CellPress

Neuron
Article

Incremental Integration of Global Contours through Interplay between Visual Cortical Areas

Minggui Chen,[1,4] Yin Yan,[1,4] Xiajing Gong,[2] Charles D. Gilbert,[3] Hualou Liang,[2] and Wu Li[1,*]
[1]State Key Laboratory of Cognitive Neuroscience and Learning, IDG/McGovern Institute for Brain Research, and Center for Collaboration and Innovation in Brain and Learning Sciences, Beijing Normal University, Beijing 100875, China
[2]School of Biomedical Engineering, Drexel University, Philadelphia, PA 19104, USA
[3]Laboratory of Neurobiology, The Rockefeller University, New York, NY 10065, USA
[4]Co-first Authors
*Correspondence: liwu@bnu.edu.cn
http://dx.doi.org/10.1016/j.neuron.2014.03.023

SUMMARY

The traditional view on visual processing emphasizes a hierarchy: local line segments are first linked into global contours, which in turn are assembled into more complex forms. Distinct from this bottom-up viewpoint, here we provide evidence for a theoretical framework whereby objects and their parts are processed almost concurrently in a bidirectional cortico-cortical loop. By simultaneous recordings from V1 and V4 in awake monkeys, we found that information about global contours in a cluttered background emerged initially in V4, started ~40 ms later in V1, and continued to develop in parallel in both areas. Detailed analysis of neuronal response properties implicated contour integration to emerge from both bottom-up and reentrant processes. Our results point to an incremental integration mechanism: feedforward assembling accompanied by feedback disambiguating to define and enhance the global contours and to suppress background noise. The consequence is a parallel accumulation of contour information over multiple cortical areas.

INTRODUCTION

To generate a coherent representation of visual objects, global contours delineating their outlines must be assembled properly. This contour integration process follows the Gestalt rule of good continuation (Wertheimer, 1923). The classical bottom-up point of view posits that V1 performs the local analysis of orientation within their small receptive fields (RFs) and that the integration of line segments into complex shapes occurs in higher-order visual cortical areas with increasing complexity in neuronal response properties. But against this idea is evidence for the intermediate-level process of contour integration occurring as early as V1 (Bauer and Heinze, 2002; Gilad et al., 2013; Kapadia et al., 1995; Li et al., 2006). The morphology of long-range horizontal connections in V1 is also suitable for mediating interactions between neurons with a similar orientation preference (Bosking et al., 1997; Gilbert and Wiesel, 1979; Rockland and Lund, 1982; Stettler et al., 2002). This intracortical circuitry is postulated to underlie contour integration by many computational models (e.g., Li, 1998; Piëch et al., 2013; Ursino and La Cara, 2004).

It has been shown that responses of individual V1 neurons are closely correlated with perceptual saliency of global contours (Li et al., 2006), suggesting that V1 is intimately involved in contour integration. However, the contour-related neuronal responses in V1 are delayed with respect to the initial response latency. Such a delay could be due to recurrent lateral interactions mediated by slow horizontal connections (Bringuier et al., 1999), to feedback signals from higher cortical areas (Buffalo et al., 2010; Poort et al., 2012), or to the time required for state changes in the cortical network that are independent of conduction times (Piëch et al., 2013).

Neurons in higher-order visual areas have larger RFs that are selective for more complex shapes from their response outsets (Desimone et al., 1984; Pasupathy and Connor, 1999). Though this could be accomplished by selective weighting of local feedforward inputs, long-range horizontal connections also confer the ability for neurons in early cortical areas to integrate information over large areas of the visual field; even within V1, neurons show selectivity for complex shapes (McManus et al., 2011). The problem of contour integration becomes more challenging in complex scenes, which requires linking line segments belonging to the same object and separating them from background components. This is difficult to achieve by feedforward processes alone because of potential ambiguities. To overcome this difficulty, an alternative theoretical framework involves creating a coarse template in higher cortical areas and feeding this template back to early areas, which then can selectively enhance neural responses to image components related to object contours and suppress the influence of unrelated features (Epshtein et al., 2008; Roelfsema, 2006; Ullman, 1984). Such a countercurrent disambiguating process is consonant with the Gestalt rules governing perceptual organization, whereby the global form of an object determines how the brain interprets its composing parts (Wertheimer, 1923).

Although previous studies have suggested the involvement of both lower and higher cortical areas in contour integration, their respective roles and interactions in the integration process remain largely unknown. To address this issue, we used microelectrode arrays to record simultaneously in monkey V1 and

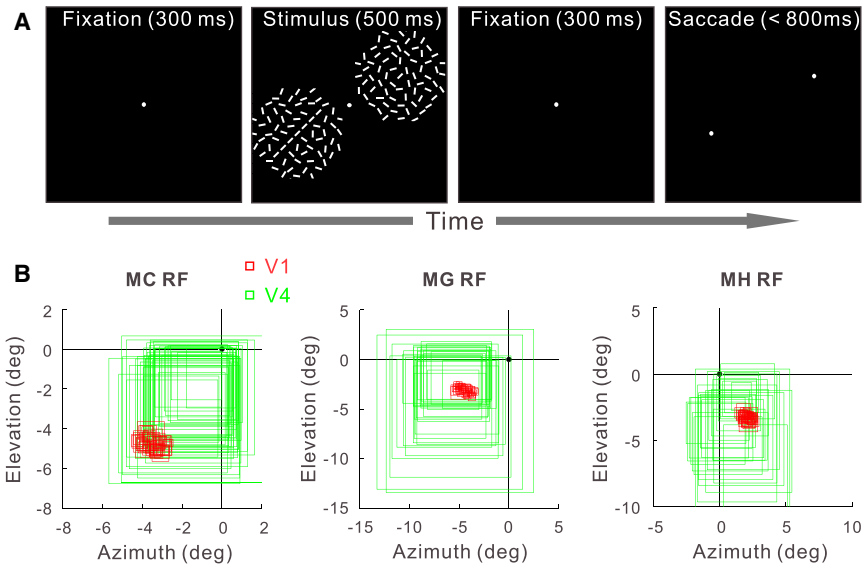

Figure 1. The Experimental Design

(A) A contour-detection trial.

(B) Distribution of the V1 and V4 RFs in the three animals (MC, MG, and MH). Each square indicates the RF of one recording site. See also Figure S1.

V4. We examined the spatiotemporal properties of V1 and V4 neurons in response to the same global contours. Our results suggest two countercurrent processing streams that operate synergistically to augment the contour signals and abate the background noise, resulting in an increment of global contour information in the cortico-cortical loop.

RESULTS

To generate the stimulus with a complex background (Figure 1A), a circular area was divided into square grids, each containing a randomly oriented bar. A global contour was formed by collinearly aligning a number of adjacent bars along a row of grids. The contour orientation was adjusted by rotating the stimulus pattern around its center.

We trained three monkeys (named MC, MG, and MH) in a two-alternative forced-choice task (Figure 1A), in which a contour pattern with an embedded contour and a noise pattern without any contour were presented simultaneously. The animal was required to make a saccade to indicate the location of the contour pattern that was randomly presented at either of the two stimulus locations. Different stimulus conditions in an experiment were randomly mixed within a block of trials, and each condition was repeated in 30 trials unless otherwise stated. While the animal was performing the task, we recorded neuronal activity with microelectrode arrays implanted in corresponding retinotopic regions in V1 and V4 (Figure 1B; Figure S1 available online). Using some rigorous offline sorting algorithms, we selected, for each electrode (i.e., a recording site), the best isolated cluster of spikes, which was further classified as belonging to either a single- or multi-unit (see Supplemental Experimental Procedures; Figure S1B).

The recorded V1 and V4 RFs had mean eccentricities of 5.08° ± 0.92° (mean ±SD) and 4.07° ± 1.77°, respectively, and mean sizes of 0.67° ± 0.19° and 5.23° ± 2.35° (minimum responsive fields mapped by gratings, Figure 1B). Each experiment was repeated in at least two of the three animals. As the results from different animals were qualitatively similar (Figure S2B), data from the same experiment in different animals were pooled unless otherwise indicated. We also combined single- and multi-units in all of the analyses to ensure the power of statistical tests, but pooling the data did not affect the observed results (see data from single units presented in Figure S2A).

Contour-Related Responses in V1 and V4

We examined contour-related responses in V1 and V4 using visual stimuli made of 0.25° by 0.05° bars distributed in 0.5° by 0.5° grids. The diameter of the stimulus pattern was 4.5°. The number of collinear bars in the contour pattern was randomly set to 1, 3, 5, or 7 within a block of trials to control the saliency of the embedded contours. In trials when the number of collinear bars was one, the two stimulus patterns were identical noise patterns; in this case, the animal had to perform the task by guessing at a chance level of 50%. The background bars were rerandomized in each trial.

To examine an individual V1 site, the embedded contour matching the neuron's preferred orientation was either centered on the RF (referred to as a contour site, C-site) (Figure 2, upper left group of insets) or placed laterally 1° away from the RF center (referred to as a background site, B-site) (Figure 2, upper right group of insets). In the latter case, the V1 RF was located on the background, and a background bar was always placed in its center at the optimal orientation. These manipulations made the individually tested V1 sites comparable in this experiment. To examine a V4 recording site, whose RF was on average 7.8 times the size of V1's RF and basically enclosed the entire stimulus pattern, the optimally oriented contour was centered in the RF. Neuronal responses to the noise pattern (i.e., 1-bar contour pattern) were used as a baseline for examining contour-related responses in V1 and V4.

