial mode of geometric selectivity was in place between 61 and 111 ms following stimulus onset. More precisely, if the statistical power of the test reaches saturation when the time window includes only selective portions of the response, then selectivity should first arise at about 72 ms. Interestingly, this time corresponds to the peak of the neuronal onset response (see Fig. 7 and Fig. S8 for the shape of the PSTH).

**Five-Bar Timing Analysis.** We extended our temporal analysis of geometric selectivity to five-bar contours, in addition to the three-bar analysis provided in the main text. Because different neurons tended to steer the optimization program into different regions of the stimulus space for five-bar stimuli, we could not directly average the five-bar tuning surfaces across cells. Instead, we averaged neural responses to the five-bar versions of the target and distracter stimuli, which were used across a large subset of our data. In Fig. S8, we compare the time course of the neural response to these stimuli, averaged over data from all three monkeys, as a function of task condition. We plot the PSTH for the five-bar line in red and the mean PSTH for the circular and wave shapes in blue. Fig. S8A shows the neural responses recorded from 38 neurons in the line detection task; Fig. S8B plots the activity of 67 different neurons recorded under the circle or wave cues. The task-dependent geometric selectivity for the five-bar line emerged rapidly during the rebound from inhibition, with the same time course as the three-bar shape selectivity. The reciprocal selectivity for curved shapes under the circle/wave detection task is not apparent in Fig. S8B because the shapes preferred by neurons under those task conditions were usually different from the cue shapes themselves. Moreover, the blue trace represents the average response over all of the curved targets, which masks strong responses of individual neurons to one cue shape or another.

**Eye Tracking Analyses.** We performed two analyses to confirm that the neural shape selectivity was not an artifact of systematic biases in the monkeys' eye movements during the stimulus presentation, within the delay period of a trial (Fig. 1B). First, we plotted the eye position distributions for each monkey, using separate sets of trials corresponding to the strongest and weakest five-bar stimuli (Fig. S9). The distributions are 2D spatial plots, showing the probability that the monkeys' gaze fell within each region of the fixation window. In our experiments, the fixation windows ranged from 0.4° to 0.5° in radius (trials with eye movements outside this window were aborted), but the majority of the gaze directions

fell within 0.2° of the fixation point. The shape and spatial extent of the eye position distributions did not differ between presentations of optimal and suboptimal contours, so the differences in neural responses to those stimuli cannot be attributed to systematic deviations in eye position. Moreover, the total amount of eye movement in each set of trials was virtually identical, indicating that monkeys did not make more microsaccades or refixations during one set of stimulus conditions than in the other. The precision of our eye tracking system is sufficient to detect saccades as small as 0.05° in either the vertical or the horizontal direction (1).

The concern about stimulus-related differences in fixation is based on the assumption that those differences engender systematic differences in the neural response. In our second analysis, we therefore computed the trial-by-trial correlations between the neural response and several distinct features of eye movements for each monkey (Fig. S10). We found no correlations beyond chance level between neuronal firing rates and any feature of the monkeys' eye movements, including the total length of the eye position trajectory, the maximum displacement of the gaze direction from the fixation point, and the mean displacement of the eye position from the fixation point. The results demonstrate that, even if there were small differences in eye position or movement between stimuli, they would not have biased our measurements of neural selectivity.

## SI Discussion

**Caveats for the Stimulus Optimization Algorithm.** It is important to introduce some caveats pertaining to the construction of "optimal" contours in our experiments. Throughout this work, we refer to the best stimuli created by the stimulus generation program as optimal contours, but they are not necessarily optimal over the entire "universe" of contour shapes. They were the most effective stimuli created under the constraints of the optimization program, but a rigorous search for a truly optimum stimulus would require an impossibly exhaustive search of all possible stimuli. The assumption of mirror symmetry, the stepwise procedure for building successively longer contours, and the weak smoothness constraint that are implicit in the method could all prevent the detection of truly optimal contours for some neurons. (For instance, the algorithm ignores the possibility that the optimum configuration of the inner contour bars might change when new contour bars are added.) Moreover, it is not clear that neurons with complicated response properties, like those in V1, have a single optimal stimulus. Indeed, many of the cells in our population have three- and five-bar tuning functions with more than one prominent peak of facilitation. Any plausible method of searching a sufficiently large stimulus space for an optimum will necessarily be incomplete, and the notion that neurons have a single optimal stimulus may be overly simplistic. The basic value of the "optimization" program is that it guides the stimulus generation into regions of the stimulus space where the responses of a given neuron are informative. It can reliably detect subsets of complex stimuli that strongly activate neurons, even if it cannot detect all such stimuli. The traditional approach of using static stimulus sets, particularly for complex stimuli, suffers from the drawback that the chosen stimuli may not yield any insight into the selectivity of an arbitrary neuron. Neurons were selective for shapes with a limited range of curvature and orientation, so using a predefined set of stimuli would sample most neurons' responses in a flat region of their tuning surface. As an extreme example, consider the cell in Fig. S4B. It has clear tuning for wave-like shapes in the five-bar tuning surface, which were discovered by the optimization algorithm, but it did not re-

spond to any of the contours used as the static five-bar stimulus set (*Bottom Inset*). Our optimization program can explore a much broader range of stimuli than previous approaches applied to V1, and its generality reveals new sets of contextual interactions that could be missed by other approaches.

**Role of Expectation in Our Experiments.** Expectation, and the various attentional mechanisms that are often bound with it, may play a significant role in our experiments. We investigated the modulation of neural responses, not to the target shapes the monkeys were cued to detect, but rather to arbitrary contours that were presented over the RF while the monkey was primed to detect the upcoming targets. This experimental design gives our study relevance to theories of predictive coding and “countercurrent stream” models of cortical function (2–5). Interestingly, the few studies that have shown task-dependent sharpening of neural tuning (in V4 and MT, where the influence of cognitive effects is less controversial) have done so in the context of expectation (6, 7), whereas theories that emphasize gain control mechanisms are based on spatial- and feature-based attentional studies (8, 9). Although attention may be important for focusing the brain’s limited computational resources on behaviorally relevant stimuli, expectation may be useful for resolving ambiguities arising from the projection of a fourdimensional world onto the 2D retinal surface (2).

**Mechanism of Task-Dependent Shape Selectivity in V1.** One alternative explanation for our data is that traditional spatial- or even object-based attention could explain the results, by selectively boosting the neural responses to salient stimuli in the RF that resemble the cued target. Under this view, V1 neurons would respond strongly to stimuli that are reminiscent of the cue only because the monkey notices them and then allocates attention either to the spatial location of the RF or to the stimulus shape as a whole. Non-cue-like stimuli would not elicit the animal’s attention, so the privileged allocation of attention by cue-like stimuli would cause gain-induced peaks on the tuning surface. This interpretation fundamentally differs from our own, which holds that shape selectivity is generated via the selective expression of lateral interactions between V1 neurons, not by preferential changes in attentional allocation that cause neuronal gain changes.

