Neuron
Article

The Role of Attention in Figure-Ground Segregation in Areas V1 and V4 of the Visual Cortex

Jasper Poort,[1] Florian Raudies,[2] Aurel Wannig,[1] Victor A.F. Lamme,[3] Heiko Neumann,[4] and Pieter R. Roelfsema[1,5,*]
[1]Netherlands Institute for Neuroscience, an institute of the Royal Netherlands Academy of Arts and Sciences, Meibergdreef 47, 1105 BA Amsterdam, The Netherlands
[2]Department of Cognitive and Neural Systems, Boston University, Boston, MA 02115, USA
[3]Cognitive Neuroscience Group, Department of Psychology, University of Amsterdam, 1018 WB Amsterdam, The Netherlands
[4]Institute of Neural Information Processing, University of Ulm, Ulm 89069, Germany
[5]Department of Integrative Neurophysiology, Centre for Neurogenomics and Cognitive Research, VU University Amsterdam, De Boelelaan 1085, 1081 HV Amsterdam, The Netherlands
*Correspondence: p.roelfsema@nin.knaw.nl
http://dx.doi.org/10.1016/j.neuron.2012.04.032

SUMMARY

Our visual system segments images into objects and background. Figure-ground segregation relies on the detection of feature discontinuities that signal boundaries between the figures and the background and on a complementary region-filling process that groups together image regions with similar features. The neuronal mechanisms for these processes are not well understood and it is unknown how they depend on visual attention. We measured neuronal activity in V1 and V4 in a task where monkeys either made an eye movement to texture-defined figures or ignored them. V1 activity predicted the timing and the direction of the saccade if the figures were task relevant. We found that boundary detection is an early process that depends little on attention, whereas region filling occurs later and is facilitated by visual attention, which acts in an object-based manner. Our findings are explained by a model with local, bottom-up computations for boundary detection and feedback processing for region filling.

INTRODUCTION

Visual perception starts in early visual areas with the detection of elementary features like the orientation and color of image elements by neurons with small receptive fields. This piecemeal analysis is very different from our subjective perception. We perceive objects composed of many features that activate large, distributed neuronal populations in visual cortex. Our visual system reconstructs objects from these distributed representations by grouping the image elements of objects and by segregating them from the background. A neural correlate of this reconstruction process is observed in the primary visual cortex (area V1), where neurons enhance their response when their receptive field (RF) is on a figure compared to when it is on the background, an effect known as figure-ground modulation (FGM) (Lamme, 1995; Marcus and Van Essen, 2002; Zipser et al., 1996). FGM labels image elements of a figure with enhanced activity so that they are grouped in perception (Roelfsema, 2006; Roelfsema and Houtkamp, 2011). Our understanding of the neural mechanisms for FGM is limited. It is unknown if this signal depends on interactions within V1 or whether it reflects an interaction between V1 and higher visual areas. Furthermore, it is unclear if the labeling process occurs for all figures, or only for those that are relevant for behavior. Finally, the functional role of these contextual influences in V1 is not well understood. How is the pattern of FGM reflected in behavior?

We wished to elucidate the neuronal interactions that give rise to FGM. Previous neurocomputational models have proposed two complementary mechanisms for the segregation of a figure from the background (Mumford et al., 1987). The first ''boundary-detection'' mechanism detects abrupt changes in features at locations where figures and background abut and the second ''region-filling'' mechanism joins similar image elements into larger figural regions (Ullman, 1984). These two processes give rise to apparently conflicting constraints on the neuronal connectivity (Grossberg and Mingolla, 1985; Roelfsema et al., 2002). On the one hand, algorithms for boundary detection use inhibition between neurons with nearby RFs tuned to the same feature (Grossberg and Mingolla, 1985; Itti and Koch, 2001; Li, 1999). Neurons with RFs in the middle of a homogenous region receive strong inhibition from their neighbors so that their activity is weak, whereas neurons with RFs at boundaries receive less inhibition so that their activity is stronger (Figure 1A, left). Algorithms for region filling, on the other hand, require that neurons tuned to similar features excite each other. If the representation of some of the figural image elements is enhanced, the excitatory connections spread the enhanced activity to neurons with a similar feature preference, coding elements of the same figure. A number of previous studies supported separate mechanisms for FGM at the figure boundary (edge modulation) and figure center (center modulation) (Huang and Paradiso, 2008; Lamme et al., 1998a, 1999; Scholte et al., 2008), but other studies disputed the existence of the region-filling process within V1 (Rossi et al., 2001; Zhaoping, 2003).

Another unresolved but possibly related issue is the role of task-driven attention in figure-ground segregation. The Gestalt

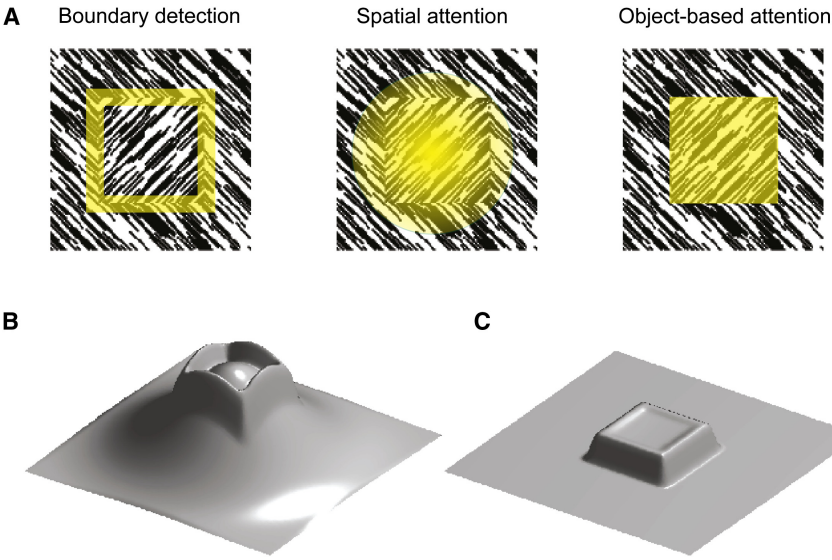

Figure 1. Possible Interactions between Figure-Ground Segregation and Attention

(A) An early, local process detects boundaries (left). Attention could act as a spotlight (middle) or be object based (right).

(B) Figure-ground modulation might add to the attentional effect in case of an attentional spotlight.

(C) Figure-ground modulation predicted by object-based attention.

psychologists (Koffka, 1935; Rubin, 1915; Wertheimer, 1923) delineated several bottom-up factors for figure-ground organization. They found that small, convex, and symmetric image regions are usually perceived as figures whereas large, concave and asymmetric regions are often perceived as background (Kanizsa and Gerbino, 1976; Koffka, 1935). But there is also an important influence of top-down factors (Peterson et al., 1991). For example, if you attend to a region of an ambiguous figure-ground display, this increases the probability that you perceive it as figure (Driver and Baylis, 1996; Vecera et al., 2004). It is not known how these bottom-up and top-down factors interact with each other (Driver et al., 2001; Qiu et al., 2007; Scholl, 2001). Does top-down attention act as a spotlight (Posner et al., 1980) and increase neuronal activity at the approximate location of the figure or does it act in object-based manner (Duncan, 1984) to specifically highlight image elements of the figure, in accordance with a region-filling process (Figure 1A)? It is also not well understood how attention interacts with the boundary-detection process. Attention might enhance neuronal activity in an additive manner (Figure 1B) or selectively boost the representation of figure's interior (Figure 1C).

To address these questions, we investigated neuronal activity in V1 in a texture-segregation task and also recorded simultaneously activity in V4, a higher area that is a source of feedback to V1 and is important for figure-ground segregation (Allen et al., 2009; De Weerd et al., 1994; Merigan, 1996). To determine the role of attention (Desimone and Duncan, 1995; Reynolds and Chelazzi, 2004; Treue, 2001), we required the monkeys to either attend the figures or pay attention elsewhere.

We report that attention acts in an object-based manner to enhance FGM in V1 and V4. Attention is important for the V1 center modulation, but less so for V1 edge modulation, indicating that edge detection is largely preattentive, while region filling primarily occurs for figures that are relevant. We reproduced our results with a hierarchical neural network model that uses a local inhibitory intra-areal process for boundary detection and excitatory feedback from higher areas for region filling. These results therefore indicate how the cortex resolves the apparently conflicting computational constraints. In addition, the experiment reveals a functional role of FGM in area V1 in eye movement planning (Moore, 1999; Supèr et al., 2004). The monkeys had to make precise eye movement to the figure center and the spatial profile of FGM predicted their saccadic endpoint, while the timing of V1 FGM predicted the onset of the saccade. These results imply that attention refines the representations of relevant objects in early visual areas, which makes them more useful for the guidance of behavior.