For each recording site we determined its response latency to the onset of the noise pattern from the mean cumulative response curve (defined as the visual-response latency, to be differentiated from the contour-response latency defined later) (see Experimental Procedures; Figure S1D). We also determined the responsiveness of each site by computing the *d*-prime (*d*′) based on the difference in distributions of neuronal firing rates before (−300 to 0 ms) and after (0–500 ms) the display of the noise pattern (defined as the visual-response *d*′, to be differentiated from the contour-response *d*′ defined later). On average, V1 neurons started responding to the noise pattern 44 ms after its onset (Figure 2B, red and green dots; Figure 2C, left red and green bars), and they had a mean visual-response *d*′ of 2.37 (Figure 2D, red and green dots; Figure 2E, left red and green bars);

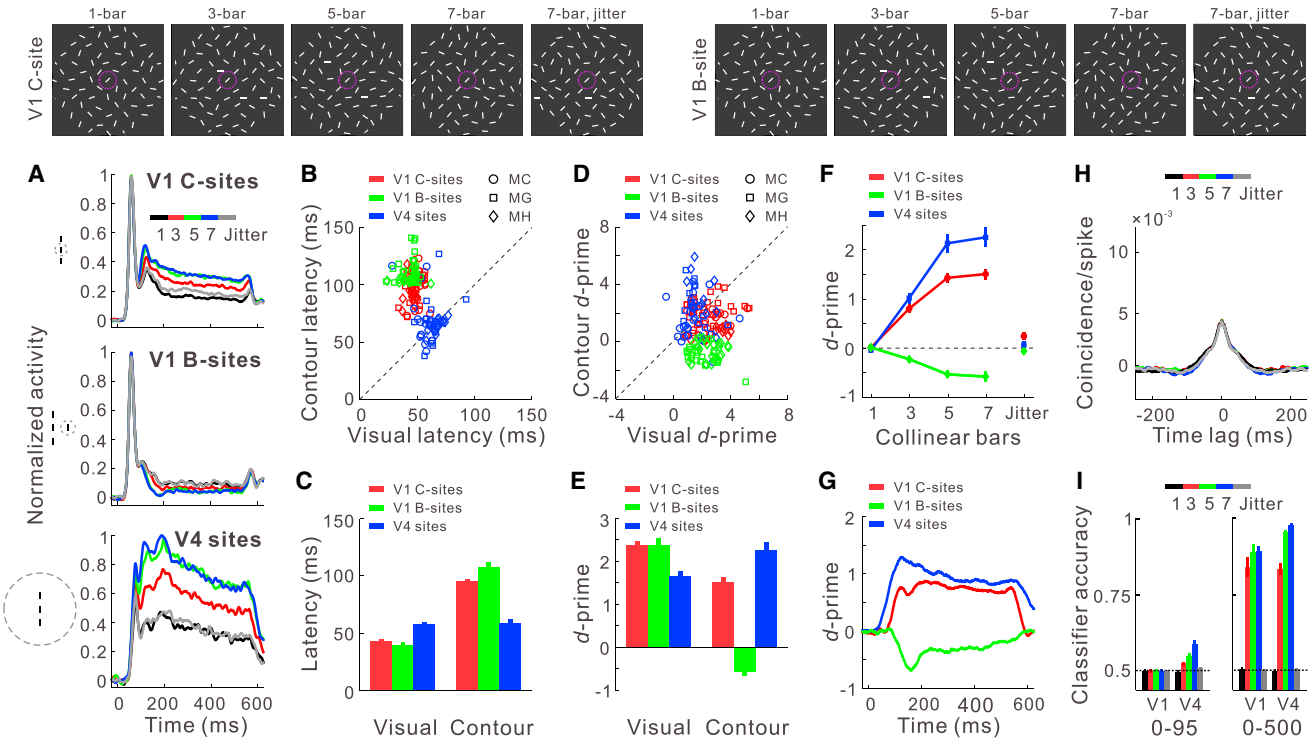

Figure 2. Contour-Induced Neuronal Responses in V1 and V4

(A) Normalized and averaged PSTHs constructed from responses to contours of different lengths (upper insets) for V1 recording sites on the contour (defined as V1 C-sites, n = 79; 23, 24, and 32 from MC, MG, and MH respectively), on the background (defined as V1 B-sites, n = 51; 22 and 29 from MG and MH), and for V4 sites (n = 67; 23, 15, and 29 from MC, MG, and MH). A jittered 7-bar contour was included for comparison. Time 0 indicates stimulus onset. Among all the recording sites, 37% and 63% were classified as single- and multi-unit (Figure S1B). Data from single- and multi-units in different animals were pooled, as they were qualitatively similar (Figure S2).

(B and C) Comparisons of the visual- and contour-response latencies for individual sites (B), and the mean latencies averaged across recording sites and animals (C). The contour responses were based on the 7-bar contour.

(D and E) Comparisons of the visual- and contour-response d' for individual sites (D), and the mean d' values averaged across recording sites and animals (E).

(F) Mean contour-response d' as a function of contour length. The rightmost isolated data points indicate the jittered 7-bar contour condition.

(G) Time course of the mean contour-response d' (7-bar contour) constructed by binning neuronal responses with 50 ms windows sliding in 1 ms steps.

(H) Cross-correlation analysis at different contour lengths. The crosscorrelograms were calculated from spike trains within −300 to 95 ms for pairs of V1 contour sites with preferred orientations close to the contour orientation (deviation <30°, n = 1,944 pairs pooled from the three animals). Results from paired V1 sites with different orientation and location relationships were very similar in each animal during this early response phase (Figure S3A). More detailed cross-correlation analyses are also shown in Figure S3B for V1 sites during the late response phase and in Figure S4 for paired V4-V4 and V1-V4 sites.

(I) Classification accuracy of a linear SVM classifier based on responses of all V1 or V4 sites within 0–95 or 0–500 ms. Error bars represent ± SEM. Different stimulus conditions were mixed within a block of trials, and each condition was repeated in 30 trials, except for the data used in cross-correlation analysis, where 75 trials were used in each condition. See also Figures S2–S4 and Table S1.

V4 neurons had a longer mean latency of 58 ms (Mann-Whitney U test on the latency of individual sites, $p < 10^{-10}$) (Figure 2B, blue dots; Figure 2C, left blue bar) and a smaller d' of 1.68 ($p < 10^{-9}$; Figure 2D, blue dots; Figure 2E, left blue bar). This simple comparison indicates an average delay of 14 ms required for feedforward conduction of signals from V1 to V4 and also suggests a general decrease of stimulus-driven activity along the visual pathway (Poort et al., 2012).

As the number of collinear bars increased from one to seven, V1 contour sites with RFs lying on the contour showed a progressive increase in activity (Figure 2A, top; Spearman's rank correlation between the contour lengths and mean neuronal responses within 0–500 ms, $\rho = 0.66$, $p < 10^{-10}$; referred to as contour facilitation). In contrast, V1 background sites with RFs lying on the background showed a progressive decrease in activity (Figure 2A, middle; $\rho = -0.38$, $p < 10^{-9}$; referred to as background suppression). Similar to V1 contour sites, V4 sites showed enhanced responses with increasing contour length (Figure 2A, bottom; $\rho = 0.65$, $p < 10^{-10}$). Neuronal responses to the 7-bar contour pattern were compared with those to the noise pattern to determine the significance of contour-related responses (Mann-Whitney U test at 5% alpha level). On average, contour facilitation was evident in 90% of V1 contour sites and in 85% of V4 sites; significant background suppression was observed in 54% of V1 background sites.

We later quantified the latencies and strengths of contour-related responses. The difference in peristimulus time histogram (PSTH) between the 7-bar contour pattern and the noise pattern

was used to construct a cumulated differential response curve, from which the latency of contour-related responses was measured (defined as the contour-response latency, to be differentiated from the visual-response latency defined earlier) (see Experimental Procedures; Figure S1E). Contour responses in V4 emerged immediately at the response outset (59 ms contour-response latency versus 58 ms visual-response latency, Wilcoxon signed-rank test, $p = 0.47$) (Figure 2C, two blue bars; for data from individual sites, see Figure 2B, blue dots around the diagonal). These contour-related responses in V4 significantly preceded V1 contour facilitation (Mann-Whitney U test, $p < 10^{-10}$; Figure 2C, right red bar, 95 ms V1 contour-response latency), which in turn preceded V1 background suppression ($p < 0.01$; Figure 2C, right green bar, 112 ms).

To quantify how well a recording site could differentiate the contour pattern from the noise pattern, the difference in distributions of mean neuronal firing rates (0–500 ms) in response to the contour and noise patterns was calculated as d' (referred to as the contour-response d', to be differentiated from the visual-response d' defined earlier). On average, the absolute value of contour-response d' increased with contour length for all types of recording sites (Figure 2F; Spearman's rank correlation all $\rho s > 0.87$, $Ps < 10^{-10}$). For the 7-bar contour pattern (the rightmost data points on the three curves shown in Figure 2F, which are replotted as the right group of bars in Figure 2E), the mean contour-response d' of V4 sites was 2.28 (Figure 2E, right blue bar), which was larger than that of V1 contour sites (1.54, Mann-Whitney U test, $p < 0.01$; Figure 2E, right red bar). The latter was larger than the absolute value of contour-response d' of V1 background sites (-0.58, $p < 10^{-7}$; negative value indicates background suppression; Figure 2E, right green bar).

Our data showed that from V1 to V4 there was an increase in latency and a decrease in d' for visually evoked responses, but the opposite was true for contour-induced responses (compare Figure 2C with 2E). These results suggest that contour grouping initially starts in V4 with a fast feedforward process, which is followed by a countercurrent process that engages V1 neurons on the contour for amplification of the signals as well as those on the background for suppression of the noise. After the initiation of contour responses in V1, the contour signals in both V1 and V4 continued to build up rapidly and reached a maximum (Figure 2G), implying that this incremental integration process involves bidirectional interactions between visual areas.