There are several compelling reasons to favor our interpretation over the alternative. First, V1 neurons express task-dependent selectivity in the very first stage of the optimization program, for three-bar stimuli. It is known from previous results that contours of this length, in the same range of retinal eccentricities, are not perceptually salient (1). Even in a task that requires the monkey to directly detect the presence of a three-bar linear contour embedded in a random field, the detection performance is barely above chance level (1). However, in our task, the geometric contours were not behaviorally relevant and were not predictive of the upcoming target shapes. Second, the most excitatory stimuli were not always the most salient. Third, even if the monkey notices particular contour shapes and then allocates attention to these, different neurons should experience attentional enhancement to the *same* shapes—the shapes that the monkey notices. In our data, however, neurons that preferred cocircular contours under the circle detection task, for example, did not prefer the *same* circular contours. For instance, some neurons were strongly selective for small, closed circles whereas others preferred broad arcs. Moreover, the optimal contours for neurons tended to resemble the targets only in a general sense (curved versus linear) rather than in precise form or orientation. If the geometric selectivity arose from the salience of stimuli that resemble the cue, then the best shapes should have been the cues themselves. A final argument against the alternative interpretation stems from timing considerations. Psychophysical evidence suggests that primates require ~200 ms for object identification and any resultant engagement of atten-

tion to the stimulus (10, 11). Moreover, neurophysiological data demonstrate that spatial attention enhances V4 neurons with an average latency of ~200 ms, whereas the corresponding modulations in V1 are weaker and occur with much longer latencies (12). However, the geometric selectivity we observed in our neurons occurred significantly earlier. Initial modes of broad tuning were expressed within 72 ms of the stimulus onset, and mature, task-dependent patterns of tuning emerged after 110 ms. Taken together, we believe these arguments rule out the concern about conventional attentional effects.

## SI Materials and Methods

**Electrophysiology.** We performed extracellular recordings from single units in V1 of three adult male monkeys (*Macaca mulatta*). We used glass-coated tungsten microelectrodes within an impedance range of 0.8–2.5 MΩ at 1 kHz. Recordings were made from a chamber filled with sterile 4% agarose gel (Sigma-Aldrich; Type I-A, Low EEO), which provided enough mechanical stability to hold isolated units for >1 h. The gel provided a seal between the dura mater, at the bottom of the recording chamber, and the x-y stage for positioning the microelectrode, at the top. All recordings were carried out in the superficial layers of the striate cortex, corresponding to eccentricities between 0.85° and 6.58°. Before beginning the optimization algorithm, we measured the RF position and orientation preference of the recorded neuron using grating and/or bar stimuli; these parameters were later used to position and orient the geometric stimuli.

**Data Acquisition and Stimulus Design.** The stimuli were created by a visual stimulus generator (VSG2/5; Cambridge Research Systems) at a viewing distance of 138 cm. They were displayed on a NANAO monitor (FlexScan F2-21) at a resolution of 1,024 × 769 pixels and with a frame refresh rate of 105 Hz. Neuronal activity was recorded with a spike sorting and acquisition system (Plexon), and the stimulus-elicited spikes from each trial were simultaneously counted by a computer board (PCI-DAS1002; Measurement Computing) for online data analysis.

The stimuli in our experiments used a previously reported paradigm (1, 13, 14), whereby a contour configuration was embedded in a contextual background of line elements. Each stimulus consisted of two circular arrays of 0.55° × 0.55° compartments, into which oriented line elements were positioned. The line elements in each patch were 0.2° × 0.05° in dimension and were displayed at 50% Michelson contrast, with a luminance of 18 cd/m<sup>2</sup>. The arrays were 11 or 13 compartments in diameter, and the bar elements were partitioned into two sets: (i) the contextual line elements, which were randomly oriented and positioned with a small random jitter about the center of their respective compartments, and (ii) the contour elements, whose position and orientation were brought into a particular pattern of geometric alignment with the bar in the central grid cell. The contours that were generated during the automated search for each neuron’s optimum shape were composed of these contour elements.

**Task Design.** The monkeys initiated each trial of the delayed match-to-sample task by gazing at a 0.08° fixation point, while their eye position was monitored at 30 Hz with an infrared tracking system\*. After a 500- to 600-ms prefixation epoch, the monkeys were required to maintain their gaze within a window (0.8°–1.0° in diameter) around the fixation point. After the prefixation period, the monkeys were presented with one of three contours (Fig. 1A, 1) to serve as a cue for the forthcoming task: a closed circle, a sinusoidal wave, or a straight line (Fig. 1A, 2). The experiments were conducted in large blocks of ~1,000 trials in which the cued

\*Matsuda K, et al. (2000). A new system for measuring eye position on a personal computer. *Soc Neurosci Abstr* 26:744.2.

contour was always the same. (The presentation of the cue at every trial was therefore redundant, but was included to reinforce the monkey's familiarity with the cued shape and his attention to the task.) After cuing the monkey for 380 ms and following an additional blank frame (9.5 ms), two fields of line segments were drawn on the computer monitor (Fig. 1B, 1). Embedded in the center of each field was a geometric stimulus, taken from a set of stimuli whose length and shape were adjusted at discrete time points during the recording session (for details see below). The neuronal responses recorded during this stimulus presentation epoch were used by the optimization algorithm to tailor the stimulus set to the neuronal activity. After a variable period of stimulus exposure for each trial (ranging from 191 to 1429 ms), the embedded stimuli were abruptly replaced by two seven-bar "target" contours (Fig. 1C, 1), which were shown for 48–191 ms. (The range of stimulus exposure times and the duration of target presentation were parameters that we varied between experiments. We anticipated that making the time of target presentation both short and unpredictable might increase the monkeys' attentiveness during the stimulus presentation, before the behavioral discrimination. Changing those parameters, however, had no apparent effect on neuronal responses, so we used a constant period of stimulus exposure, 475 ms, for monkey C.) To form the target contour in each patch, seven line elements from the patch, including lines from both the geometric stimulus and the contextual background, were transiently rearranged.

One of these flashed contours could be the cue presented at the beginning of the trial, whereas the other was a distracter contour. (For most of the experiments, the same distracter was used for all trials and under all task conditions. In other experiments, we used either random distracter shapes or distracters that were randomly "jumbled" or distorted versions of the cue, but we never observed any dependence of the animals' behavior or the neural responses on the nature of the distracter.) At the end of each trial, the two stimulus fields were replaced by two dots (saccade targets; Fig. 1C, 2); the monkey's task was to make a saccade toward the dot location where the cued contour was flashed. In most of the experiments, one-third of the trials were catch trials, in which the same distracter contour was flashed in both patches (i.e., the cued contour did not reappear). These trials were included to ensure that the monkey attended to the spatial location of both contextual patches; the correct decision during a catch trial was to look toward a third saccade target, which was present in all trials together with the other two saccade targets (Fig. 1C, 2).

In the majority of our experiments, we allowed the monkey to perform the task with the same cue for the entire recording session, thereby enabling the progressive construction of long contours. In a smaller set of experiments, we probed a neuron's geometric selectivity while the monkey was cued on a particular target shape and then repeated the optimization algorithm after cueing the monkey on a different contour.

**Automated Stimulus Generation.** The optimization algorithm is an iterative procedure for constructing the symmetric configuration of line elements that maximizes the response of the recorded neuron. At the outset of the recording session, the optimal configuration contained only one contour element: the central bar in the RF, surrounded by the field of random contextual lines. The central bar was optimally positioned and oriented to match the neuron's RF properties. At the first iteration, the algorithm searched for

the optimal positions and orientations of two adjacent contextual bars, each spaced a center-to-center distance of 0.55° from the central bar. The contextual bars were always placed symmetrically on either side of the bar in the RF. After settling on the optimal arrangement of the first two adjacent bars, the algorithm made their configuration, together with the central bar, permanent throughout the remainder of the experiment. Every succeeding stimulus shown to the monkey then contained the optimal three-bar configuration at its core. The next iteration of the algorithm searched for the best arrangement of the next two bar elements, each of which was spaced 0.55° from either end of the already established contour (Fig. S2). Proceeding in this way, the algorithm could execute an arbitrary number of iterations, each time extending the number of bars in the constructed contour. However, time constraints and diminishing neuronal selectivity for distal contour elements obliged the termination of the program after it settled on the optimal five- or seven-bar contour.