RESULTS

Behavioral Task and Accuracy

We trained three monkeys to perform two tasks with identical visual stimulation (see Figure 2 and Experimental Procedures). The animals first directed their gaze to a fixation point. After 300 ms of fixation, we presented a textured background (consisting of either 45° or 135° oriented line elements) and a 4° square figure with elements of the orthogonal orientation (Figures 2A and 2B). In the hemifield opposite to the figure, we presented two white curves on top of the textured background. On alternating days, the monkeys performed different tasks: on figure-detection days they made a saccade to the center of the figure, while on curve-tracing days the figure was irrelevant and they made a saccade to a red circle at the end of the curve connected to the fixation point (Figure 2C). Their average accuracy was 98% correct in the figure-detection task and 94% in the curve-tracing task.

Figure-Ground Modulation in Area V1

We recorded multiunit spiking activity with chronically implanted electrode arrays (Figure S1 available online). Figure 3 shows the activity of neurons at an example V1 recording site. We placed the figure at one of 23 positions (spaced 0.5° apart) so that the RF sometimes fell on the figure center (blue in Figure 3A), on the edge of the figure (red) or on the background (black). Because of the many conditions we averaged neuronal activity across seven sessions in both tasks. Figure 3B shows the neuronal activity in the figure-detection task. It can be seen that the responses evoked by the figure-center and edge were stronger than the response evoked by the background (t test on responses 200–600 ms, both $Ps < 0.05$). Figure 3D shows

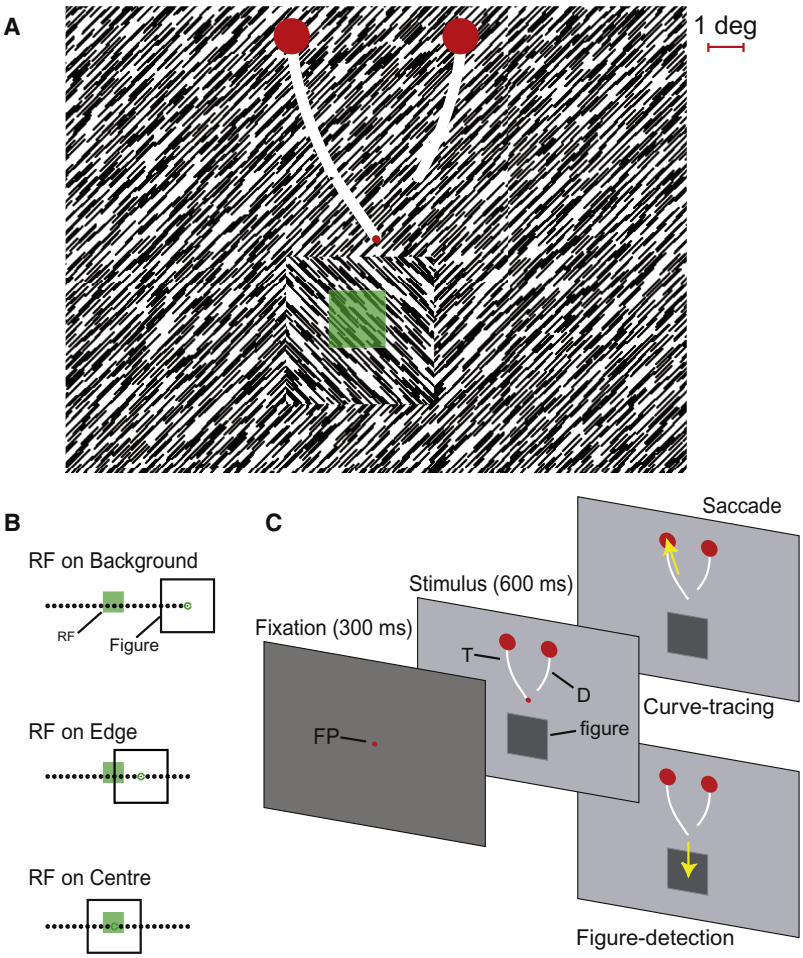

Figure 2. The Behavioral Paradigm

(A) The monkeys saw an orientation-defined square figure superimposed on a background with an orthogonal orientation, and two curves connected to larger red circles. The smaller red circle in the center is the fixation point. Green square, example RF of a V1 recording site. Length of the oriented line elements was increased in this figure to enhance visibility.

(B) The square Figure (4° × 4°) could appear at one of 23 locations. Black dots represent possible locations of the figure center. In some of the conditions the RF (green square) fell on the background (upper), on the edge (middle) or on the figure center (bottom).

(C) Figure-detection and curve-tracing task. The stimulus appeared after 300 ms of fixation and after an additional 600 ms the fixation point (FP) disappeared and the monkey had to make a saccade. The figure-detection task required a saccade to the center of the square figure. The curve-tracing task required a saccade to the circle that was connected to the fixation point by one of the curves (target, T) while the monkey could ignore the other curve (distracter, D).

See also Figure S1.

the FGM, which was computed by subtracting the average response to the background from the single-condition responses, as a function of time and figure position. The FGM at the edges started early and the FGM at the figure-center occurred later. Figures 3C and 3E illustrates the activity at the same recording site when the monkey did not attend the figure because he carried out the curve-tracing task. The FGM at the edge positions was also observed (t test, $p < 0.05$), but FGM was now virtually absent for the center positions (t test, $p > 0.05$).

These effects of attention on edge and center FGM were reproduced across a total of 59 V1 recording sites in three monkeys. In the figure-detection task, the response to the figure center and edge were enhanced relative to the background by 65% and 76%, respectively (in a window from 200–600 ms, Figure 4A). In the curve-tracing task, the edge modulation was also strong (52% increase in the response); however, the center response fell in between the response to the edge and the response to the background (29% increase, Figure 4B). These effects were present until the time of the saccade (right panels of Figures 4A and 4B).

Figures 4C and 4D show the space-time profile of FGM for attended and nonattended figures (bottom panels show responses aligned to stimulus onset, top panels responses aligned to saccade onset). Edge modulation started early, consistent with previous results (Lamme et al., 1999; Nothdurft et al., 2000) and was followed by a gradual filling in of the figure center, but this filling-in process was only partial for unattended figures. When aligning the responses to the saccade, it becomes clear that FGM in the figure detection task ramps up until the saccade is made, at which stage all elements of the figure are labeled with an enhanced response.

To investigate the reliability of these effects, we performed a repeated-measures ANOVA with factors RF position (center or edge) and task (figure detection or curve tracing), on the FGM across recording sites in successive 50 ms time windows (Figure 4E). From 75 ms after stimulus presentation onward, edge modulation was stronger than center modulation (main effect of RF position, dark gray area; $F_{1,58} > 9.5$, $p < 0.05$; with Bonferroni correction) that was maintained until the monkey's response. From 225 ms onward, there was also a main effect of task and a significant interaction (both *Ps* < 0.05) between RF position and task (light gray region in Figure 4E), because the center modulation depended more on attention than edge modulation. This interaction persisted until the onset of the saccade (Figure 4E, right panel) and the effect of attention on FGM was largest just before the eye movement was made (Supplemental Information, Figure S2E).