One may argue that the enhancement of V4 responses might simply result from spatial summation of iso-oriented optimal bars. To exclude this possibility, we laterally offset alternate collinear bars by 0.125° to disrupt the collinearity of a 7-bar contour (demonstration in Figure 2, top insets, the "jitter" condition). This small jitter eliminated the delayed contour signals in V1 as well as the early and late contour signals in V4 (Wilcoxon signed-rank test, all $Ps > 0.46$; the "jitter" condition in Figures 2A and 2F), indicating that continuity is critical for integration of contour segments in both V1 and V4. This observation also suggests that the underlying processes in V1 and V4 are interdependent.

One may also argue that the earlier emergence of contour signals in V4 (59 ms) than V1 (95 ms) might be a consequence of synchronized feedforward inputs from V1, although the initial responses of individual V1 neurons (within 44–95 ms after stimulus onset) did not carry any contour information. However, by crosscorrelation analysis of the spike trains of pairs of V1 sites in this early response phase, no significant difference was seen across different contour lengths, regardless of the orientation preference and the RF location of the V1 sites (Figure 2H, permutation test for the Euclidean distance between any two crosscorrelograms (CCGs), all $Ps > 0.82$). Therefore, the early contour signals in V4 were not due to precisely aligned spikes in time across V1 neurons.

Using a decoding method (support vector machine [SVM]), we further examined whether the simultaneously recorded V1 sites carried any global contour information within their population responses and thus could contribute to the early contour signals in V4. A linear SVM classifier was trained to predict whether the stimulus pattern was a contour or noise pattern based on neural activities in single trials. The SVM classifier took account of the information carried by individual recording sites and the correlation structure among them. The classification accuracy was at the chance level based on early (0–95 ms) V1 responses (Figure 2I, left panel; not significantly different from 0.5, Wilcoxon signed-rank test, all $Ps > 0.39$). In contrast, the classifier's performance was significantly above the chance level based on V4 responses during this early period (Figure 2I, right group of bars in left panel; all $Ps < 0.003$ for contour patterns with more than one collinear bars), and it reached above 0.85 based on V1 or V4 responses during the entire period of stimulus display (0–500 ms, Figure 2I, right panel); in addition, the classification accuracy was correlated with the contour length (Figure 2I, all but the leftmost group of bars; Spearman's rank correlation all $\rho s > 0.67$, $Ps < 10^{-10}$). Since early V1 responses of individual neurons and of the population as a whole could not be used to predict the embedded contours, this suggests that the initial contour-related V4 responses could not be inherited from V1, but the considerable further increase in V4 classification accuracy after the onset of V1 contour responses suggests a close interplay between these two cortical areas in amplification of global contour signals.

Spatiotemporal Properties of Contour-Related Responses in V4

The emergence of contour signals from V4 response onset and the absence of contour information in V1 around this period of time suggest initiation of contour integration in V4 via pooling feedforward inputs. If this speculation holds, the contour signals in V4 would reflect the basic response properties of V4 neurons. We tested this hypothesis by examining the orientation and location dependence of the contour-related responses in V4.

To examine the orientation dependence of the contour signals in a V4 site, we fitted its orientation tuning curve with a Gaussian function. We placed a 7-bar contour in the RF center and set the contour at the optimal orientation (Gaussian center), at the orientations 1.17 SD and 1.96 SD away from the Gaussian center (neuronal responses dropped by 50% and 85% respectively), or at the orthogonal orientation.

When the contour was rotated away from the preferred orientation, neuronal responses rapidly decreased (Figure 3A). The same was true of the contour-response d' (Figures 3B and 3C,

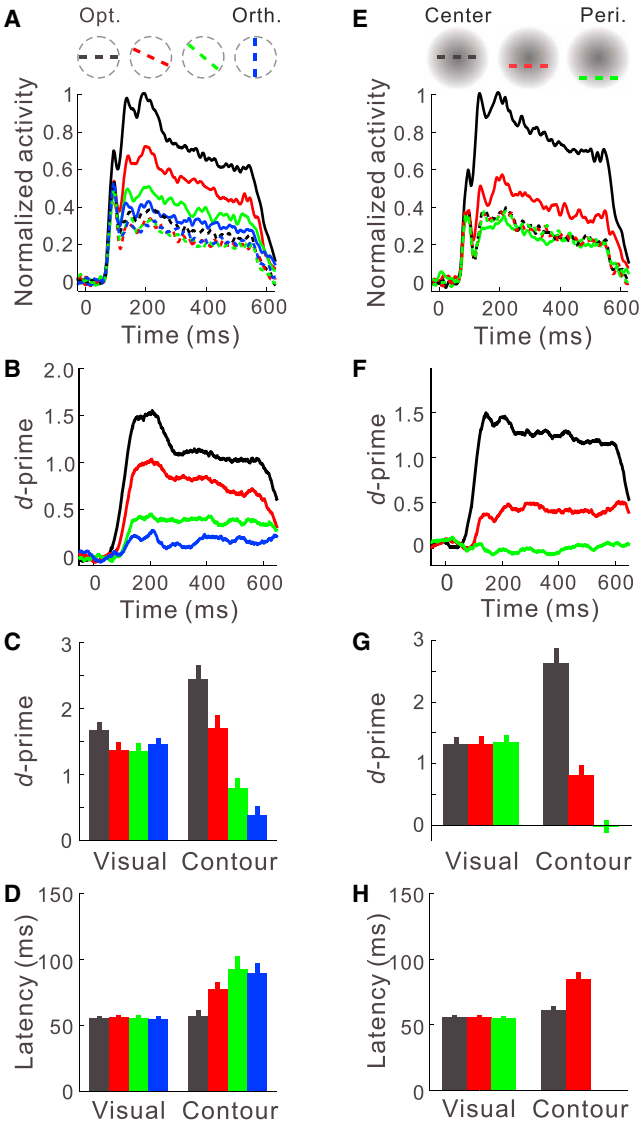

Figure 3. Comparisons of Contour Signals in V4 across Orientations and Locations

(A) Orientation dependence of V4 contour-related responses (n = 54; 20, 16, and 18 from MC, MG, and MH respectively, with single- and multi-units pooled). Population PSTHs were constructed from neuronal responses to contours centered in the RF at different orientations (different colors, see insets), which were chosen relative to the optimal orientation of the recorded site. Solid and dashed PSTHs correspond to the 7-bar contour pattern and the 1-bar contour pattern (i.e., noise pattern), respectively.

(B) Time course of contour-response d' calculated from the data shown in (A).

(C) Mean visual- and contour-response d' based on neuronal responses from 0–500 ms shown in (A).

(D) Mean visual- and contour-response latencies at four contour orientations.

(E–H) Similar to (A)–(D), but showing location dependence of contour-related responses (n = 59; 25, 13, and 21 from MC, MG, and MH). Note that the contour-response latency was not measurable in the RF periphery (absent in [H]) due to a contour-response d' of ~0 (see [E]–[G]). Error bars represent ±SEM.

right group of bars; the mean d' during 0–500 ms, Spearman's rank correlation $\rho = -0.53$, $p < 10^{-10}$). At the nonoptimal orientations, the contour-response latency significantly increased (Figure 3D, right group of bars; $\rho = 0.65$, $p < 10^{-10}$). On the other hand, the visual-response d' (Figure 3C, left group of bars) and the visual-response latencies (Figure 3D, left group of bars), which were measured based on neuronal responses to the noise pattern (equivalent to the 1-bar contour pattern), were independent of the orientation of the stimulus pattern (Friedman test for repeated-measures, all Ps > 0.15), indicating that the orientation dependence of contour-response d' and latency is not an artifact due to different spatial arrangements of the background elements when the entire stimulus pattern was rotated around its center to generate different contour orientations.

Not only were the contour signals in V4 dependent on neuron's orientation selectivity, but they were also contained within the classical RF, as revealed by the following experiments.

An optimally orientated 7-bar contour was placed at the RF center, or at the locations away from the RF center where neuronal excitability dropped by 50% and 85%. When the contour was shifted away from the RF center, we observed a marked decrease in neuronal responses (Figure 3E) and contour-response d' (Figure 3F; Figure 3G, right group of bars; the mean d', Spearman's rank correlation $\rho = -0.68$, $p < 10^{-10}$). In particular, when the global contour was placed at the location where RF excitability dropped by 85%, V4 neurons were unable to differentiate the contour and noise patterns (Figures 3E and 3F, green curve; Figure 3G, right green bar; the mean d' not significantly different from 0, Wilcoxon signed-rank test, $p = 0.83$). Shifting the contour to the peripheral position where RF excitability was reduced by 50% was also associated with a significant increase in contour-response latency (Figure 3H, two bars on the right; $p < 10^{-5}$). Unlike the contour-related responses, the visual-response d' (Figure 3G, left group of bars) and latencies (Figure 3H, left group of bars) in response to the noise pattern (i.e., 1-bar contour pattern with one optimally oriented bar embedded at any of the three locations in the RF) were indistinguishable (Friedman test, all Ps > 0.62).

We previously showed that the property of contour integration in V1 involves contextual influences from outside the classical RF (Li et al., 2006). Here we attempted to determine the contribution of extra-RF influences on contour integration in V4 by using very large stimulus patterns. Keeping the stimulus size constant (50° in diameter), we proportionally scaled the hidden grids dividing the stimulus and the bar inside each grid so that the same V4 RF enclosed different number of grids/bars (insets in Figures 4A–4D). At each stimulus scale, the number of collinear bars was gradually increased from one until the contour extended up to seven times the RF size. Since very large stimulus patterns were used in this experiment, some modifications of the stimulus display had to be made: the viewing distance was reduced from 100 to 32 cm; the fixation point was moved from the CRT center to one corner so that long contours could be generated in the screen center. Because the display could hold only one stimulus pattern, the animal was simply required to perform a fixation task.