At each iteration, the algorithm generated an initial stimulus set intended to comprehensively sample the stimulus space for the next pair of line segments to add to the contour. The routine repeatedly tested each of these stimuli. After 5 or 7 randomly interleaved repetitions of each stimulus, the algorithm removed half of the stimuli from the test set, eliminating those stimuli that elicited the lowest firing rates from the recorded neuron. The remaining stimuli were tested for 15–21 additional trials, and the stimulus that elicited the highest firing rate was selected. The algorithm then carried out a "refinement" step, constricting the region of the stimulus space it would search around the selected stimulus. The refinement step tested a new, smaller set of stimuli that were closer to the selected stimulus than the stimuli in the previous set. After testing each of these new stimuli 20 or 28 times, together with the best stimulus from the previous set, the program selected the most effective stimulus and could execute another refinement step by further constricting the stimulus space around the current optimum. At each of these stimulus space constrictions, the differences in the position and orientation of the contour elements in the test stimuli were progressively decreased until a good approximation to the optimal configuration was found. We typically configured the optimization routine to use 1 refinement step per iteration before fixing the best configuration and moving on to the next set of contour bars.

The initial stimulus set at each iteration contained 32 automatically generated stimuli, with four different bar orientations at each of eight spatial locations on each side of the contour (Fig. S2). Let  $\ell$  be the collinear axis that extends outward from the bar at the right edge of the existing contour; let the point  $p$  be the center of that bar. Now consider the circular arc  $C$ , such that the distance between  $p$  and  $C$  is 0.55°,  $C$  is bisected by  $\ell$ , and the endpoints of  $C$  form a central angle of 110° with  $p$ . The eight spatial locations on the right side of the contour were evenly spaced between the endpoints of  $C$ , covering a broad range of positions; and the four orientations at each position were 0°, 45°, 90°, and 135°, relative to the orientation of the bar in the RF. The bars on the left end of the contour were mirror symmetric to the bars on the right edge.

Let  $q$  and  $\theta$  be the optimum position and orientation of the outermost contour bar from a previous stimulus set,  $S$ . Further, let  $d$  be the difference in polar angle between adjacent bar positions, and let  $\Delta$  denote the smallest difference between different bar orientations, for the stimuli in  $S$ . If  $\Omega$  represents the set of all previously tested stimuli, then the new stimulus set generated by a refinement step would be

$$R = \{(\theta + \Delta, q + d), (\theta, q + d), (\theta - \Delta, q + d), (\theta + \Delta, q), (\theta - \Delta, q), (\theta + \Delta, q - d), (\theta, q - d), (\theta - \Delta, q - d)\} / \Omega,$$

where the set operation  $\{\cdot\}/\Omega$  denotes all of the members of the bracketed set that are not members of the set  $\Omega$ . If  $R$  was the empty set, then the algorithm would replace it with

$$R = \{(\theta + \Delta/2, q + d/2), (\theta, q + d/2), (\theta - \Delta/2, q + d/2), \\ (\theta + \Delta/2, q), (\theta - \Delta/2, q), (\theta + \Delta/2, q - d/2), \\ (\theta, q - d/2), (\theta - \Delta/2, q - d/2)\}.$$

In addition to the stimuli that were tailored to the neural responses, the optimization algorithm often included several predefined, geometrically pertinent shapes in the stimulus set. When generating the first set of three-bar contours, the stimuli often included the innermost three bars of the cued target shape (and the cue rotated by 180°). For monkey C (but not A or B), the algorithm also included in the initial stimulus set a one-bar stimulus, consisting only of the bar in the RF, surrounded by the random field. This stimulus was identical to the three-bar stimuli, except that none of the bars in the random field were brought into a particular alignment with the bar in the RF. Finally, when generating the first set of five-bar stimuli, the program included the innermost five bars (middle segments) of all three cues, the false target distracter, and their 180° rotations.

**3D Tuning Surfaces.** Geometric stimuli were mapped onto the 2D plane by specifying each stimulus as a function of its outermost line segments. Because the stimuli were all mirror symmetric, only one of the two outermost contour bars was needed to specify each stimulus. To compare the tuning surfaces between neurons, we expressed the stimuli in a rotated Cartesian coordinate system, where the central bar in the RF was collinear with the horizontal axis. The outermost bar on the right arm of any stimulus in this coordinate system was then chosen to specify the stimulus geometry. Each stimulus was specified by two angular coordinates: the orientation of the rightmost contour bar with respect to the horizontal ( $\alpha$ ) and the polar angle of the bar's position relative to the center of the RF ( $\beta$ ) (Fig. S2).

In the three-bar tuning surfaces, positive values of the  $\beta$  coordinate always represent stimuli with the same “orientation” as the cued target and the distracters (i.e., stimuli that “face the same direction” as the cued contour). Conversely, negative values of  $\beta$  describe stimuli with the reverse orientation (i.e., stimuli that are rotated by 180° relative to the cued target).

**Five-Bar Population Analysis with Sammon’s Mapping.** Sammon’s mapping (15) projects high-dimensional objects, like our five-bar stimuli, into two dimensions by minimizing the following cost function:

$$E_S = \frac{1}{\sum_i \sum_{j>i} d_{ij}} \sum_i \sum_{j>i} \frac{(d_{ij} - d'_{ij})^2}{d_{ij}}.$$

Here,  $d_{ij}$  is our measure of “distance” between stimulus  $i$  and stimulus  $j$ ; and  $d'_{ij}$  is the Euclidean distance between the 2D points that represent stimuli  $i$  and  $j$ .

We define the distance between any two stimuli as the amount of “energy” required to deform one stimulus configuration into the other. We represent the stimuli as one-dimensional, piecewise linear functions of arc length,  $t$ . Let  $\vec{S}_i(t) : [0, 1] \mapsto \mathbf{R}^2$  be the function representing stimulus  $i$ , which consists of the contour bars that make up the original stimulus, interconnected by straight line segments that adjoin adjacent bars. The center of each stimulus is always taken to reside at the origin, and each stimulus is rotated to the same orientation. Let the difference between stimuli  $i$  and  $j$  be represented by  $\vec{D}_{ij}(t) = \vec{S}_i(t) - \vec{S}_j(t)$ , and let  $(\vec{D}_{ij}(t))_x$  and  $(\vec{D}_{ij}(t))_y$  be the  $x$  and  $y$  components of the

difference function  $\vec{D}_{ij}(t)$ . Then our measure of the elastic deformation distance between stimuli  $i$  and  $j$  is

$$d_{ij} = \int_0^1 (\vec{D}_{ij}(t))_x^2 + (\vec{D}_{ij}(t))_y^2 dt \\ + \int_0^1 \left( \frac{d(\vec{D}_{ij}(t))_x}{dt} \right)^2 + \left( \frac{d(\vec{D}_{ij}(t))_y}{dt} \right)^2 dt.$$

**Permutation Test.** Let  $\{\vec{F}_1, \vec{F}_2, \dots, \vec{F}_N\}$  and  $\{\vec{G}_1, \vec{G}_2, \dots, \vec{G}_M\}$  be two sets of tuning surfaces that were averaged together to form the mean surfaces  $\vec{T}_1$  and  $\vec{T}_2$ . Here, each surface is represented as a vector that samples the surface height at the stimulus locations that were tested, over all of the experiments, from the corresponding monkey(s). The permutation test between  $\vec{T}_1$  and  $\vec{T}_2$  was performed by pooling all of the individual surfaces into the set  $\{\vec{F}_1, \vec{F}_2, \dots, \vec{F}_N, \vec{G}_1, \vec{G}_2, \dots, \vec{G}_M\}$  and then randomly reassigning each surface into one of two subsets, of length  $N$  and  $M$ . The surfaces in these randomly permuted subsets were averaged together, and a test statistic computed on these mean surfaces was compared with the statistic computed on  $\vec{T}_1$  and  $\vec{T}_2$ . The random permutation and averaging procedure was performed 25,000 times; the  $P$  value was reported as the probability that the test statistic from the permuted data was as small as that from the original datasets.