Next, we analyzed how well neurons at individual recording sites distinguished between figure and ground on single trials by computing *d*-primes (from 200 to 600 ms, see Experimental Procedures). The average *d*-prime of the center modulation was 0.32 if the figure was ignored and it increased by 68% to a value of 0.53 if it was attended (Figure 4F, paired t test $p < 10^{-6}$). We also observed a significant albeit weaker effect

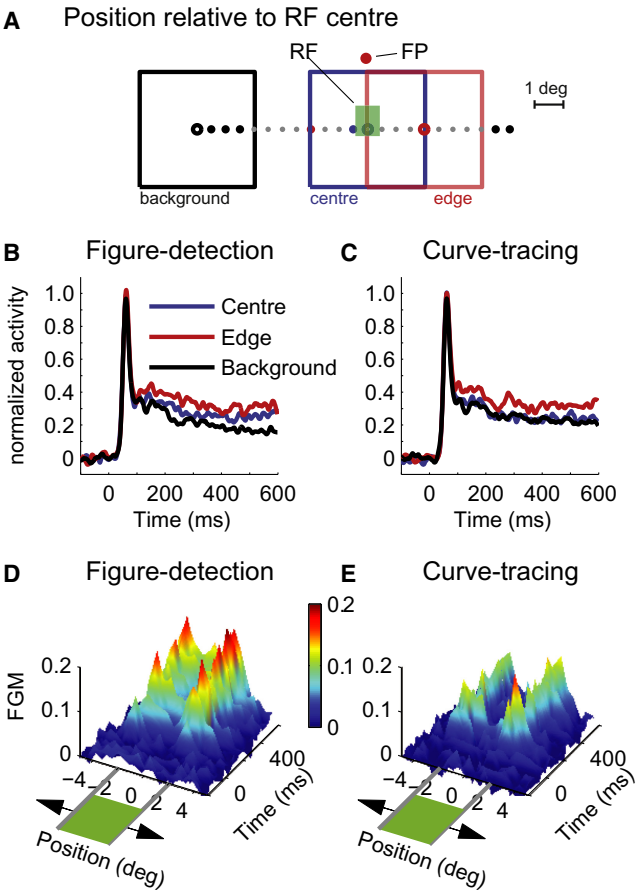

Figure 3. Effect of Attention on FGM of an Example V1 Recording Site

(A) The dots represent the possible locations of the center of the figure relative to the RF (green rectangle). The black, red, and blue dots represent conditions with the RF on background, edge, and figure center, respectively. Outlines drawn in corresponding colors represent example figure positions. FP, fixation point.

(B and C) Average neuronal response with the RF on the figure center (blue), edge (red), and background (black) in the figure-detection (B) and curve-tracing task (C).

(D and E) Space-time profile of FGM in the figure-detection (D) and curve-tracing task (E). The x axis shows time relative to stimulus onset, and the y axis figure position. If y equals 0, the RF center falls on the figure-center, and at y = −2, 2 the RF falls on the edge.

of attention on the *d*-prime of edge-FGM that increased from 0.53 to a value of 0.61 (15% increase, Figure 4G, paired t test $p < 10^{-6}$).

Figure-Ground Modulation in Area V4

Our results show that top-down attention increases FGM in V1. We also investigated neuronal activity in V4, which is one of the sources of feedback to V1. In these experiments we used the same figure size as in V1 because we wanted to compare the two areas in the same task. However, RFs in V4 were much larger than those in V1 and it was often impossible to confine the V4 RF to the interior of the figure. In V4 we therefore did not distinguish between activity evoked by the figure center and edge but only compared conditions with the RF on the figure and on the background.

Figure 5A illustrates how the figures were positioned relative to the RF of an example V4 recording site (for RF mapping, see Supplemental Information). In both the figure-detection and curve-tracing task, the figure evoked stronger responses than the background (average across five sessions per task), but FGM was strongest if the monkey attended the figure (t test, $p < 0.05$, Figures 5B and 5C). The space-time representation revealed that FGM occurred for all stimuli with the RF on the figure (Figures 5D and 5E). FGM was particularly strong and early if one of the edges fell on the center of the RF, whereas the effect of attention was expressed during a later phase of the response.

We obtained similar results across a population of 46 V4 recordings sites in the three monkeys (Figure 6). In both tasks, the figure evoked stronger activity than the background in V4, and this FGM was apparent at a relatively early phase of the response. Figures 6C and 6D shows the space-time profile of the FGM. FGM in the two tasks was similar in the first phase of the response, but it became stronger in the figure-detection task from 175 ms onward (Figure 6E, $p < 0.05$, paired t tests in successive 50 ms windows with Bonferroni correction), which is 50 ms earlier than the attention effect in V1. An analysis across individual V4 recording sites showed that attention increased the FGM *d*-prime by 64% from an average of 0.68 in the curve-tracing task to 1.12 in the figure-detection task (paired t test $p < 10^{-6}$, Figure 6F). As in area V1, FGM as well as the effect of attention in V4 also increased just before the saccade (Figure S2F).

We carried out a number of control analyses which showed that the response modulations in V1 and V4 could not be explained by differences in eye position between conditions (Supplemental Information, Figure S3).

Comparison of Neuronal Activity in Areas V1 and V4

To quantify the strength of visually driven activity we used the signal-to-noise ratio of the response transient evoked by the appearance of the texture (peak response divided by SD of spontaneous activity, see Experimental Procedures). The average signal-to-noise ratio in V1 was 9.3, which was more than twice as strong as the value of 4.1 in V4 (4.1, t test $p < 10^{-6}$). In contrast, FGM was stronger in V4 than in V1. In the figure-detection task, the FGM *d*-prime was 1.12 in V4, significantly larger than the center modulation (*d*-prime = 0.53) and edge modulation in V1 (*d*-prime = 0.61) ($p < 0.0001$ for both comparisons). In the curve-tracing task, the average FGM *d*-prime in V4 was 0.68, which was significantly larger than the *d*-prime for the V1 center modulation of 0.32 ($p < 0.0001$) but only marginally larger than the edge modulation in V1 (0.53; $p = 0.06$).

As a measure for the attentional effect, we computed the increase in the FGM *d*-prime if attention was directed to the figure compared to when it was not (a few recording sites with weak FGM in the curve-tracing task were excluded: one case in V4, two cases for V1 center, and one for V1 edge modulation). Attention increased FGM in V4 by 100%, on average, which was similar to the increase of 130% of the center modulation in V1

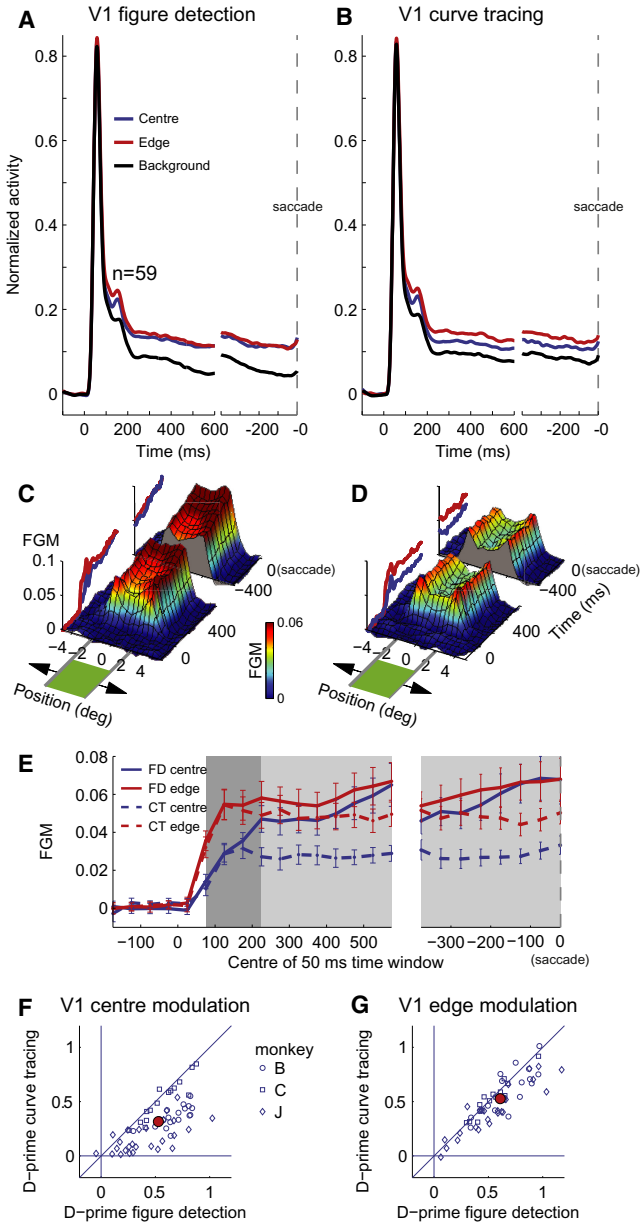

Figure 4. Population Analysis in Area V1

(A and B) The response evoked by image elements of the figure-center, edge, and background in the figure-detection (A) and curve-tracing task (B) averaged across 59 V1 recording sites, aligned to stimulus onset (left panel) and saccade onset (right panel).

(C and D) Space-time profile of the FGM (difference between figure and background response) in the figure-detection (C) and curve-tracing task (D). x axis, time relative to stimulus onset and saccade. y axis, figure position relative to the center of the RF. Red and blue curves on the left show FGM for center (blue) and edge positions (red) (enlarged in Figures S2A and S2B).