At the largest stimulus scale when the RF held only one contour element and all the other elements were outside, V4 neuronal responses were independent of the number of collinear bars (Figure 4A), and the contour-response d' was close to 0 regardless of the contour length (Figure 4E, black curve;

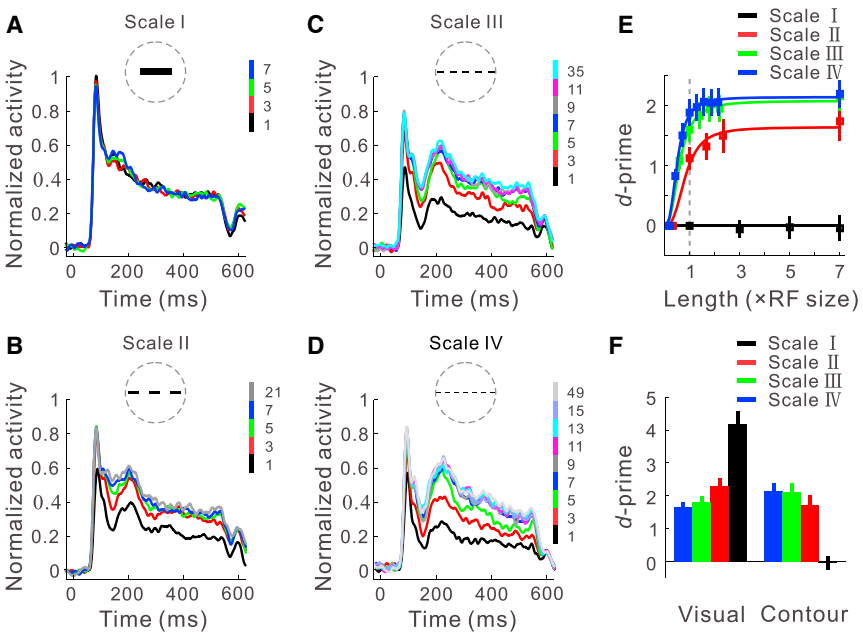

Figure 4. Contour Integration in V4 at Different Stimulus Scales

(A–D) Population PSTHs at four stimulus scales (see insets, Scale I–IV), which were chosen in such as a way that the RF could hold 1, 3, 5, or 7 collinear bars (n = 34; 14 and 20 from MG and MH). The size of bars composing the stimulus and the spacing between adjacent bars were proportionally scaled, while the stimulus size was kept unchanged. The contour length was increased until it was seven times the RF size. The color legends represent different number of collinear bars.

(E) The mean contour-response d' as a function of contour length at each of the stimulus scales shown in (A)–(D). The contour length (number of collinear bars) was normalized to RF size (vertical dashed line indicates 1 × RF size).

(F) The visual-response d' (left group of bars) calculated from neuronal responses to the noise pattern (i.e., 1-bar contour), and the contour-response d' (right group of bars) for the longest contour (7× RF size), at the four stimulus scales. Error bars represent ±SEM.

Friedman test, p = 0.83). At smaller scales, neuronal responses (Figure 4B-4D) and contour-response d' (Figure 4E) increased sharply as one increased the number of short collinear bars within the RF, but the enhancement became negligible after the collinear bars extended outside the central RF region, where neuronal excitability was less than 15% of that in the RF center. In particular, at the smallest stimulus scale when the RF contained 7 short collinear bars (Figure 4D), these within-RF contour elements contributed to ~90% of the maximal contour-response d' (1.91 versus 2.14, not significantly different, p = 0.48) (Figure 4E, blue curve). These results indicate that V4 neurons are unable to integrate contour elements outside their classical RFs.

A comparison between different stimulus scales also showed that the contour-response d' decreased with increasing stimulus scale (Figure 4F, right group of bars; same as the rightmost data points in Figure 4E; $\rho = -0.47$, $p < 10^{-9}$), suggesting that large RFs in V4 are not necessarily suited for integration of long contour segments. This result was not due to reduced responsiveness of V4 neurons to stimulus patterns composed of long, and thus sparse, bars, because the visual-response d' actually increased with the stimulus scale (Figure 4F, left group of bars; Spearman's rank correlation $\rho = 0.43$, $p < 10^{-7}$).

Spatiotemporal Properties of Contour-Related Responses in V1

Contour signals in V1 lagged behind those in V4 by several tens of milliseconds. We explored the properties of these delayed V1 response components, which turned out to be distinct from V4 contour responses in terms of orientation and location dependencies.

A 7-bar contour was set at one of four orientations ranging from the preferred to the orthogonal orientation of a V1 site (insets in Figures 5A and 5B). These four orientations were chosen from the Gaussian fit of the orientation tuning curve: the Gaussian center (optimal), the orientations at 1.17 SD and 1.96 SD away from the Gaussian center, and the orientation orthogonal to the optimal.

Simply by looking at the response magnitudes of V1 contour sites (Figure 5A), we observed that neuronal responses decreased as the contour was rotated away from the preferred orientation, seemingly akin to V4 (compare Figures 5A and 3A). However, by examining the contour-response d', a measure of neuronal ability to discriminate the contour pattern from the noise pattern, we found almost constant d' across all contour orientations (the upper group of curves in Figure 5C; see also Figure 5D, dark solid line, left y axis, Friedman test, p = 0.40). This result is in striking contrast with that in V4 (compare Figure 5C, upper curves, with Figure 3B and Figure 5D, dark solid curve, with Figure 3C, right group of bars). The contour-response latencies in V1 were also constant across contour orientations (p = 0.38, Figure 5D dark dashed line, right y axis), unlike V4 (compared with Figure 3D right groups of bars). One can notice that the collinear contour centered on V1 RF activated strong responses even at the orthogonal orientation (Figure 5A). This seemingly poor orientation selectivity was most likely due to the presence of a background of randomly oriented bars. In fact, when tested with gratings, the V1 sites showed a typical distribution of orientation selectivity (Figure S1C).

Contour-induced background suppression in V1 was also largely independent of the orientation of the contour placed outside the RF (Figure 5B). The contour-response d' (Figure 5C, lower group of curves; Figure 5D, gray solid line at the bottom, left y axis) and the contour-response latency (Figure 5D, gray dashed line, right y axis) were similar across different contour orientations (Friedman test, p = 0.68 and 0.36, respectively, for the mean d' and latency, though the background suppression seems a little stronger at the optimal orientation) (Figure 5C).

The above results indicate that the global contour induced a general facilitation of neurons encoding the contour and a

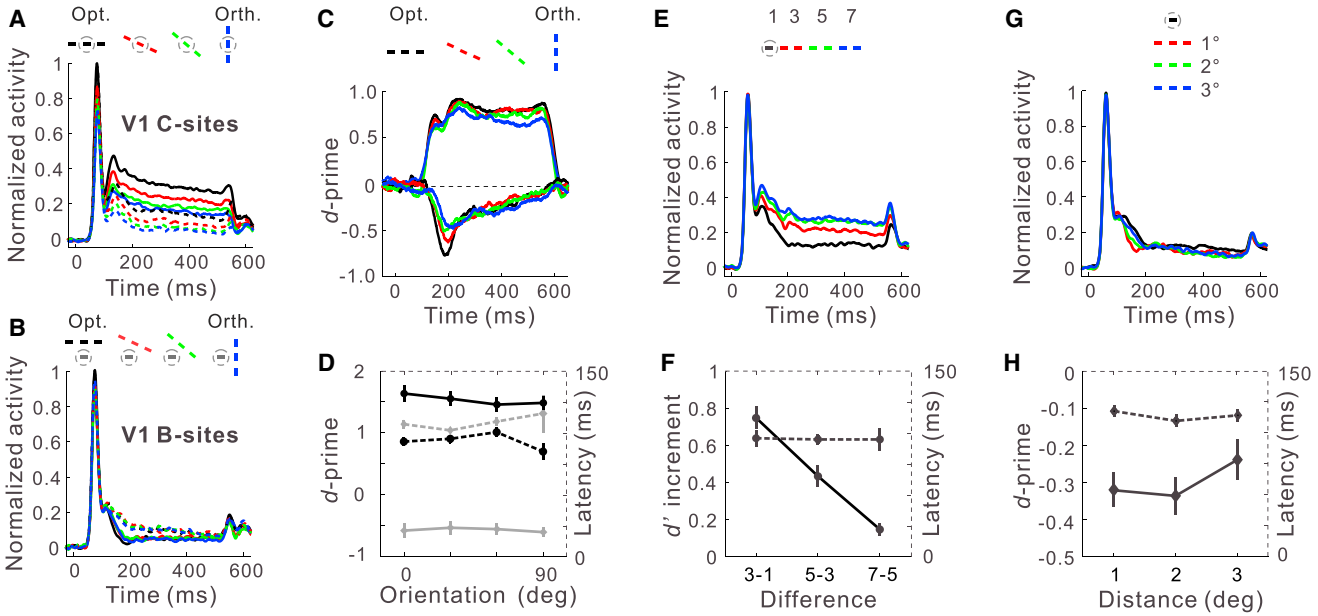

Figure 5. Comparisons of Contour Signals in V1 across Orientations and Locations

(A and B) Population PSTHs of V1 contour sites ([A], n = 76; 20, 24, and 32 from MC, MG, and MH) and background sites ([B], n = 59; 30 and 29 from MG and MH) at different contour orientations (different colors, see insets). To test a contour site, the contour was centered on the RF (insets in [A]); to test a background site, the contour was embedded 1° away from the RF center, and an optimally oriented background bar was always placed in the RF center (insets in [B]). Solid and dashed PSTHs correspond to the 7- and 1-bar contour patterns, respectively.

(C and D) Orientation independence of contour detectability in V1.

(C) Time course of contour-response *d′* for V1 contour sites (upward curves) and background sites (downward curves).