Let  $\vec{T}_{1,i}$  and  $\vec{T}_{2,i}$  be the mean neural responses to stimulus  $i$ . Additionally, let  $D_i$  be the distance between stimulus  $i$  and the nearest cocircular stimulus. If  $D_i$  is computed as the elastic deformation energy between stimulus  $i$  and its nearest cocircular neighbor, then our test statistic is

$$TS = \sqrt{\sum_{\{i|D_i \leq \delta\}} (\vec{T}_{1,i} - \vec{T}_{2,i})^2}.$$

This test statistic measures the Euclidean distance between the mean tuning surfaces along the cocircularity lines. We used  $\delta = 0.2$  (arbitrary units) to focus the statistical comparisons close to the cocircularity lines, but the results from the permutation tests were not sensitive to this parameter. The components of the neural response vectors near the cocircularity line were linearly mapped into the range  $[0, 1]$ , so that

$$\max_i \vec{T}_{1,i} = \max_i \vec{T}_{2,i} = 1 \\ \min_i \vec{T}_{1,i} = \min_i \vec{T}_{2,i} = 0$$

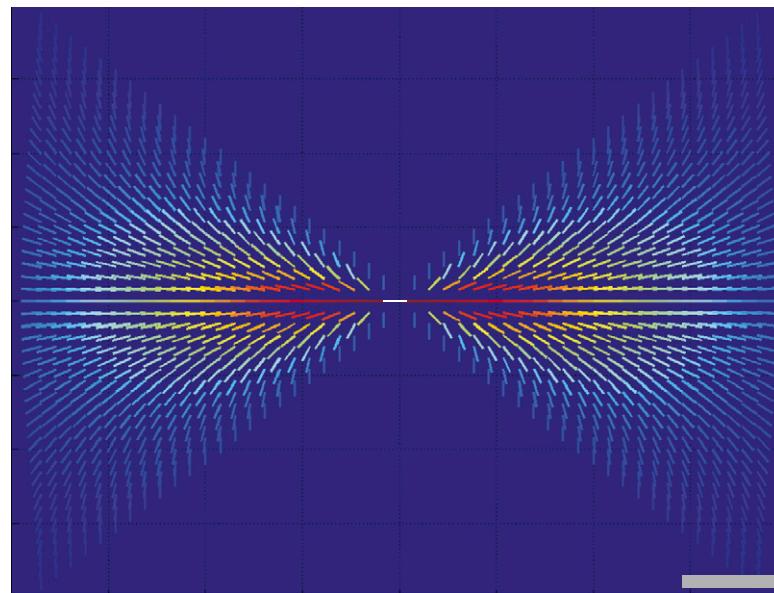
for all  $i$  such that  $D_i \leq \delta$ .

**Two-Way ANOVA.** We applied Friedman’s test (Matlab Statistics Toolbox; The MathWorks) to examine the time course of geometric selectivity. Friedman’s test is a nonparametric alternative to the standard two-factor ANOVA; it avoids the normality assumption, but it still assumes that all of the data derive from distributions with the same variance. Because the variance in neuronal spiking is proportional to the mean response, we applied Friedman’s test to the square root of the neuronal spike counts, rather than to the spike counts themselves. When applied to data that are distributed according to a Poisson or a Gamma distribution, the square root operation is a variance stabilizing transformation: It adjusts the data so that samples with different means will have approximately the same variance. The transformed spike counts in a given time window were then grouped according to two “factors”: the stimulus that elicited the spikes (factor A) and the neuron from which the spikes were recorded (factor B). In our dataset, we had 32 “levels” of factor

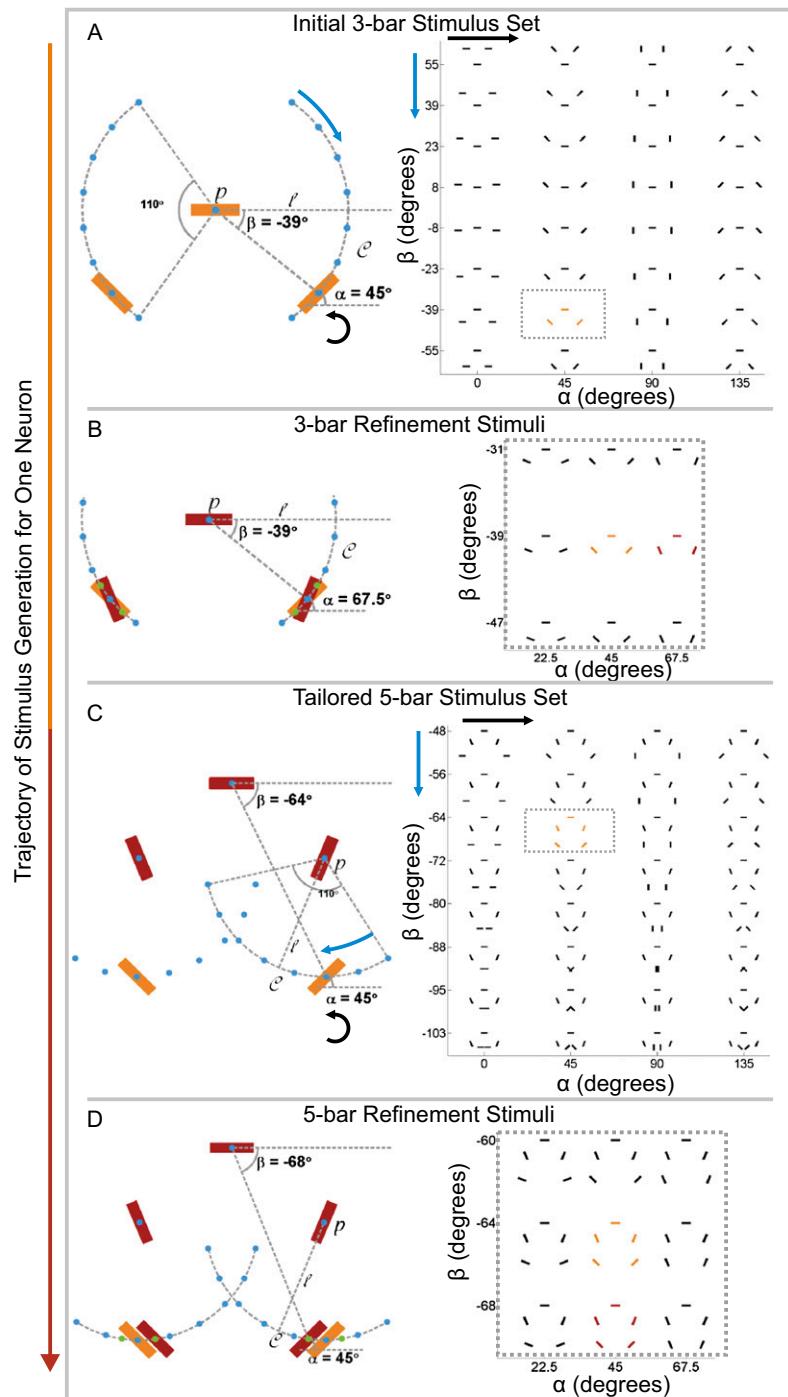
A (i.e., 32 different stimuli), 22 levels of factor B, and 40 observations per treatment (i.e., 40 trials per stimulus, for each neuron). Friedman's test corrects for the extraneous variability introduced by the different levels of factor B and tests for the presence of a "main effect" due to factor A, over all of the

supplied observations. In the context of our data, this means that Friedman's test corrects for the variability of neural responses from cell to cell (and task to task) and tests the null hypothesis that all 32 stimuli elicit the same neural response (within a given time window).