(E) FGM at the figure-center (blue lines) and edge (red lines) in the figure-detection task (FD, solid line) and curve-tracing task (CT, dashed line) in successive 50 ms windows aligned on stimulus onset (left) and saccade (right). Values on the x axis refer to the center of the time window and error bars show ± 2 SEM. Dark gray area, time points with significant main effect of figure position (repeated-measures ANOVA, $p < 0.05$, with Bonferroni correction). Light gray area, time points with a main effect of task ($p < 0.05$) and also a significant interaction between figure position and task.

(F and G) FGM *d*-prime in the figure-detection (x axis) and curve-tracing task (y axis) when the RF fell on the figure-center (F) or edge (G) (time window, 200–600 ms after stimulus onset). The larger red data points show means. See also Figures S2, S3, and S4.

(t test, $p > 0.4$). Attention increased edge modulation by only 19%, which was weaker than the effect on V1 center modulation (paired t test, $p < 0.001$) and V4 FGM (t test $p < 0.05$). Thus, attention had a strong influence on FGM in V4 and also on the V1 representation of the figure center, but a comparatively weak effect on the V1 edge representation.

The monkeys were proficient in the curve-tracing task with an average accuracy of 94%. We therefore considered the possibility that larger attentional effects on FGM occur with a more demanding curve-tracing task that removes more resources from the texture-defined figure. We therefore performed an additional experiment with the same texture stimuli while we varied the difficulty of the curve tracing-task (Figure S4). However, we found that the magnitude of the attentional effects did not depend strongly on task difficulty (Figure S4).

Functional Role of FGM

The figure-detection task demanded precise saccades because the eye had to land in a 2.5° window in the center of the 4° figure. Thus, eye movement planning had to rely on the selection of the figural elements and the subsequent determination of the figure center, e.g., by computing the spatial average of all figural elements. We hypothesized that eye movement planning could benefit from FGM in V1, because this signal provides a high-resolution representation of the figural elements. If so, FGM might predict the timing and the accuracy of the saccade and the spatial profile of FGM might predict the saccadic landing point.

The monkeys had to maintain fixation for 600 ms after stimulus onset and they often predicted the offset of the fixation point, because the saccade followed after a median delay of 100 ms, which is shorter than the typical reaction time in a detection task. To investigate the relationship between saccade planning and FGM we divided trials into fast responses (<100 ms after fixation point offset) and slow responses (>100 ms) and compared the FGM (Figure 7A). We observed that the magnitude of V1 FGM in the center of the figure gradually increased toward the saccade. This ramping of V1 activity occurred earlier in fast trials than in slow trials: the FGM *d*-prime was strongest in the fast trials in a window of 400–600 ms after stimulus onset ($p < 0.05$, Figure S5E). This effect also occurred, albeit weaker, for the V1 edge modulation ($p < 0.05$) but was not apparent in V4 (Figure S5E).

We next investigated the relation between FGM and the saccade landing position. We measured the deviation from the median saccade landing position for every stimulus position (Figure 2B) on every trial and selected the 25% of the trials where the saccadic endpoint deviated most to the left but still landed in the 2.5° target window (blue arrows in Figure 7B) and the 25% of the trials where the saccade deviated most to the right (red arrows). In the remaining 50% of trials the saccadic endpoint was relatively close to the center (green arrows). Figures 7C

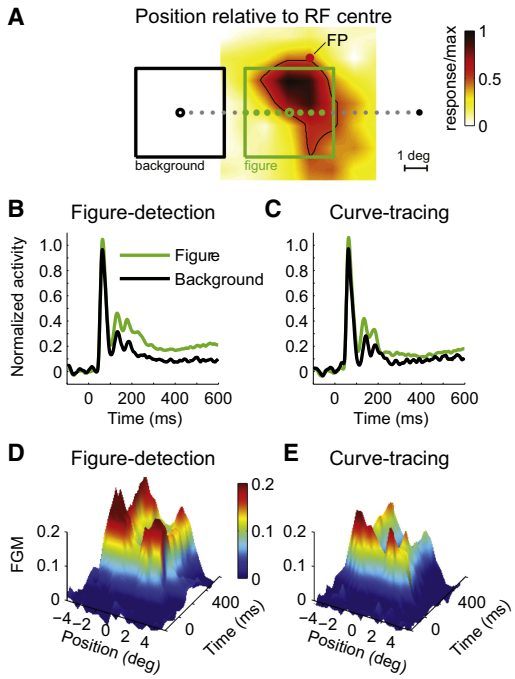

Figure 5. Effect of Attention on FGM of an Example V4 Recording Site

(A) Position of the figure relative to the RF. The red dot represents the fixation point, and the small circles represent the figure center in the different conditions (black dots represent background conditions and green dots figure conditions). The color map shows the RF as determined with a mapping task and the thin black line shows the region where the activity was more than 50% of the maximum response.

(B and C) Response evoked by the figure (green circles in A) and background positions (black circles) in the figure-detection (B) and curve-tracing task (C).

(D and E) Space-time profile of the FGM in the figure-detection (D) and curve-tracing task (E).

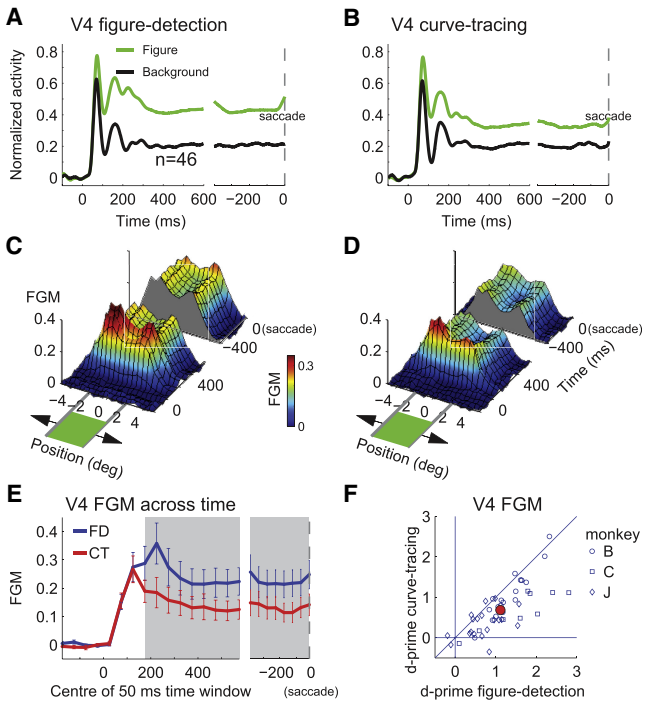

Figure 6. Population Analysis in Area V4

(A and B) Responses evoked by the figure and background in area V4 in the figure-detection (A) and curve-tracing task (B), averaged across 46 V4 recording sites, aligned to stimulus onset (left) and saccade (right).

(C and D) Space-time profile of the FGM in the figure-detection (C) and curve-tracing task (D). Here, we only included V4 recording sites for which all 23 figure positions had been measured (n = 37) (see also Figures S2C and S2D).

(E) Average FGM in successive 50 ms time windows in the figure-detection (FD, blue line) and curve-tracing task (CT, red line). Gray region, significant difference between FGM in the two tasks (paired t test, $p < 0.05$, with Bonferroni correction). Error bars show ± 2 SEM.

(F) FGM *d*-prime across recording sites in the figure-detection (x axis) and curve-tracing task (y axis) (time window 200–600 ms after stimulus onset). Red data point shows mean *d*-prime.

See also Figures S2, S3, and S4.

and 7D shows the spatiotemporal profile of V1 FGM in the trials with deviating saccades. If the saccade deviated to the left, FGM was higher on the left side of the figure and if the saccade deviated to the right FGM was strong on the right side of the figure (paired t tests, $p < 0.05$). In the trials where the saccade ended close to the center, FGM was more homogeneous (Figure 7E) and stronger ($p < 0.05$, see Supplemental Information). Accordingly, the strength of FGM in area V1 predicted saccadic accuracy (Figure S6). We observed similar effects in V4 where an increase of FGM on the left predicted that the saccade would deviate to the left, and an increase in FGM on the right predicted a deviation of the saccade to the right (Figure S6). These results suggest that the profile of FGM is read out for the accurate planning of saccades toward the center of the figure.