(D) Mean contour-response *d′* (left y axis, solid lines) and latencies (right y axis, dashed lines) for V1 contour sites (dark lines) and background sites (gray lines). These results were derived from the data shown in (A) and (B).

(E and F) Comparisons of contour signals between two adjacent contour lengths when the number of collinear bars was increased from 1 to 3, 3 to 5, and 5 to 7 on one side of the RF (n = 54; 30 and 24 from MG and MH).

(E) Population PSTHs at different contour lengths (1, 3, 5, and 7 bars).

(F) Solid line (left y axis) shows the mean *d′* between neuronal responses to two adjacent contour lengths (3 versus 1, 5 versus 3, and 7 versus 5 number of collinear bars); dashed line (right y axis) shows the time point where the two PSTHs diverged for two adjacent contour lengths. Note that the spacing between contour elements was 0.5°; therefore, for a 7-bar contour, the outermost collinear bar was 3.0° away from the RF center.

(G and H) Contour-induced background suppression at different contour-to-RF distances (n = 62; 30 and 32 from MG and MH).

(G) Population PSTHs when a 7-bar contour was placed at 1° (red), 2° (green), and 3° (blue) from the RF center. The PSTH corresponding to the noise pattern (black) is also shown for comparison.

(H) Mean contour-response *d′* (solid line, left y axis) and latencies (dashed line, right y axis) of background suppression. Error bars represent ±SEM.

general inhibition of neurons responding to the background, independent of orientation tuning of neurons.

The constant delays of V1 contour facilitation and background suppression across all contour orientations were observed when the distance was fixed between the RF center and the global contour center (insets in Figures 5A and 5B). To examine whether this distance could affect the timing of contour facilitation, the contour element at one end of the contour was placed inside the classical RF, and the collinear bars were appended to the other end (Figure 5E inset). The addition of collinear elements exerted an additional facilitatory effect relative to the previous contour length. Although this effect decreased with distance from the RF, it still was seen a couple of degrees from the RF boundary (Figure 5F, solid line, left y axis; Wilcoxon signed-rank test, the *d′* between two adjacent contour lengths significantly larger than 0, all *P*s $< 10^{-5}$). The delay of contour-related response, however, did not change with the distance (Figure 5F, dashed line, right y axis; Friedman test, p = 0.89). The contour-induced inhibition of V1 background sites also showed a similar effect: the contour placed at different distances away from the RF (Figure 5G inset) inhibited the neuronal responses with indistinguishable latencies (p = 0.48; Figure 5H, dashed line, right y axis), although the suppression strength depended on the contour-to-RF distance (Figure 5H, solid line, left y axis; Spearman's rank correlation $\rho = 0.19$, $p < 0.01$).

As an additional test of the distance of contour-related interactions in V1, we increased the size of grids dividing the stimulus pattern, and thus the center-to-center distance between adjacent bars from 0.5° to 2.5°, without changing the size of component bars (Figure 6A). A 2.5° distance in the visual field corresponded to ~6.9 mm distance in V1 based on the cortical magnification factor (2.74 mm/°, calculated from the distance between electrodes and the recorded RF locations in the visual field) (Figure 1B). At this separation, both contour facilitation (Figure 6B, left) and background suppression (Figure 6B, right) were still dependent on the number of collinear bars that were spaced

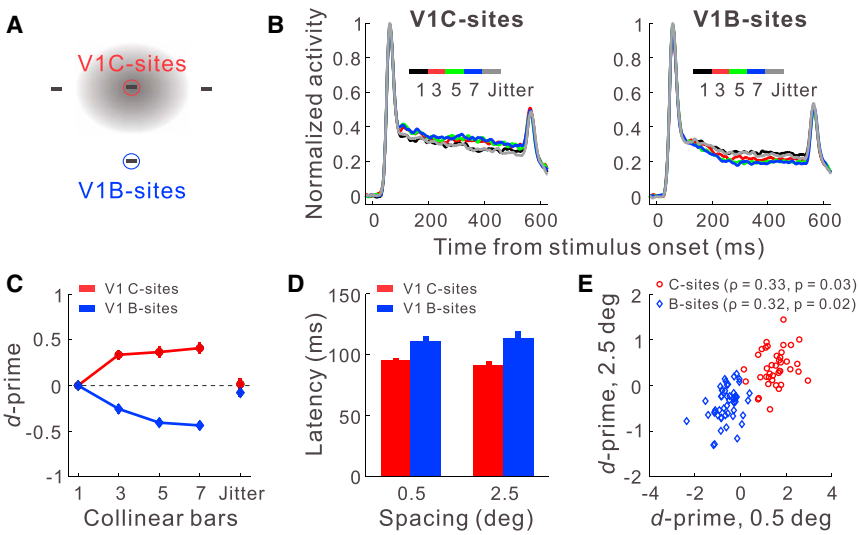

Figure 6. Long-Range Contour Integration in V1

(A) Illustration of a large spacing (2.5°) between neighboring bars. The gray area represents the putative horizontal projections of a V1 neuron; the small circle in the center represents its classical RF. When examining a background site, the contour was placed 2.5° laterally from its RF.

(B) Population PSTHs of V1 contour sites (left, n = 50; 25 and 25 from MG and MH) and background sites (right, n = 66; 37and 29 from MG and MH) at different contour lengths with 2.5° interelement spacing. The jittered 7-bar contour condition is included for comparison.

(C) Same data in (B) were replotted to show the mean contour-response *d′*.

(D) Comparison of the mean contour-response latencies at two interelement spacings.

(E) Correlation between contour-response *d′* at two different interelement spacings for individual V1 sites. Error bars represent ±SEM.

far apart. The magnitude of contour-response *d′* increased with contour length for both contour sites (Figure 6C, red curve; Spearman's rank correlation $\rho = 0.52$, $p < 10^{-10}$) and background sites (Figure 6C, blue curve; $\rho = -0.53$, $p < 10^{-10}$). A disruption of the collinearity by a lateral jitter (0.625°, scaled to five times the jitter used for 0.5° bar spacing in Figure 2F) abolished the contour-related signals (Figure 6C, right-most isolated data points; not significantly different from 0, Wilcoxon signed-rank test, all *P*s > 0.69). Moreover, the contour-response latency was still similar to that observed at the bar spacing five times smaller (0.5°) for both the contour facilitation (91 ± 4.03 ms versus 95 ± 1.51 ms, Mann-Whitney U test, p = 0.38; Figure 6D, two red bars) and background suppression (114 ± 4.51 ms versus 112 ± 4.67 ms, p = 0.76; Figure 6D, two blue bars). Furthermore, the strengths of contour signals at these two spacings were significantly correlated for individual V1 sites in both contour facilitation (Pearson correlation $r = 0.33$, $p < 0.03$; Figure 6E, red) and background suppression ($r = 0.32$, $p < 0.02$; Figure 6E, blue). The similar response properties of V1 neurons at both small and large contour spacings suggest that the same neural mechanism is used for integration of contour segments across a large range of contour spacings and within a large visual-field area.

Coherence Analyses within and across V1 and V4

The asynchronous onset of contour-related signals in V1 contour sites, V1 background sites, and V4 sites suggests complex interactions within the cortical network. To examine the interactions among V1 and V4 neurons, we performed spike-spike coherence analyses, which gave a measure of similarity between neuronal outputs in each frequency component.

The data used for the coherence analysis, as well as for cross-correlation analysis, were collected in an experiment with several modifications made to avoid sampling bias and to ensure reliability. First, the embedded contour took one of four fixed orientations (vertical, horizontal, 45°, and 135°), and at each orientation a fixed set of five contour positions (0.5° apart) were tested to cover the cluster of V1 RFs. Second, the orientations and positions of individual bars in the complex background were fixed across trials rather than rerandomized to reduce variation of neuronal responses. Third, the number of repetitions of each stimulus condition was increased from 30 to 75.

For a given contour orientation and location, the simultaneously recorded V1 and V4 sites were classified using the criteria described in Figure 7A. In brief, both V1 and V4 sites were first classified as the parallel sites, whose preferred orientations were close to the contour orientation, and as the orthogonal sites. The selected V1 sites were further separated into contour and background sites. To analyze interareal interactions, pairs of V1and V4 sites were divided into two categories: pairs with overlapping and nonoverlapping RFs. This classification allowed us to analyze intra-areal and interareal interactions among V1 and V4 neurons with similar or different tuning properties and spatial relationships.

The coherence value was a function of the frequency of neuronal spike trains (e.g., Figure 7B). The difference in mean coherence values (averaged across 0–100 Hz) between the contour- and noise-pattern conditions was used to measure the change in coherence strength induced by the global contour.

We first measured the coherence between V1 sites using spike trains from 0–500 ms. Compared with the noise-pattern condition, we observed a significant reduction in coherence between paired V1 sites when a contour passed through their RFs (Wilcoxon signed-rank test, all *P*s $< 10^{-10}$; Figure 7C, left group of bars), in spite of increased neuronal responses by the contour. Such a reduction in coherence was generally observed across pairs of V1 contour sites, regardless of their orientation preferences, implying increased independence between responses of V1 neurons with RFs lying on the contour.