1. Li W, Pięch V, Gilbert CD (2006) Contour saliency in primary visual cortex. *Neuron* 50: 951–962.
2. Summerfield C, Egner T (2009) Expectation (and attention) in visual cognition. *Trends Cogn Sci* 13:403–409.
3. Bar M, et al. (2006) Top-down facilitation of visual recognition. *Proc Natl Acad Sci USA* 103:449–454.
4. Epshtain B, Lifshitz I, Ullman S (2008) Image interpretation by a single bottom-up top-down cycle. *Proc Natl Acad Sci USA* 105:14298–14303.
5. Ullman S (2007) Object recognition and segmentation by a fragment-based hierarchy. *Trends Cogn Sci* 11:58–64.
6. David SV, Hayden BY, Mazer JA, Gallant JL (2008) Attention to stimulus features shifts spectral tuning of V4 neurons during natural vision. *Neuron* 59:509–521.
7. Ghose GM, Bearl DW (2010) Attention directed by expectations enhances receptive fields in cortical area MT. *Vision Res* 50:441–451.
8. Reynolds JH, Chelazzi L (2004) Attentional modulation of visual processing. *Annu Rev Neurosci* 27:611–647.
9. Maunsell JHR, Treue S (2006) Feature-based attention in visual cortex. *Trends Neurosci* 29:317–322.
10. Fabre-Thorpe M, Delorme A, Marlot C, Thorpe S (2001) A limit to the speed of processing in ultra-rapid visual categorization of novel natural scenes. *J Cogn Neurosci* 13:171–180.
11. Maldonado P, et al. (2008) Synchronization of neuronal responses in primary visual cortex of monkeys viewing natural images. *J Neurophysiol* 100:1523–1532.
12. Buffalo EA, Fries P, Landman R, Liang HL, Desimone R (2010) A backward progression of attentional effects in the ventral stream. *Proc Natl Acad Sci USA* 107:361–365.
13. Kapadia MK, Ito M, Gilbert CD, Westheimer G (1995) Improvement in visual sensitivity by changes in local context: Parallel studies in human observers and in V1 of alert monkeys. *Neuron* 15:843–856.
14. Li W, Gilbert CD (2002) Global contour saliency and local colinear interactions. *J Neurophysiol* 88:2846–2856.
15. Kohonen T (2001) *Self-Organizing Maps* (Springer, Berlin), 3rd Ed.



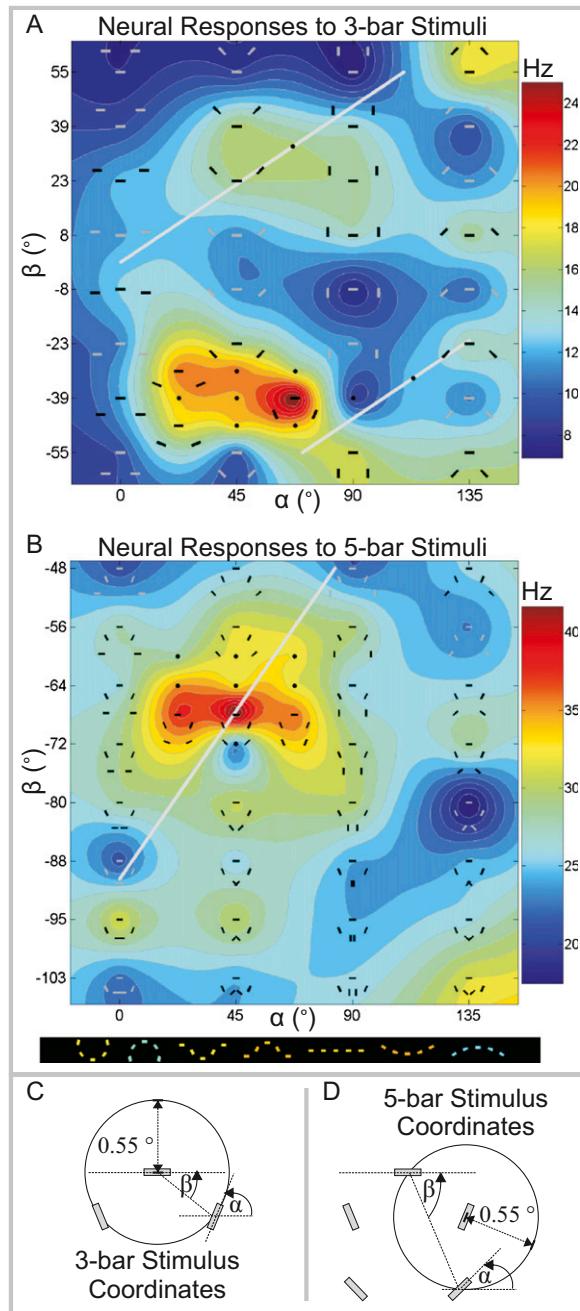
**Fig. S1.** A depiction of the theoretical association field (AF), showing the optimal contextual interactions for the white line at the center of the figure. In the AF theory, the saliency of smooth contours derives from a broad set of local interactions between line segments; any line segment can be perceptually linked with a range of other segments, and salient contours are characterized by local segments whose interactions sum up over the length of the contour. Given the relative positions of any two line segments, the strength of the interaction, or linkage, between them is a function of their relative orientations. Here, given the position and orientation of the white bar, we show the orientation of the line segment at each spatial position that elicits the strongest interaction with the white bar; the color of each segment specifies the strength of that linkage (warm colors, strong connection; cold colors, weak connection). In the standard AF model, the optimal interactions are between "cocircular" segments, which lie tangentially along the same circular arc. The strongest of these effects are between nearby segments that lie along a circle with an infinite radius (i.e., a straight line); as the radius of the circular arc connecting the two segments decreases, the linkage becomes weaker. Psychophysically, the strength of the interaction between line segments refers to the tendency of the visual system to bind them together into a perceptual whole. Physiologically, it refers to the connection strength between the neurons that are activated by the line segments. (Gray scale bar: 0.5°.)