Latency of Visual Responses, FGM and Attention Effects in Areas V1 and V4

The relative timing of the neuronal activity evoked by the line elements, the FGM and the attention effects provides insight into the chain of events underlying figure-ground segregation. To measure the timing of visually driven activity, we fitted a curve to the average visual responses and took the time point where it reached 33% of its maximum as an estimate of latency (Roelfsema et al., 2007) (colored traces in Figure 8A, see Supplemental Information). The latency of the visual response in V1 was 40 ms and the latency in V4 was 52 ms, and a bootstrap analysis indicated that this latency difference was significant ($p < 0.01$). To measure the latency of FGM in the two areas, we fitted the same type of curve to the difference between the responses evoked by the figure and background (Figure 8A). The edge modulation in V1 had a latency of 60 ms and was followed by FGM in V4 at latency of 67 ms. These latencies were both later than the visual response in V4 ($p < 0.05$), and the difference between them was marginally significant ($p = 0.06$). Finally, the V1 center modulation occurred with a latency of 95 ms, significantly later than V1 edge-FGM and V4 FGM (both *Ps* < 0.05).

An analysis of latency across individual recording sites confirmed these effects. Activity in area V1 started with the visual response, which was followed by edge-FGM (Figure 8B,

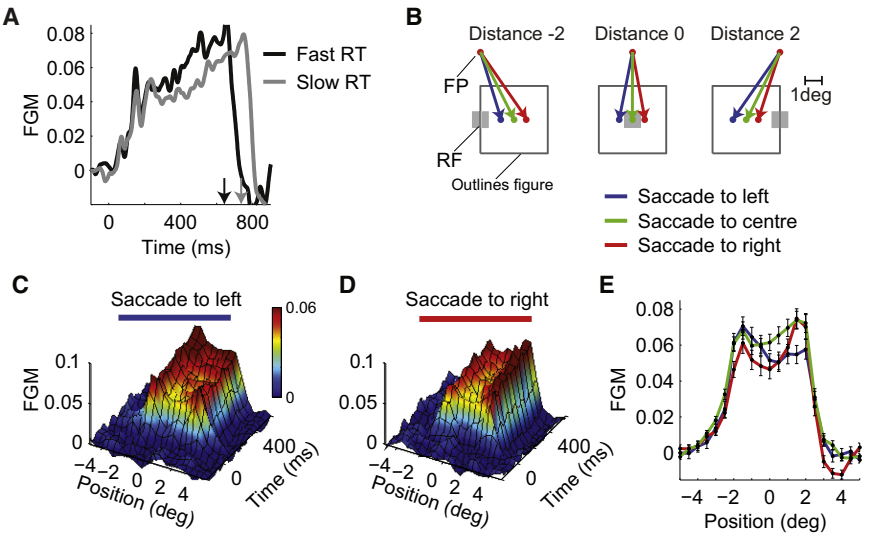

Figure 7. The Functional Role of Figure-Ground Modulation in Area V1

(A) FGM in V1 at the figure center for trials with reaction faster and slower than 100 ms. Arrows indicate the average onset of the saccade for the fast (black) and slow (gray) trials.

(B) The RF (gray square) was at one of 23 positions relative to the figure. The blue arrow indicates the median saccade vector in trials where the saccadic endpoint deviated to the left (median deviation −0.63°), the green arrow the mean saccade vector in trials where the saccade endpoint was central (median deviation 0°), and the red arrow the median saccadic endpoint for saccades deviating to the right (median deviation 0.61°).

(C and D) Spatiotemporal profile of V1 FGM in the figure-detection task for trials with saccades deviating to the left (C) and saccades deviating to the right (D), averaged across 59 V1 recording sites.

(E) Average response in time window from 400 to 600 ms for saccades deviating to the left (blue line), saccades to the center (green line), and saccades deviating to the right (red line). Error bars show ± 1 SEM.

See also Figures S5 and S6.

$p < 10^{-6}$, paired t test), which was, in turn, followed by center-FGM (Figure 8C, $p < 10^{-4}$, paired t test). Also in area V4 there was a significant delay between the visual response and the FGM (Figure 8D, $p < 10^{-4}$, paired t test).

We determined the timing of the attention effect by subtracting the FGM in the curve-tracing task from that in the figure-detection task (ΔFGM). The effect of attention occurred after 159 ms in V4, which was earlier than the attention effect on center-FGM in V1 with a latency of 204 ms ($p < 0.05$). We did not obtain a reliable latency measurement for the weak effect of attention on edge-FGM. The effects of attention on the V1 center-FGM and V4 FGM were significantly later than the V1 center-FGM (both *Ps* < 0.01). Thus, the results show an orderly progression of processing phases in the texture-segregation task, with visually driven activity preceding FGM, which was modulated by attention at an even later phase of the response (Figure 8E).

Modeling the Neuronal Interactions for Figure-Ground Segregation

Our findings indicate that boundary detection is an early, automatic process whereas the filling in of the figure-center with FGM in V1 depends on attention. These two processes have opposite requirements for the interactions between neurons (Roelfsema et al., 2002), with iso-orientation inhibition for boundary-detection and iso-orientation excitation for region filling. We created a neurodynamical model to test if these connection schemes can be combined in a single, hierarchical neural network (see Supplemental Information for details). The input into the model was an orientation defined figure on a textured background that is first represented in $V1_m$ ("m" stands for model) and was then propagated to $V2_m$ and $V4_m$ by feedforward connections. Each model area contained two maps, one for each orientation (Figure 9A) and higher areas represented the image at a courser resolution.

To achieve boundary detection, we implemented local center-surround interactions for iso-orientation suppression within each of the areas. This suppression is strongest at locations where the orientation is homogeneous and weakest at orientation discontinuities (Figure 9B), as can be seen if the activity of the two orientation maps is summed (third column in Figure 9A). Iso-orientation suppression is also strong at the figure-center in $V1_m$ and $V2_m$ where the orientation is homogeneous so that the activity is initially similar to that evoked by the background. Area $V4_m$ has a lower spatial resolution so that the representation of the edges is more diffuse, causing early FGM across the entire figure representation (Figure 9A).

The model uses feedback connections that excite neurons tuned to the same orientation to fill the entire figural region in lower areas with FGM (Figure 9A, fourth column). The figure orientation is represented in area V4 with enhanced activity, and these V4 cells excite neurons in lower areas that represent the same orientation, causing FGM to also fill the interior of the figure (Figure 9B bottom). Figures 9C and 9E illustrates how the boundary-detection and region-filling mechanisms work in concert to account for the space-time profile of FGM in $V1_m$ and $V4_m$ with early FGM at the figure boundaries and later FGM in the figure center.

To model the attention effect, we assumed that areas higher than $V4_m$ control the efficiency of $V4_m$ boundary detection. If the figure is not attended, FGM in $V4_m$ is weaker, and the model propagates this effect to V2 and V1 where center-FGM is reduced (Figure 9D). In contrast, the effect of attention on $V4_m$ has little influence on edge-FGM because it is computed locally, within $V1_m$. We also modeled the effect of a lesion in all areas higher than V1 that removes the feedback completely, and this abolished center-FGM whereas edge-FGM was preserved (Figure S7) in accordance with a previous study (Lamme et al., 1998a). These modeling results confirm that a fast and local