Similar to paired V1 contour sites, the coherence between a contour site and a background site was also weakened by the global contour (all *P*s $< 10^{-7}$; Figure 7C, middle group of bars). However, unlike paired contour sites or paired contour and background sites, the coherence between paired V1

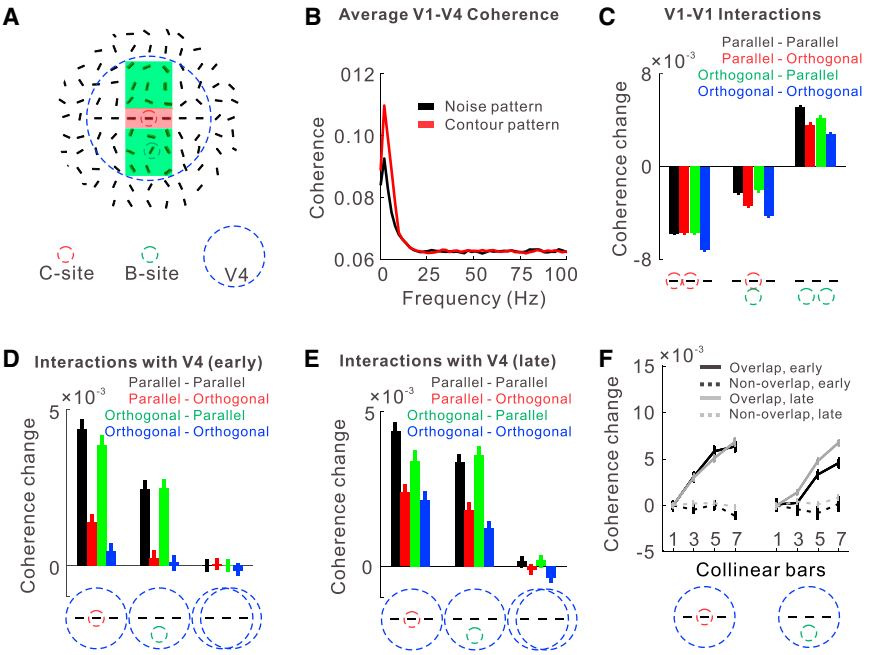

Figure 7. Spike-Spike Coherence between Recording Sites

(A) Classification of recording sites. The simultaneously recorded V1 and V4 sites were first classified based on the deviation of their optimal orientations from the contour orientation—the parallel sites (deviation <30°) and the orthogonal sites (deviation > 60°); the remaining sites were discarded. For initially selected V1 sites, only those with RF centers located within a ±0.75° band perpendicularly intersecting the contour center (green and pink areas) were included in the analysis; these sites were separated into the contour sites (pink area, RF-contour distance ≤0.25°) and the background sites (green area, RF-contour distance between 0.25° and 1.75°). For initially selected V4 sites, only those with central RF regions (±1.17 SD of the Gaussian envelope) intersecting the axis of the global contour were used. To analyze V1-V4 interactions, paired V1 and V4 sites were classified into two categories: pairs with overlapping RFs, in which the V1 RF center was located within the central region (±1.17 SD) of the V4 RF, and pairs with nonoverlapping RFs, in which the V1 RF center was at least 2.58 SD away from the V4 RF center.

(B) An example showing the spike-spike coherence as a function of frequency. In this example, the coherence between V1 and V4 sites with overlapping RFs was calculated for the contour pattern (with different contour lengths pooled) and noise pattern, respectively. The presence of the global contour remarkably enhanced the V1-V4 coherence at low frequencies (<20 Hz).

(C) Contour-induced changes in mean coherence (0–100 Hz) between V1 contour-contour sites (left), contour-background sites (middle), and background-background sites (right). Four different colors indicate different relationships of preferred orientations of paired sites with respect to the contour orientation.

(D and E) Contour-induced changes in mean coherence between V1 contour sites and V4 sites (left), V1 background sites and V4 sites (middle), and V4-V4 sites (right) during the early (−300 to 95 ms, [D]) and late (95–500 ms, [E]) response periods. For examining V1-V4 interactions, only paired sites with overlapping RFs were used.

(F) Comparison of contour-induced coherence change between V1 and V4 sites with overlapping and nonoverlapping RFs at different contour lengths. Left: coherence between V1 contour sites and V4 sites. Right: coherence between V1 background sites and V4 sites. The mean coherence value in the 1-bar contour condition was subtracted. Overlapping V1 and V4 RFs (insets) were found only in MG and MH, and nonoverlapping RFs were found only in MC (Figure 1B). Error bars represent ±SEM. The numbers of different types of paired sites in each animal are listed in Table S1. In Supplemental Information, we provide the coherence analyses and power spectra of V1 and V4 spikes trains in individual animals (Figure S5). We also conducted Granger causality analysis on paired sites with various location and orientation relationships (Figure S6); the results were generally consistent with coherence analysis except that Granger causality may provide possible directional information about the neuronal interactions. See also Figures S5 and S6 and Table S1.

background sites was enhanced (all *P*s < 10^{-10}; Figure 7C, right group of bars), which could be useful in strengthening local inhibition to exclude the background noise.

Neurons in V4 were most sensitive to contours at their preferred orientations from response outset, suggesting that V4 neurons can selectively integrate feedforward inputs. This can be further tested by examining the coherence between V1 and V4 sites with overlapping RFs using neuronal responses from −300 to 95 ms, which excluded the contribution of V1 contour signals that emerged only after this period of time. V4 sites with preferred orientation close to the contour orientation (i.e., the parallel sites) showed a significant increase in coherence with both V1 contour and background sites (all *P*s < 10^{-10}; Figure 7D, left and middle black and green bars). However, distinct from V4 parallel sites, V4 orthogonal sites with preferred orientation deviated far from the contour orientation showed much smaller contour-induced changes in coherence with any classes of V1 sites (Figure 7D, left and middle red and blue bars; all *P*s < 10^{-10}). These results suggest that, although V1 neurons cannot differentiate the contour and noise patterns during this early response period, V4 neurons are able to integrate contour segments by pooling feedforward inputs.

After the emergence of contour-related signals in late V1 response components (95–500 ms), the enhanced coherence between V4 parallel sites and V1 contour or background sites was still present (all *P*s < 10^{-10}; Figure 7E, left and middle black and green bars), suggesting interplay between cortical areas for contour integration during this delayed response period.

Further evidence for such interareal interplay comes from a comparison with nonoverlapping V1 and V4 RFs. We calculated the mean coherence values at different contour lengths, with the baseline coherence in the 1-bar contour condition subtracted. With increasing contour length, a monotonic increase in V1-V4 coherence was observed in both early and late response phases for paired sites with overlapping RFs (Figure 7F, solid curves; Spearman's rank correlation, all ρs > 0.46, p < 10^{-10}). For V1 and V4 sites with nonoverlapping RFs, however, their coherence was independent of contour length (Figure 7F, dashed lines; Friedman test, all *P*s > 0.31). This comparison suggests that interareal interactions at the same retinotopic locations were

important for generating and amplifying contour signals in both V1 and V4.

DISCUSSION

Feedforward and Feedback Processes in Contour Grouping

The results presented here point toward a model of contour integration involving intercortical interactions, with enhancement in responses to contours resulting from a countercurrent stream of processing between areas V1 and V4. The onset of contour responses in V1 is delayed relative to that seen in V4, though both continue to evolve in parallel after that time. Still unresolved is whether the enhancement in V1 is due to the feedback input alone or whether it involves an interaction between feedback and intrinsic connections within V1. We have suggested previously that feedback may serve to gate horizontal connections within V1 (Gilbert and Li, 2013; Gilbert and Sigman, 2007; McManus et al., 2011; Piëch et al., 2013).

It is generally accepted that the feedforward, feedback, and horizontal connections targeting a visual cortical neuron contribute collectively to its rich response properties (Gilbert and Li, 2013; Lamme et al., 1998). While the feedforward inputs endow the neuron with selectivity for simple stimulus attributes, the lateral and feedback connections can dynamically modify its response properties according to stimulus context and behavioral goal, and confer selectivity for more complex stimulus geometries. This contrasts with the traditional feedforward model that posits V1 as a preprocessor, representing local stimulus attributes (Hubel and Wiesel, 1959), and higher areas such as V4 as form processor, representing complex shapes (Brincat and Connor, 2006; Schiller, 1995).

The interaction among cortical neurons during contour integration is likely to be a complex one, in part because of the dissociation between the contour-related properties of V1 and V4 neurons: those in V1 show facilitation for global contours of all orientations, whereas those in V4 show facilitation only for optimally oriented contours; V1 shows a push-pull relationship between foreground and background with profound inhibition of neurons with RFs on the background pattern, whereas V4 only shows a decrement in responses to contours that are shifted away from the RF center. Different response properties of V1 and V4 neurons suggest their respective contributions to contour detection. Because of the large RF size, the contour-related responses in V4 are likely to provide only a crude signal that approximates the position and shape of the contour but that is nevertheless sufficient to provide the feedback signal that enhances contour-related responses in V1 and suppresses irrelevant background inputs. The push-pull response mode in V1 may in turn allow V4 to analyze contour shape unhindered by distracters in the visual image; in addition, the contour signals in V1 may confer higher spatial resolution for scene segmentation by providing more precise information concerning contour position, alignment of contour elements, the contour/background boundary, and details of contour shape.

The above processing scheme is consistent with a general theoretic framework (Epshtein et al., 2008; Hochstein and Ahissar, 2002; Jehee et al., 2007; Roelfsema, 2006; Ullman, 1984) that is efficient in analyzing and disambiguating complex visual scenes based on bottom-up and top-down recurrent processing. A modeling study has shown that a cortex-like hierarchical network is able to reliably and rapidly recognize objects and their parts almost simultaneously by a single feedforward sweep followed by a feedback sweep (Epshtein et al., 2008). Some models have shown how the top-down interactions may operate by gating local circuits within V1, accounting for both contour enhancement and background suppression (Li, 1998; Piëch et al., 2013).

Comparisons with Previous Studies

The source of contextual signals in V1 has been hotly debated. On the one hand, many orientation-dependent contextual effects, including contour integration governed by the Gestalt law of continuity, are thought to be mediated by V1 horizontal connections, because these connections tend to link neurons with nonoverlapping RFs but similar orientation preferences (Bosking et al., 1997; Gilbert and Wiesel, 1979; Rockland and Lund, 1982; Stettler et al., 2002). On the other hand, some researchers attribute the long-range contextual influences, such as surround inhibition (Angelucci et al., 2002; Bair et al., 2003), surface segmentation (Poort et al., 2012), and border ownership (Sugihara et al., 2011; Zhang and von der Heydt, 2010), to feedback projections from higher cortical areas.