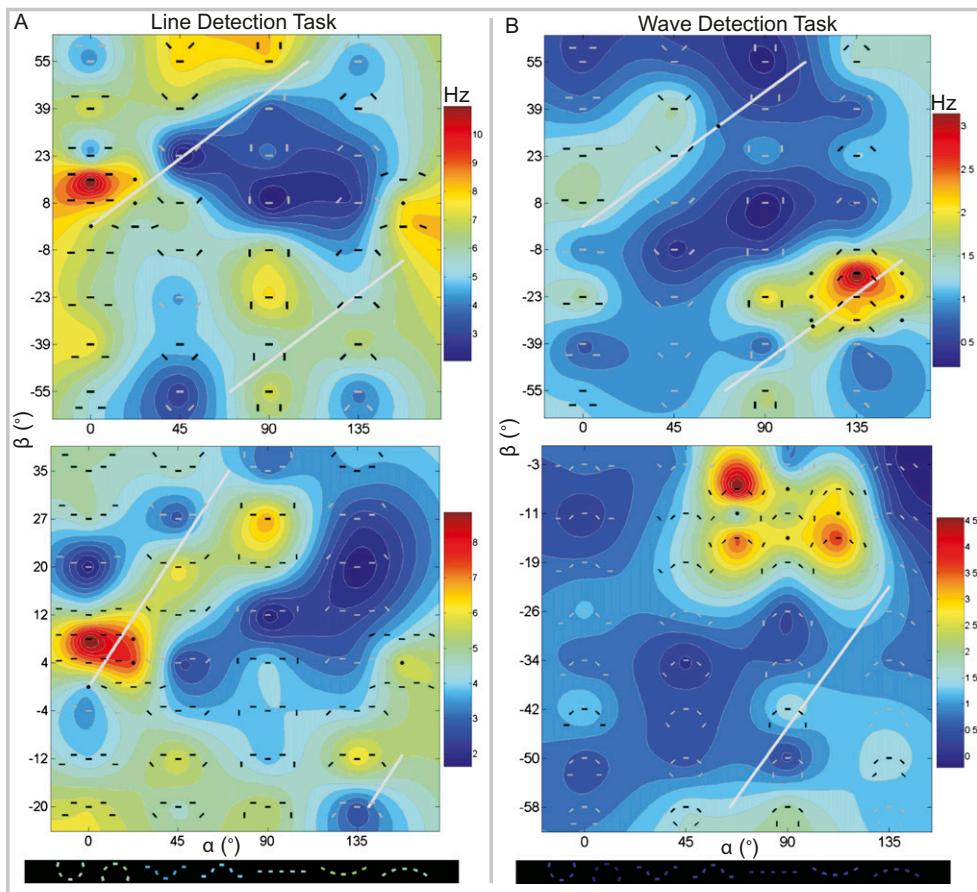
**Fig. S2.** Optimization algorithm for stimulus generation. (A–D) The geometric stimuli that were created by the optimization algorithm for an example neuron recorded under the wave detection task. (The cell's responses to these stimuli are shown in Fig. S3.) The Right side of each panel illustrates the stimuli that were generated during each stage of the optimization algorithm; the Left side depicts how these stimuli were created. As always, the stimuli are drawn in a rotated coordinate system, where the horizontal axis is rotated to the cell's preferred orientation. All stimuli are defined by two coordinates:  $\alpha$ , the orientation of the bar on the right arm of the contour; and  $\beta$ , the polar angle (position) of the right contour bar with respect to the central bar in the classical RF. The stimuli always had bilateral symmetry, so the position and orientation of the bar on one end of the contour determine the geometry on its opposite end. The letters  $p$ ,  $l$ , and  $c$  refer to variables that are defined in *SI Materials and Methods*. (A) The initial stimulus set was identical for all neurons. It consisted of 32 distinct three-bar stimuli. The central bar was centered in the classical RF; the bars outside the minimum response field took one of eight positions (Left, blue dots), with four possible orientations at each position ( $0^\circ$ ,  $45^\circ$ ,  $90^\circ$ , and  $135^\circ$ ), making for 32 different stimuli (Right). The stimulus from this set that elicited the strongest response from the example neuron is shown in orange ( $\alpha = 45^\circ$ ,  $\beta = -39^\circ$ ). (B) The stimulus set generated after the best contour was selected from the original stimuli. This set of refinement stimuli consisted of 8 new stimuli, in addition to the best stimulus from the previous stage. The new stimuli took one of three positions, indicated by the green points and the preferred blue position from the preceding stage (Left), and one of three bar orientations: ( $22.5^\circ$ ,  $45^\circ$ , or  $67.5^\circ$ ). Closely spaced in both position and orientation, the stimuli were chosen to refine the estimate of the cell's preferred three-bar stimulus. The final estimate (red) had the same bar position as the orange contour, but a bar orientation of  $67.5^\circ$  rather than  $45^\circ$ . (C) The five-bar stimuli generated during the third stage of the optimization algorithm, using the optimum three-bar stimulus as a seed configuration to search for longer contours. The stimulus set was

Legend continued on following page

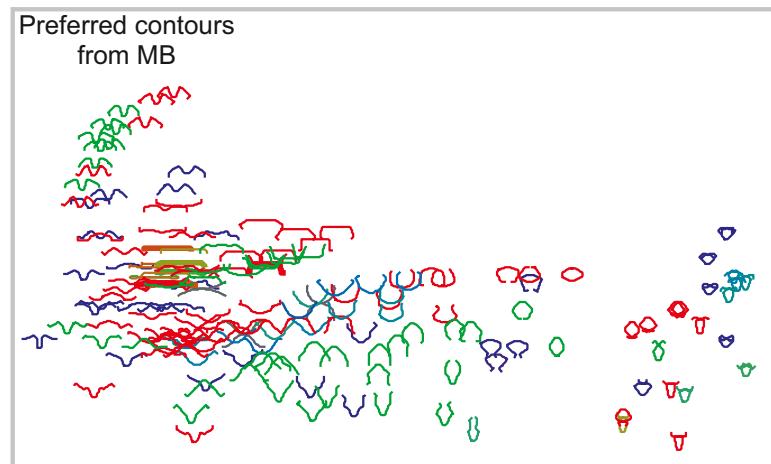
constructed as in A, but by using the position and orientation of the bar on the right side of the optimum three-bar contour, rather than the central bar in the RF, as the basis for the relative positioning and orientation of the fourth and fifth bars. Orange: The best configuration from this stimulus set. (D) The five-bar refinement stimuli, created around the best stimulus from the previous stage. Red, the improved estimate for this cell's five-bar optimum.



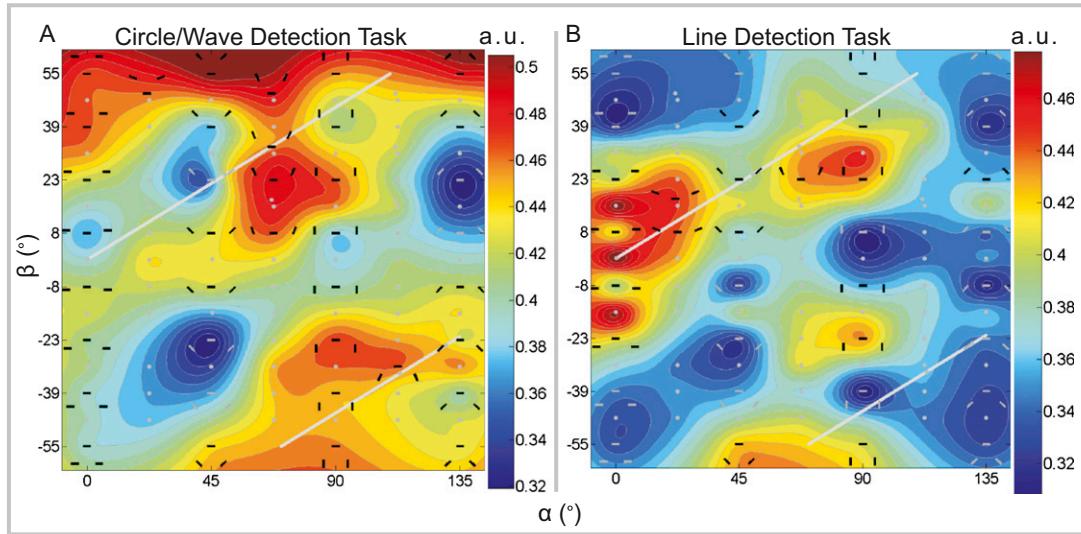
**Fig. S3.** Representation of neural shape selectivity: The responses of a single neuron (from monkey C) to the dynamically generated stimuli shown in Fig. S2. (A) The neural response function, represented as a heat map, for the three-bar stimuli. (B) The response map for the generated five-bar stimuli. (A and B) The heat maps were created by plotting the neural response to each tested stimulus and interpolating the data over the stimulus plane with Hardy's multiquadratics. The magnitude of the neural response is given in Hertz. As in Fig. 2, the stimuli are defined by a pair of angular coordinates ( $\alpha, \beta$ ), and the neural responses are plotted as a function of these defining stimulus parameters. For ease of visualization, the tested geometric stimuli are drawn over the data; for each contour, the position of the central bar corresponds to the ( $\alpha, \beta$ ) coordinates of the stimulus. The contours from both the initial and the refinement stimulus sets are drawn, but wherever the positions of stimuli in the  $\alpha$ - $\beta$  plane are too close to clearly show all of the stimuli, the most effective stimuli are drawn explicitly whereas the others are indicated with points. (The intensity of the stimuli atop the tuning surfaces was adjusted to improve their visibility.) The gray "cocircularity" lines are the loci of stimuli, with various radii of curvature, whose outermost contour bars are cocircular with the central bar in the classical RF. This neuron exhibited clear peaks of facilitation near the cocircularity lines on the three- and five-bar heat maps; the optimum five-bar contour for the cell resembled a closed circular contour and was situated directly on the cocircularity line. (B) The contours below surface are the five-bar versions of the target and distracter contours; the neural response to each contour is indicated by the color of the stimulus bars. (C and D) Schematics illustrating the meaning of the stimulus coordinates for three- and five-bar stimuli, using the optimum three-bar (C) and five-bar (D) contours as examples. The geometric stimuli are always drawn as though they were horizontally oriented, but the actual stimuli were rotated so that the central bar matched the preferred orientation of the recorded neuron.