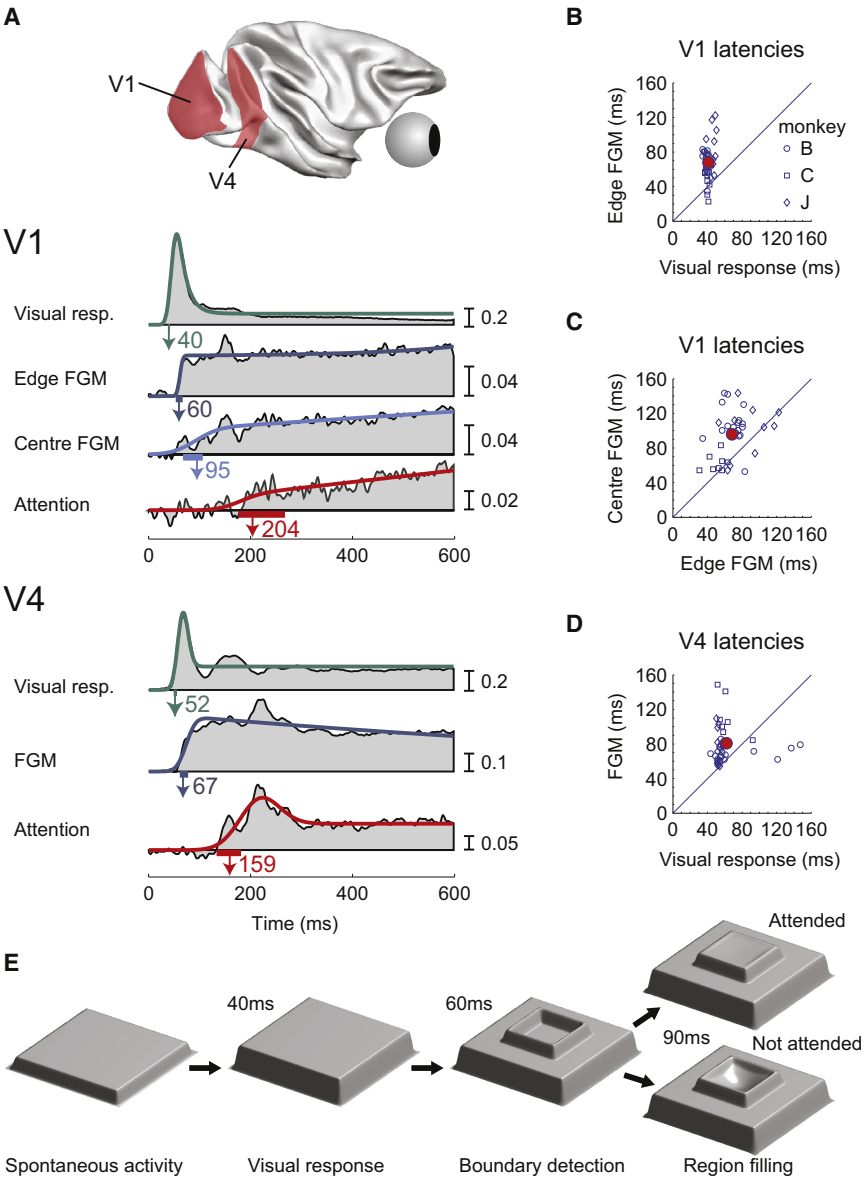

Figure 8. Timing of Visual Response, Figure-Ground Segregation, and Attention

(A) Comparison of the latency of the visual response, FGM and the attentional effect in V1 (top) and V4 (bottom). Colored traces represent curves that were fitted to the visual response (green), FGM (blue), and attentional effect (red). Arrows show the latency of these effects, and bars show the 95% confidence interval. Monkey brain image courtesy of Rainer Goebel (created with BrainVoyager software).

(B) Comparison of the latency of the visual response (x axis) and the edge modulation (y axis) across V1 recording sites. Red data point shows the mean.

(C) Comparison between the latency of edge-FGM (x axis) and center-FGM (y axis) in V1.

(D) Comparison between the latency of the visual response (x axis) and FGM (y axis) in V4.

(E) Schematic illustration of the successive phases of neuronal activity. First, features are registered across the entire visual field (visual response). Second, feature discontinuities are detected (boundary detection). This is followed by region filling: the image elements of an attended object are labeled with enhanced activity but this process is incomplete for nonattended objects.

boundary-detection mechanism based on iso-orientation inhibition, combined with a slower region-filling mechanism that uses iso-orientation excitation in feedback connections explains the space-time profile of FGM and also the influence of attention.

DISCUSSION

Here, we investigated the representation of orientation-defined figures in V1 and V4. By systemically shifting the figure position relative to the RFs and by varying behavioral relevance we obtained insights into the mechanisms for figure-ground segregation. Our results support theories that propose two complementary processes for figure-ground segregation (Grossberg and Mingolla, 1985; Mumford et al., 1987; Roelfsema et al., 2002). The first process detects boundaries between image regions with different features, and the second joins regions with similar features that usually belong to the same object. We observed an early enhancement of neuronal activity at the boundaries between figure and background at multiple spatial scales (in V1 and V4) and found that the neuronal correlates of boundary detection depend only weakly on attention. Boundary detection is followed by filling of the interior of the figure with enhanced neuronal activity, and this later process has a stronger dependency on attention. The connectivity schemes for boundary detection and region filling differ, because the former requires iso-orientation inhibition and the latter iso-orientation excitation. Our modeling results show that these conflicting constraints can be met by different processes with their own topology of connections and time course (Roelfsema et al., 2002; Scholte et al., 2008). We suggest that boundary detection relies on a local iso-orientation inhibition scheme, whereas region filling is the result of corticocortical feedback connections that implement iso-orientation excitation.

By simultaneously recording neuronal activity in two visual cortical areas in monkeys, we delineated the sequence of events in the texture-segregation task (Figure 8E), which fit well with neuroimaging results in humans (Scholte et al., 2008). First, information about the stimulus features is propagated from the LGN to V1 and then onward to V4. Slightly later, the representation of figure edges is enhanced in V1 as is the figure representation in

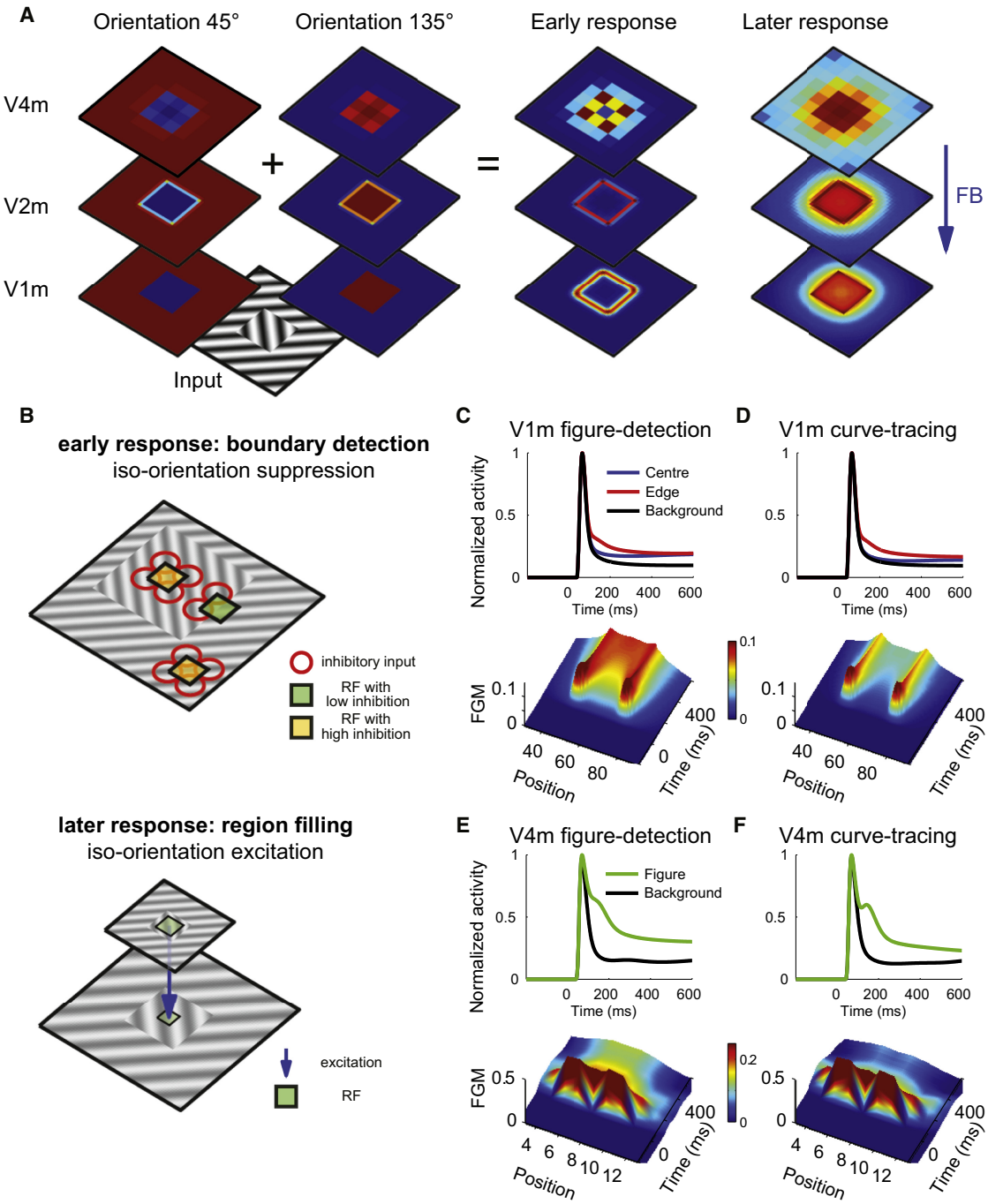

Figure 9. Model of Figure-Ground Segregation and the Effect of Attention in Visual Cortex

(A) The model areas $V1_m$, $V2_m$, and $V4_m$ respond to an orientation-defined square figure (bottom). In each area, the input is represented in two activity maps encoding stimulus elements with an orientation of 45° (first column) and 135° (second column). The third column shows the response of the model after 50 ms summed across the two orientation maps. The orientation discontinuities are detected during this phase of the response (warm colors). The decrease in spatial resolution in higher areas (in particular in $V4_m$) causes a more diffuse FGM across the entire figure representation. The right column shows the activity after 400 ms, when feedback connections (FB, blue arrow) have propagated the enhanced activity back to lower areas causing a filling-in of center of the figure with FGM.