One argument for feedback origin of V1 long-range contextual modulation is based on an additional delay required for some contextual influences to take place in V1 relative to the outset of visual responses. However, this argument has been questioned given the evidence that there is a large overlap of response latencies among cortical areas and that within interconnected cortical loops there are complex recurrent neuronal interactions across many nodes (for a review, see Lamme and Roelfsema, 2000). Another argument for feedback origin of V1 contextual effects is based on constant delays of some modulatory effects independent of the distance of contextual stimuli from the classical RF (Bair et al., 2003; Zhang and von der Heydt, 2010). We also observed a constant delay for contour facilitation (or background suppression) regardless of the location of V1 RF on the contour (or on the background). The constant delay implies that contour-induced push-pull modulatory effects do not require a propagation of contour signals from one location to another within V1. While one may speculate that this constant and significant delay could be due to a backward propagation of contour signals from V4 to V1, it is important to keep in mind that a shift of V1 horizontal network from one stable state to another could also take a significant amount of time that is independent of contour length, as suggested by a modeling study (Piëch et al., 2013).

Our finding of orientation independence of contour-induced modulatory effects in V1 is somewhat surprising, but this can be explained by a careful review of the properties of V1 horizontal connections (Bosking et al., 1997; Gilbert and Wiesel, 1979; Rockland and Lund, 1982; Stettler et al., 2002). At the longest range, the horizontal connections link cortical columns of similar orientation, but the lateral connections tend to be nonspecific as to orientation locally (Stettler et al., 2002). Additionally, the horizontal connections have a roughly equal extent in all directions

visuotopically. This allows facilitation along collinear contours, which is then conveyed, by local connections, to neurons of all orientations. At this point, one cannot differentiate this scenario from facilitation provided by feedback, since the feedback to V1 is nonspecific as to orientation (Stettler et al., 2002).

A number of studies have shown that synchronous neuronal spikes are unrelated to the processes of contour integration (Roelfsema et al., 2004), surface segmentation (Lamme and Spekreijse, 1998), and other feature binding (Dong et al., 2008; Palanca and DeAngelis, 2005; Thiele and Stoner, 2003). Similarly, we did not see consistent changes in neuronal synchrony that was correlated with the contour length within V1 (Figures 2H and S3) or V4 (Figures S4A and S4B) or between them (Figures S4C and S4D); instead, a close correlation between neuronal firing rates and contour saliency was observed, suggesting that the firing rate is a more reliable code for contour integration than precise spike synchrony.

Previous studies have shown that, similar to V1, response properties of V4 neurons are subject to modulations by stimulus context placed outside their classical RFs (e.g., Desimone and Schein, 1987). Nevertheless, the current study showed that the contour responses in V1 involved contextual influences outside the classical RF, whereas those in V4 were contained within the RF and could be mediated by integration of feedforward inputs. The existence of horizontal connections, in all cortical areas, however, suggests the existence of contextual influences in V4 that may operate in stimulus domains other than contour integration.

Incremental Integration as a General Mechanism of Perceptual Organization

The coarse-to-fine incremental integration of visual scene components within the hierarchical cortical network, which involves two countercurrent streams of processing and features a reversed global-to-local process, seems to be a general mechanism in grouping and segmenting of visual scenes. Similar delayed response components carrying global figural information have been observed in V1 of awake monkeys for contour integration (Bauer and Heinze, 2002; Li et al., 2006, 2008), surface segmentation (Poort et al., 2012; Roelfsema and Spekreijse, 2001; Zipser et al., 1996) and perceptual filling-in of subjective contours (Lee and Nguyen, 2001). These delayed V1 responses are closely correlated with the animal's performance on perceptual tasks. Moreover, interference by transcranial magnetic stimulation of human V1 during this delayed period, even well after higher cortical areas are activated, significantly impairs subjects' ability in categorization of natural scene images (Koivisto et al., 2011). These studies suggest that feedback modulation from higher to lower visual areas plays a critical role in conscious perception of global forms.

All the earlier findings above using complex visual stimuli could be put under the framework of incremental integration of image components mediated by countercurrent streams of processing: the feedforward processing captures the gist of visual scenes in higher-order visual cortex; the weak figural signal is enhanced, and the background noise is reduced by feedback modulation of neuronal responses in early visual cortex; the countercurrent processing streams lead to parallel augmentation of global figural information in multiple cortical areas with different levels of detail.

EXPERIMENTAL PROCEDURES

Animal Preparation

Three adult male monkeys (*Macaca Mulatta*, 6.5–10.5 kg) were used. Ethical approval was granted by the Institutional Animal Care and Use Committee of Beijing Normal University, with all procedures in compliance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals.

Following standard surgical procedures (see Supplemental Experimental Procedures), the animals were first implanted with a titanium head-post for restraining head movements. After extensive training in the fixation and contour detection tasks (similar to Li et al., 2006), microelectrode arrays were implanted in V1 and V4 using a pneumatic inserter (Blackrock Microsystems).

Contour Stimulus and Behavioral Task

The stimulus patterns were generated by a visual stimulator (ViSaGe MKII, Cambridge Research Systems) on a gamma-corrected CRT monitor (Vision Master Pro-514, Iiyama; 1,200 × 900 pixels at 100 Hz). The stimulus, viewed at a distance of 100 cm, comprised of antialiased white (12.4 cd/m^2) bars on a uniform gray (4.1 cd/m^2) background.

Each stimulus display in a trial consisted of a contour pattern and a noise pattern, which were presented symmetrically around the fixation point (Figure 1A). An infrared tracking system (K. Matsuda et al., 2000, Soc. Neurosci., conference) was used to sample eye positions at 30 Hz. The distributions of recorded eye positions were not significantly different across experimental conditions. A trial began when the animal fixated within a circular window of 0.5° in radius around the fixation point. After the animal kept its fixation for 300 ms, the stimulus was presented for 500 ms followed by 300 ms blank interval (Figure 1A). Then two 0.2° targets appeared. The animal was required to make a saccade within 800 ms to the target corresponding to the contour pattern location in exchange for a reward. Trials were aborted if the animal's gaze was out of the fixation window before the saccade targets were displayed.

Electrophysiological Recordings

Neuronal spikes were recorded using microelectrode arrays (Multiport) and data acquisition system (Cerebus) made by Blackrock Microsystems. The array in V1 or V4 contained 6 × 8 electrodes (~0.5 mm in length spaced 0.4 mm apart). Spikes were detected by applying a voltage threshold with a signal-to-noise ratio of 3.5, and their waveforms were sampled at 30 kHz. The recorded spikes from each electrode were first sorted into clusters by self-customized software. Spikes in the cluster with the largest signal-to-noise ratio were taken as the spiking activity of the recorded site, which was further classified as a single- or multi-unit using a parametric index (Figure S1B; see Supplemental Experimental Procedures).

The RFs of all the V1 and V4 recording sites were determined using a narrow band (0.3° wide) of square-wave gratings (1.5 cycles/°, drifting at 2 cycles/s, 90% Michelson contrast with mean luminance of 21.2 cd/m^2) (see Supplemental Experimental Procedures for more details). The mean firing rates recorded by each electrode at different horizontal or vertical positions of the grating stimuli were fitted with a Gaussian function. The RF center was measured as the Gaussian center; the RF width was defined as 2 × 1.96 SD of the Gaussian with the width of the grating stimulus (0.3°) subtracted. After mapping the RF, we used a large, circular grating patch (7° in diameter) to determine the orientation tuning curve of each V1 or V4 site and to measure the optimal orientation and tuning width by Gaussian fitting. The goodness of fit was estimated using R^2; only recording sites with $R^2 > 0.8$ in both the RF profiles and the orientation tuning curve were examined.

To test the contour integration ability within V4 RF (Figure 3E), more detailed RF structures were also mapped by a small, circular grating patch (0.42° in radius with similar grating parameters) that was randomly presented in grids spaced 0.6° apart. The visual field area within which the neuronal activity was above 15% of the maximum (with spontaneous activity subtracted) was defined as the V4 RF. Its center was defined as the center of mass; its size was defined as the square root of its area.

Data Analysis

Only recording sites showing significant responses to any of the contour patterns were included in the analysis; the statistical significance was determined at the level of 5% using the Wilcoxon signed-rank test by comparing the mean stimulus-evoked activity (0–500 ms) with the mean spontaneous activity (−300 to 0 ms). Each site was included only once in each analysis, except for cross-correlation and spike-spike coherence that required a large number of paired sites. To quantify the relationship between mean neuronal responses (0–500 ms) and the contour lengths, Spearman's rank correlation was computed. The statistical significance of the correlation coefficient was determined based on its Student's t approximation (Siegel, 1956).

We computed the sensitivity index d' from signal detection theory to quantify stimulus-evoked (visual-response d') and contour-induced (contour-response d') neuronal firing rates. The Wilcoxon signed-rank test (for paired data) and Mann-Whitney U test (for unpaired data) were used to assess the difference in d' between two data sets.

The PSTH for each recording site under a stimulus condition was created by binning the spike times in 1 ms intervals and averaging the binned spike counts across trials. The raw PSTH was smoothed by convolving it with a Gaussian window (SD = 5 ms). The PSTHs of each recording site in all the stimulus conditions to be compared were normalized to the largest peak of them, with spontaneous activity subtracted. The normalized PSTHs were averaged across recording sites to derive the population PSTHs.