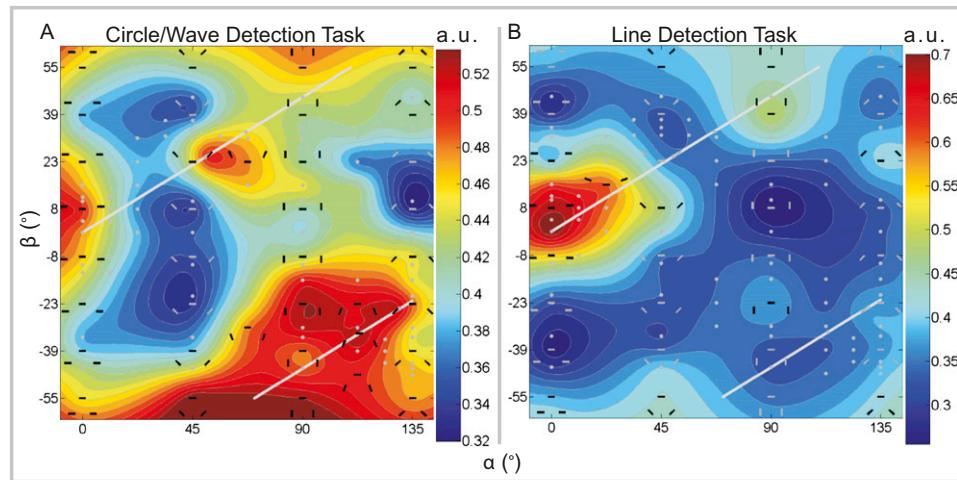
**Fig. S4.** Additional examples of linear and wave-like facilitation. (A) Three- and five-bar response functions recorded from a neuron in monkey C during the line detection task. (B) Three- and five-bar heat maps recorded from a neuron in monkey B during the wave detection task. Both neurons were highly selective for the optimal contours created by the stimulus generation algorithm and were largely quiescent to the predefined stimulus set of target and distracter contours. The neuron in A was specifically activated by an approximately collinear contour with a small lateral offset. The neuron in B has an optimum on the three-bar cocircularity line, but demonstrates a strong preference for a disjointed, wave-like optimum on the five-bar surface far from cocircularity.



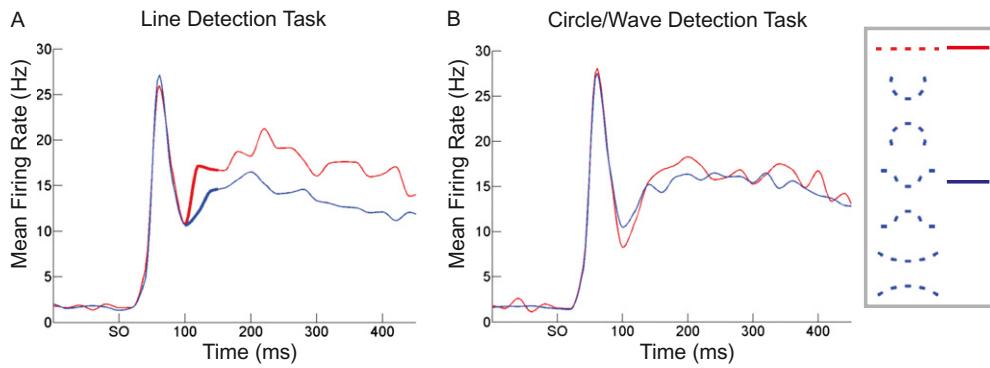
**Fig. S5.** The preferred five-bar contours generated by the neurons recorded from monkey B (MB). The stimuli are plotted according to the same algorithm used to construct Fig. 5. The distribution of optimal contour shapes from MB closely resembles that obtained from monkey C, although the task dependence of the distribution is absent in MB. [For MB, the influence of top-down control over shape processing is apparent in the population optima (Fig. S5) rather than the optima for single neurons plotted individually, as done here.]



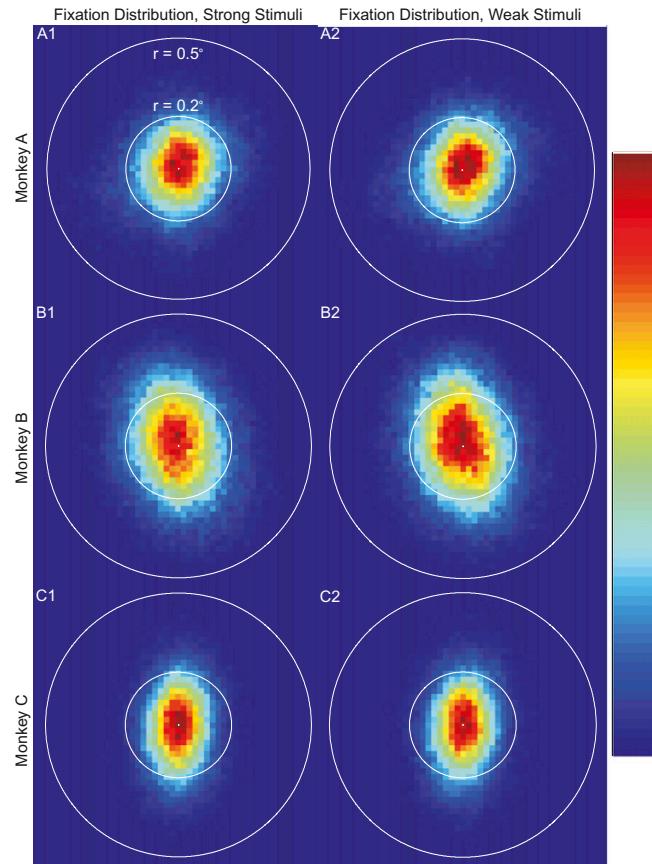
**Fig. S6.** The mean three-bar heat maps averaged over the neurons recorded from monkey B (MB) during the circle and wave detection tasks (A) and during the line detection task (B). The heat maps for each cell were normalized and averaged according to the same procedures applied in Fig. 6 A–D. The task-dependent differences between the mean neural responses are statistically significant ( $P = 0.007$ ; permutation test) and they follow the same trend as that for monkeys A and C: Under the line detection task, the population expressed a response peak near collinearity; under the curved detection tasks, the optimal network response was engendered by stimuli along the cocircularity lines. Sample sizes: (A)  $n = 33$  and (B)  $n = 30$ .