(B) Complementary mechanisms for figure-ground segregation. Top: boundary detection by iso-orientation suppression. RFs in regions with a homogeneous orientation receive strong inhibition (yellow squares) from nearby locations (red circles). The RF on the boundary (green) receives less iso-orientation suppression, and has a stronger response. Bottom: region filling by excitatory feedback connections that mediate iso-orientation excitation. Feedback connections propagate FGM from orientation selective units in higher areas to units in lower areas tuned to the same orientation.

(C–F) Responses and space-time profile of the FGM in model areas $V1_m$ and $V4_m$ in the figure-detection (C and E) and curve-tracing task (D and F).

See also Figure S7 and Table S1.

V4 (Burrows and Moore, 2009). It then takes a few tens of milliseconds before FGM also emerges in the center of the figure in V1 (Lamme et al., 1999). The effects of attention are observed at yet later time points; attention first increases FGM in V4 and it then also boosts center modulation in V1 (Ogawa and Komatsu, 2006; Roelfsema et al., 2007).

One must be cautious when inferring connectivity from latency differences alone. For example, the effect of feedback to V1 may under some conditions be faster than influences caused by horizontal connections (Bair et al., 2003). However, the difference between the mechanisms for edge- and center-FGM is supported by a number of additional observations. First, task-driven attention boosted the representation of the figure center and had less effect on the edge representation (Figure 8E). This implies that edge-FGM is largely stimulus driven, whereas center-FGM depends more on feedback from higher areas. Second, a previous study (Lamme et al., 1998a) showed that lesions in higher visual areas reduce center FGM in V1 but leave edge modulation intact (see also Hupé et al., 1998). Third, we could reproduce the timing and the spatial profile of the visual responses, the FGM and the attentional modulation in V1 and V4 with a model that detects boundaries with local inhibition and uses excitatory feedback for region filling. These results imply that the mechanisms proposed by us are sufficient to explain the data.

Edge Modulation

The enhancement of neuronal activity at boundaries occurs quickly (Lamme et al., 1999; Nothdurft et al., 2000) and is not strongly modulated by attention. Previous studies demonstrated that texture elements surrounded by dissimilar elements are more salient (Joseph and Optican, 1996). Image elements that pop out cause stronger neuronal activity in visual cortex during an early response phase (Burrows and Moore, 2009; Kastner et al., 1997; Knierim and van Essen, 1992; Lamme et al., 1999; Lee et al., 2002; Nothdurft et al., 1999; Ogawa and Komatsu, 2006) and a similar increase in V1 activity occurs at the location of an edge where the orientation changes abruptly (Nothdurft et al., 2000). These saliency effects also occur when animals ignore the stimulus (Knierim and van Essen, 1992) (but see Burrows and Moore, 2009), and even if they are anesthetized (Kastner et al., 1997; Nothdurft et al., 1999, 2000). Accordingly, image elements can pop out in psychophysics (Theeuwes et al., 2006) if they are not relevant to the task, although these effects are transient and disappear after 250 ms (Donk and van Zoest, 2008; Joseph and Optican, 1996). It is likely that edge-FGM is related to neuronal responses in V1, V2, and V4 that reflect the assignment of the edge to the figural side, because borders "belong" to figures and not to the background (Zhou et al., 2000). This border-ownership signal depends on attention (Qiu et al., 2007) and future studies could compare it to edge-FGM within a single task.

Here, we modeled the boundary-detection process with a connection scheme where units tuned to the same orientation inhibit each other (Itti and Koch, 2001; Li, 1999; Roelfsema et al., 2002) so that singletons and orientation boundaries evoke stronger activity. Our result that the edge-FGM in V1 and the FGM in V4 occurred at approximately the same time is in accordance with such a local computational scheme: the edge enhancement in V1 does not depend on feedback, in accordance with a study demonstrating that the V1 pop-out signal also occurs if V2 is not active (Hupé et al., 2001).

Center Modulation

FGM in the center of the figure depends on task relevance, which suggests that it is controlled by feedback from higher areas. Accordingly, our model implemented iso-orientation excitation in the feedback connections. Thus, if the figure orientation is 135°, V4 neurons tuned to this orientation increase their response and propagate the enhanced activity back to V2 and V1 neurons tuned to 135° so that FGM is confined to the figure. The attentional effect uses the same route and as a result it is object-based (Figures 1A and 1C).

The influence of attention on center-FGM accounts for a discrepancy in the literature. In contrast to a number of other studies (Lamme, 1995, 1999; Marcus and Van Essen, 2002; Zipser et al., 1996), Rossi et al. (2001) did not observe center modulation in area V1. Interestingly, the monkeys of their study did not have to detect the figure, except in one experiment with a monkey that discriminated between a figure at a fixed location and a homogeneous background, which is a task that could be solved by detecting one of the boundaries. In contrast, the monkeys of our study made eye movements to the center of a figure that varied in its location, which presumably required perception of the entire figure and presumably depends on FGM at the figure center (Supèr et al., 2001). Moreover, our monkeys had a lot of experience in localizing the figure, and training amplifies the modulation of V1 activity (Li et al., 2008). The influence of attention on center-FGM may have also contributed to the absence of FGM in V1 in two fMRI studies because the subjects' attention was directed away from the figure (Kastner et al., 2000; Schira et al., 2004). A previous study by Marcus and Van Essen (2002) also investigated the effects of figure-ground segregation and attention in V1 and V2 in monkeys. Attention enhanced V2 activity but it did not increase FGM and had little effect on activity in V1. However, in this study the monkeys always attended one of two similar figures and it is possible that the monkeys perceived both figures, because increases in the number of figures does not diminish FGM (Lamme et al., 1998b; Landman et al., 2003). In the present study, the monkeys could ignore the texture for an entire recording day with the curve-tracing task, which may explain why we observed stronger attentional effects and interactions between FGM and task-driven, object-based attention.

The Functional Role of Figure-Ground Modulation in V1 and V4

In our task, the monkeys had to maintain fixation for at least 600 ms. The majority of saccades in a reaction time version of the texture-segregation task are made between 250 and 400 ms after stimulus presentation (Supèr, 2006). Thus, FGM and the effect of attention start well before the initiation of the saccade. What could be the role of FGM and attention in early visual areas?

Our task demanded precise saccades because the eye had to land in a 2.5° window centered on the 4° figure. Eye movement

planning could benefit from FGM in V1, because it provides a high-resolution representation of all figural elements (Mumford, 1996). Previous studies using luminance defined figures showed that the average landing position of the eye is close to the center of the area of a figure, while the distribution of saccadic endpoints depends on shape (Melcher and Kowler, 1999; Moore, 1999). In the texture segregation task, saccade planning requires a selective averaging of the position of texture elements of the figure, discarding the background elements. V1 projects to the superior colliculus and this pathway might exploit FGM to determine the center of the figure as target of the saccade (Fries and Distel, 1983; Wurtz and Albano, 1980). In accordance with this view, FGM in V1 and V4 predicted the saccade landing position. If FGM was stronger on one side of the figure, the saccadic endpoint was biased toward that side and trials where FGM was strongest throughout the figure were associated with the most accurate saccades. These results are consistent with a previous study that demonstrated that V4 neurons predict how well the saccadic endpoint aligns with a visual stimulus (Moore, 1999). Furthermore, we found that activity in the figure center ramped up until the moment of the saccade (see also Supèr et al., 2004). This temporal profile of FGM is reminiscent of studies on decision making that observed similar effects in parietal cortex (see, for example, Gold and Shadlen, 2007). Thus, our results establish that FGM in area V1 is strongly related to the spatial and temporal features of saccadic eye movements, possibly via a route through the superior colliculus.