We adapted a two-phase regression method to measure neuronal response latencies by constructing cumulative response curves along time (Sugihara et al., 2011) (see Figures S1D and S1E; Supplemental Experimental Procedures). The latency estimates of individual recording sites could be unreliable due to their noisy PSTHs; therefore, to calculate the mean latencies of a class of recording sites, population PSTHs were used; the SEM was determined by bootstrapping, which randomly resampled the recording sites 1,000 times.

The CCG, which quantifies the ratio of coincident spikes between two recording sites, was constructed using 1 ms bins (Smith and Kohn, 2008). To remove stimulus-locked coincidences, a shift predictor was constructed by shuffling the trial number. The shuffle-corrected CCG was regarded as the strength of interneuronal synchrony.

Neuronal interactions were also examined using spike-spike coherence, a measure of similarity in each frequency component of spike trains between neurons. The power spectra of spike trains were estimated using the multitaper technique. Coherence between two signals was defined as the modulus of their cross-spectrum normalized by the geometric mean of autospectra. The coherence values over frequencies (0–100 Hz, Figure 7B) were averaged. The mean coherence in response to the noise pattern was subtracted from that to the contour pattern to estimate the contour-induced coherence change (Figures 7C–7F).

A linear classifier, known as the SVM, was used to decode the population responses (see Supplemental Experimental Procedures). The SVM classification accuracy was computed in a leave-one-out cross-validation procedure.

SUPPLEMENTAL INFORMATION

Supplemental Information includes six figures, one table, and Supplemental Experimental Procedures and can be found with this article online at http://dx.doi.org/10.1016/j.neuron.2014.03.023.

AUTHOR CONTRIBUTIONS

M.C., Y.Y., and W.L. designed and conducted the experiments. M.C., X.G., and H.L. analyzed the data. M.C., C.D.G., and W.L. wrote the paper.

ACKNOWLEDGMENTS

We thank Dajun Xing, Zheng Li, Si Wu, En Zhang, Min Huang, and Shiming Tang for valuable comments and also thank Xibin Xu, Feng Wang, Rujia Chen, and Bin Yang for technical assistance. This work was supported by National Key Basic Research Program of China (2014CB846101, 2011CBA00400), National Natural Science Foundation of China (31125014, 30970983), the 111 Project (B07008), and the Fundamental Research Funds for the Central Universities. The authors declare no competing conflicts of interest.

Accepted: March 3, 2014
Published: May 7, 2014

REFERENCES

Angelucci, A., Levitt, J.B., Walton, E.J.S., Hupe, J.-M., Bullier, J., and Lund, J.S. (2002). Circuits for local and global signal integration in primary visual cortex. J. Neurosci. *22*, 8633–8646.

Bair, W., Cavanaugh, J.R., and Movshon, J.A. (2003). Time course and time-distance relationships for surround suppression in macaque V1 neurons. J. Neurosci. *23*, 7690–7701.

Bauer, R., and Heinze, S. (2002). Contour integration in striate cortex. Classic cell responses or cooperative selection? Exp. Brain Res. *147*, 145–152.

Bosking, W.H., Zhang, Y., Schofield, B., and Fitzpatrick, D. (1997). Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. J. Neurosci. *17*, 2112–2127.

Brincat, S.L., and Connor, C.E. (2006). Dynamic shape synthesis in posterior inferotemporal cortex. Neuron *49*, 17–24.

Bringuier, V., Chavane, F., Glaeser, L., and Frégnac, Y. (1999). Horizontal propagation of visual activity in the synaptic integration field of area 17 neurons. Science *283*, 695–699.

Buffalo, E.A., Fries, P., Landman, R., Liang, H., and Desimone, R. (2010). A backward progression of attentional effects in the ventral stream. Proc. Natl. Acad. Sci. USA *107*, 361–365.

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.

Desimone, R., Albright, T.D., Gross, C.G., and Bruce, C. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. J. Neurosci. *4*, 2051–2062.

Dong, Y., Mihalas, S., Qiu, F., von der Heydt, R., and Niebur, E. (2008). Synchrony and the binding problem in macaque visual cortex. J. Vis. *8*, 1–16.

Epshtein, B., Lifshitz, I., and Ullman, S. (2008). Image interpretation by a single bottom-up top-down cycle. Proc. Natl. Acad. Sci. USA *105*, 14298–14303.

Gilad, A., Meirovithz, E., and Slovin, H. (2013). Population responses to contour integration: early encoding of discrete elements and late perceptual grouping. Neuron *78*, 389–402.

Gilbert, C.D., and Li, W. (2013). Top-down influences on visual processing. Nat. Rev. Neurosci. *14*, 350–363.

Gilbert, C.D., and Sigman, M. (2007). Brain states: top-down influences in sensory processing. Neuron *54*, 677–696.

Gilbert, C.D., and Wiesel, T.N. (1979). Morphology and intracortical projections of functionally characterised neurones in the cat visual cortex. Nature *280*, 120–125.

Hochstein, S., and Ahissar, M. (2002). View from the top: hierarchies and reverse hierarchies in the visual system. Neuron *36*, 791–804.

Hubel, D.H., and Wiesel, T.N. (1959). Receptive fields of single neurones in the cat's striate cortex. J. Physiol. *148*, 574–591.

Jehee, J.F.M., Roelfsema, P.R., Deco, G., Murre, J.M.J., and Lamme, V.A.F. (2007). Interactions between higher and lower visual areas improve shape selectivity of higher level neurons-explaining crowding phenomena. Brain Res. *1157*, 167–176.

Kapadia, M.K., Ito, M., Gilbert, C.D., and 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.

Koivisto, M., Railo, H., Revonsuo, A., Vanni, S., and Salminen-Vaparanta, N. (2011). Recurrent processing in V1/V2 contributes to categorization of natural scenes. J. Neurosci. *31*, 2488–2492.

Lamme, V.A., and Roelfsema, P.R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. Trends Neurosci. *23*, 571–579.

Lamme, V.A.F., and Spekreijse, H. (1998). Neuronal synchrony does not represent texture segregation. Nature *396*, 362–366.

Lamme, V.A., Supèr, H., and Spekreijse, H. (1998). Feedforward, horizontal, and feedback processing in the visual cortex. Curr. Opin. Neurobiol. *8*, 529–535.

Lee, T.S., and Nguyen, M. (2001). Dynamics of subjective contour formation in the early visual cortex. Proc. Natl. Acad. Sci. USA *98*, 1907–1911.

Li, Z. (1998). A neural model of contour integration in the primary visual cortex. Neural Comput. *10*, 903–940.

Li, W., Piëch, V., and Gilbert, C.D. (2006). Contour saliency in primary visual cortex. Neuron *50*, 951–962.

Li, W., Piëch, V., and Gilbert, C.D. (2008). Learning to link visual contours. Neuron *57*, 442–451.

McManus, J.N.J., Li, W., and Gilbert, C.D. (2011). Adaptive shape processing in primary visual cortex. Proc. Natl. Acad. Sci. USA *108*, 9739–9746.

Palanca, B.J.A., and DeAngelis, G.C. (2005). Does neuronal synchrony underlie visual feature grouping? Neuron *46*, 333–346.

Pasupathy, A., and Connor, C.E. (1999). Responses to contour features in macaque area V4. J. Neurophysiol. *82*, 2490–2502.

Piëch, V., Li, W., Reeke, G.N., and Gilbert, C.D. (2013). Network model of top-down influences on local gain and contextual interactions in visual cortex. Proc. Natl. Acad. Sci. USA *110*, E4108–E4117.

Poort, J., Raudies, F., Wannig, A., Lamme, V.A., Neumann, H., and Roelfsema, P.R. (2012). The role of attention in figure-ground segregation in areas V1 and V4 of the visual cortex. Neuron *75*, 143–156.

Rockland, K.S., and Lund, J.S. (1982). Widespread periodic intrinsic connections in the tree shrew visual cortex. Science *215*, 1532–1534.

Roelfsema, P.R. (2006). Cortical algorithms for perceptual grouping. Annu. Rev. Neurosci. *29*, 203–227.

Roelfsema, P.R., and Spekreijse, H. (2001). The representation of erroneously perceived stimuli in the primary visual cortex. Neuron *31*, 853–863.

Roelfsema, P.R., Lamme, V.A., and Spekreijse, H. (2004). Synchrony and covariation of firing rates in the primary visual cortex during contour grouping. Nat. Neurosci. *7*, 982–991.

Schiller, P.H. (1995). Effect of lesions in visual cortical area V4 on the recognition of transformed objects. Nature *376*, 342–344.

Siegel, S. (1956). Nonparametric Statistics for the Behavioral Sciences. (New York: McGraw-Hill).

Smith, M.A., and Kohn, A. (2008). Spatial and temporal scales of neuronal correlation in primary visual cortex. J. Neurosci. *28*, 12591–12603.

Stettler, D.D., Das, A., Bennett, J., and Gilbert, C.D. (2002). Lateral connectivity and contextual interactions in macaque primary visual cortex. Neuron *36*, 739–750.

Sugihara, T., Qiu, F.T., and von der Heydt, R. (2011). The speed of context integration in the visual cortex. J. Neurophysiol. *106*, 374–385.

Thiele, A., and Stoner, G. (2003). Neuronal synchrony does not correlate with motion coherence in cortical area MT. Nature *421*, 366–370.

Ullman, S. (1984). Visual routines. Cognition *18*, 97–159.

Ursino, M., and La Cara, G.E. (2004). A model of contextual interactions and contour detection in primary visual cortex. Neural Netw. *17*, 719–735.

Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt II. Psychol. Forsch. *4*, 301–350.

Zhang, N.R., and von der Heydt, R. (2010). Analysis of the context integration mechanisms underlying figure-ground organization in the visual cortex. J. Neurosci. *30*, 6482–6496.

Zipser, K., Lamme, V.A.F., and Schiller, P.H. (1996). Contextual modulation in primary visual cortex. J. Neurosci. *16*, 7376–7389.