**Fig. S7.** Mean three-bar heat maps, averaged over all cells that were recorded under different task conditions during a single experiment. (A) The mean response surface obtained from 13 cells while monkeys A and B performed either the circle or the wave detection task. (B) The mean neural response recorded from the same 13 cells during the line detection task. The heat maps were constructed with the analysis methods used in Fig. 6 A–D and Fig. S6. The task-dependent differences between the mean heat maps are statistically significant ( $P = 0.013$ ; permutation test) and in the same direction observed when populations of different neurons are observed under the corresponding task conditions (Fig. 6 A–D and Fig. S6). The data shown here, which include the cell in Fig. 6 E and F, demonstrate directly that the same neurons dynamically change their geometric tuning as a function of behavioral task. This dynamic tuning, when integrated over the population of neurons, shifts the optimal network response according to cognitive state. Under the line detection task, the cumulative neural response was dominated by a sharp peak of collinear facilitation. For these 13 cells, a local maximum near collinearity remained in the population response even under the circle and wave tasks, but the magnitude of the peak was lower (~0.5, compared with ~0.7) and the bulk of facilitation was dispersed along the cocircularity lines. Sample sizes: monkey A,  $n = 7$ ; monkey B,  $n = 6$ .

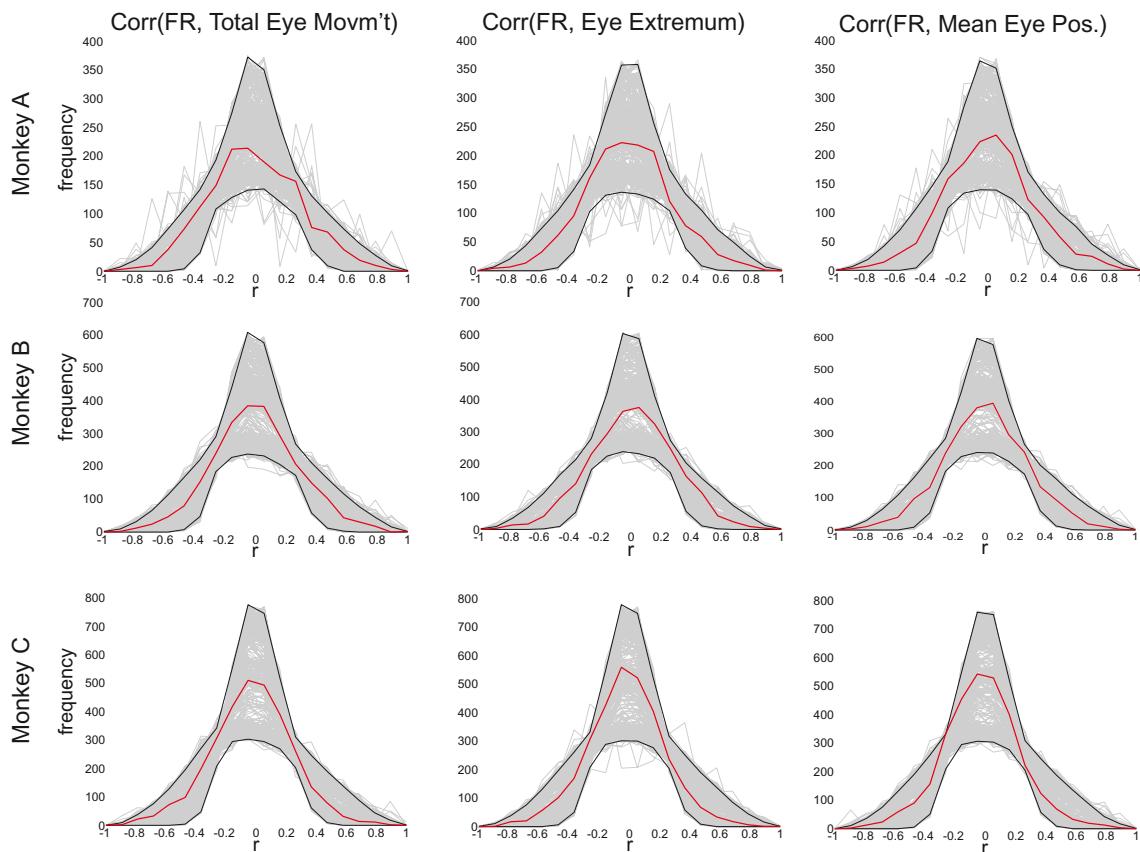


**Fig. S8.** Time course of five-bar shape selectivity. (A and B) Peri-stimulus time histograms: The response to the five-bar line is shown in red, and the mean response to the circle, wave, and distracter is shown in blue, under task conditions when the animal looked for the line (A,  $n = 38$ ), or for the circle or wave (B,  $n = 67$ ). SO, stimulus onset (time 0). The spikes elicited by each stimulus, over a population of neurons, were pooled into 20-ms time bins and fit with a smooth interpolant. Under the line detection task—but not under the circle or wave tasks—neurons were preferentially facilitated by the five-bar line, compared with the segments from the other cues and the distracter. This task-dependent selectivity emerged rapidly following the rebound from contextual inhibition. The highlighted portions of the red and blue PSTHs in A are significantly different ( $P = 0.0039$ ; one-sided Wilcoxon signed-rank test; time window, [100, 160] ms).



**Fig. S9.** The eye position distributions during the presentation of the strongest (Left) and weakest (Right) five-bar stimuli, for all three monkeys. (Left) Each distribution was formed by pooling the eye traces recorded during the display of the five most effective stimuli for each cell. (Right) Eye traces pooled from a corresponding number of the weakest stimuli. White dot, fixation point (FP); outermost white circle, fixation window, radius =  $0.5^\circ$ ; inner circle, drawn for reference, radius =  $0.2^\circ$ . Mean eye movement speed (eye trace length/time, averaged over trials) in degrees per second: A, 1, 1.43; A, 2, 1.41; B, 1, 1.30; B, 2, 1.35; C, 1, 1.22; C, 2, 1.24. The monkeys tended to focus their fixation within  $\sim 0.2^\circ$  of the FP; their behavior was the same for both strong and weak stimuli.

### Trial-by-Trial Correlations Between Firing Rate (FR) and Eye Movement



**Fig. S10.** Pearson correlation coefficients between neuronal responses to geometric stimuli and three features of the monkeys' eye traces. Rows show data for each monkey and columns show correlations between neural activity and different attributes of the eye movements. Red line: Distribution of observed correlations between sets of stimulus responses and simultaneously recorded eye traces (see below). Gray lines: Chance distributions obtained by randomly reshuffling "matching" eye traces and stimulus responses and then recalculating the correlation coefficients. Each set of data was randomly reshuffled 500 times. The distributions computed from these reshuffled data were ranked according to their peak height (a simple measure of their overall shape); the random distributions falling between the 2.5th and 97.5th percentiles are displayed. Black lines: 95% confidence bounds for chance correlations, computed point-by-point along the distributions. (Left) Correlations between neural responses and the total length of the eye position trajectory in a trial, normalized by the recording time (i.e., eye movement speed). (Center) Activity correlations with the maximum displacement of the eye position away from the fixation point within a trial. (Right) Firing rate correlations with the distance of the mean eye position from the fixation point. The correlations were all computed in the following way. Let  $n_i = \{n_{i,1}, n_{i,2}, \dots, n_{i,S}\}$  be the set of neural responses to the  $i$ th presentation of each contour in a given stimulus set ( $n_{i,j}$  is the response to the  $j$ th stimulus during trial  $i$ ).  $n_i$  is therefore a single-trial slice through the measured neural response function, consisting of a series of  $S$  trials recorded in close temporal succession (while the firing rate was stationary). The distributions show the Pearson correlation coefficients between each of these single-trial series and the simultaneously recorded eye traces, from all experiments in which the optimization algorithm constructed five-bar contours. Here, the Pearson coefficient measures the relationship between eye movements and the neural selectivity for one stimulus over another. We found no systematic correlations between any feature of the monkeys' eye movements and their neural responses, beyond the chance correlations we would expect if the relationship between response rate and eye movement were completely random.