The Interplay between Figure-Ground Segregation and Attention

Perceptual organization refers to the set of processes that are responsible for the grouping of features into objects and for the segregation of features that belong to different objects and the background. Our results strengthen previous proposals that the processes for perceptual organization influence attentional processes, and vice versa, that attention acts on texture-defined figures in an object-based manner to influence perceptual organization (Bhatt et al., 2007; Driver et al., 2001; Qiu et al., 2007; Scholl, 2001; Walther and Koch, 2006). Features of the same surface are thereby labeled in the visual cortex with enhanced neuronal activity, as if the enhanced activity binds features to reconstruct objects from distributed feature representations (Roelfsema, 2006; Roelfsema and Houtkamp, 2011). We found that this reconstruction is only partial for objects that are irrelevant for behavior, suggesting that the visual cortex leaves irrelevant representations in a more primordial state and only fully labels representations of relevant objects. These high-resolution representations in early visual areas can then be used to guide behavioral responses toward objects of interest.

EXPERIMENTAL PROCEDURES

Behavioral Tasks

Three monkeys participated in the study. The animals performed a figure-detection task and a curve-tracing task on alternate days (interleaved design) with identical stimuli. The animals were seated at a distance of 0.75 m from a monitor (width 0.375 m) with a resolution of 1,024 × 768 pixels and a frame rate of 100 Hz. A trial started as soon as the monkey's eye position was within a 1° × 1° window centered on a red fixation point (0.2°, on a gray background with luminance of 14 cd·m^{-2}). When the monkey had kept his gaze for 300 ms on the fixation point, the stimulus appeared with a square figure and two curves on a background with line elements (Figure 2A). The stimulus stayed in view, while the monkey maintained fixation for at least an additional 600 ms, and then the fixation point disappeared, cueing the monkey to make a saccade (Figure 2C). In the figure-detection task, the monkey had to make an eye movement into a target window of 2.5° × 2.5° centered on the middle of the figure square. In the curve-tracing task the monkey had to make a saccade into a target window of 2.5° × 2.5° centered on the circle that was attached to the curve connected to the fixation point (target curve, T) while ignoring the other curve (distracter curve, D). Correct responses were rewarded with apple juice.

The monkey performed one of the tasks on each day. We cued the monkey which task to perform by starting every session with trials with only the figure (without curves) or only the curves on a homogeneously textured background. After a number of trials (~10), we introduced the stimuli with the two curves and the figure. Data collection started when the performance of the monkey was above 85%. The accuracy in the figure detection and in the curve-tracing task was 97% and 92% in monkey B, 99% and 91% in monkey C, 99% and 96% in monkey J, respectively.

Visual Stimulus

The figure-ground stimulus consisted of a square figure with oriented line elements (16 pixels long, 0.44°, and 1 pixel wide) on a background with an orthogonal orientation (Figure 2A).

The two orientations that we used for the line elements (45° and 135) were counterbalanced across conditions so that the average receptive field stimulus was identical (see Supplemental Experimental Procedures for details).

The figure always appeared in the same half of the screen (bottom half for monkeys B and J, left half for monkey C). We varied the position of the center of the figure relative to the center of the RF of the recording sites on one spatial dimension (x axis for monkeys B and J, y axis for monkey C), while keeping the position on the other dimension constant (at the x or y coordinate of the RF center) so that the RF fell on the background, on the edge or on the center of the figure in different conditions. A total of 23 figure positions were presented (the distance of the RF center relative to the figure center ranged from −5.5° to 5.5° with 0.5° steps, see Figure 2B). For 9 of the 46 V4 recording sites, we did not present figures at all these positions, but we used a subset of five positions (one center, two edge, and two background positions), and the data from these recording sites were not included in the space-time plots (Figures 6C and 6D). The stimulus also contained two curves (width 0.27°, luminance 82 cd·m^{-2}) and two red circles (size 1.5°) in the hemifield opposite to the figure (upper hemifield for monkeys B and J and right hemifield for monkey C). One of the curves was connected to the fixation point (target curve) and the other curve was not (distracter curve). A small change close to the fixation point switched the target and distracter curve (in the example of Figures 2A and 2C the left curve is the target curve but in other trials the right curve was connected to the fixation point). All 23 figure positions × 2 curve configurations were presented in a randomly interleaved sequence in both tasks.

Surgical Procedure

The animals underwent two surgeries under general anesthesia that was induced with ketamine (15 mg kg^{-1} injected intramuscularly) and maintained after intubation by ventilation with a mixture of 70% N_2O and 30% O_2, supplemented with 0.8% isoflurane, fentanyl (0.005 mg kg^{-1} intravenously), and midazolam (0.5 mg kg^{-1} h^{-1} intravenously). In the first operation a head holder was implanted and a gold ring was inserted under the conjunctiva of one eye for the measurement of eye position.

In the second operation, arrays of 4 × 5 electrodes (Cyberkinetics Neurotechnology Systems Inc.) were chronically implanted in areas V1 and V4 (see Figure S1). All procedures complied with the NIH Guide for Care and Use of Laboratory Animals (National Institutes of Health, Bethesda, Maryland), and were approved by the institutional animal care and use committee of the Royal Netherlands Academy of Arts and Sciences.

Recording of Neuronal Activity

Details about the recording methods and information about the measurement of RFs in V1 and V4 can be found in Supplemental Experimental Procedures. We quantified visual responsiveness by first calculating the spontaneous mean activity, *Sp*, and the standard deviation, *s*, across trials in a 200 ms time window preceding stimulus onset. We then computed the peak response, *Pe*, by smoothing the average response over conditions with a moving window of 25 ms and taking the maximum during the stimulus period (0–600 ms after stimulus onset). The visual responsiveness index was then given by $VR = (Pe\text{-}Sp)/s$. Only recording sites with a good visual response ($VR > 3$) were included in the analyses. Because of the many conditions (23 figure positions and 2 attention conditions) we averaged responses across recording sessions on different days. We only included recording sites with at least three figure-detection and three curve-tracing sessions in the analysis (the average number of sessions in every task was 5). This resulted in a sample of 59 recording sites in V1 (18 in monkey B, 15 in monkey C, and 26 in monkey J) and 46 in V4 (18 in monkey B, 9 in monkey C, and 19 in monkey J). The average number of trials per condition was 484. We did not observe significant differences in visual responsiveness between the figure-detection and the curve-tracing tasks (paired t test; V1 and V4, $p > 0.05$).

Data Analysis

We separately analyzed the V1 responses evoked by the texture background, the figure edge and the figure center. To create the space-time plots in Figures 4C, 4D, 6C, and 6D we first computed for every RF the distance of the figure center to the RF center in all stimulus conditions and rounded these distances to the nearest multiple of 0.5°. We then computed for every figure position (0.5° steps) an average response across RFs. In area V1, background responses were obtained by averaging across conditions where the RF-center was separated from the nearest figure edge by at least 2°. Edge responses were averaged across conditions where one of the two edges fell in the RF, and center responses across conditions where the RF-center was within 0.5 degrees of the figure center. In V4, we did not distinguish between edge and figure responses, because the large RFs typically did not fit entirely within the figure. Instead, we compared figure positions where the figure completely or partially covered the RF to background positions where the figure was outside the RF. We took as background responses those conditions where the RF border was separated from the nearest figure edge by at least 2° (except for 5 recording sites with distance <1.8°, and 3 with distance <1°) and as figure responses those configurations where the RF hot spot (point in V4 RF with maximum response) was within 2° from the figure center. To compute the population responses, we first normalized the responses before averaging across recording sites by subtracting *Sp* and dividing the result by (*Pe*-*Sp*) (*Sp* and *Pe* were defined above).

As a measure of the reliability of the FGM across trials, we computed the FGM *d*-prime: $d'_{FGM} = (\bar{y}_{Fig} - \bar{y}_{Bck})/s_{Act}$, where $\bar{y}_{Fig}$ and $\bar{y}_{Bck}$ is the average response evoked by the figure and background, respectively, and s_{Act} is the pooled standard deviation across the two conditions. For each recording site, we obtained an estimate of the FGM *d*-prime during figure detection and curve tracing by averaging *d*-prime values across the corresponding sessions.

Model of Figure-Ground Modulation and Attention Effects

The model consists of areas $V1_m$, $V2_m$, and $V4_m$ (the subscript "m" stands for model). The input signal arrives in two maps. One map encodes texture elements of 45° orientation, while the other map encodes texture elements of 135° (Figure 9A). The input first arrives in $V1_m$ which then activates area $V2_m$, which, in turn, projects to area $V4_m$ (Figure S7). In the model higher areas have larger RFs. RF width (and height) in area $V4_m$ is four times larger than the RF width in $V2_m$, which, in turn, is twice as large as the width in $V1_m$. In addition to the feedforward projections, higher areas also provide feedback to lower areas (see Supplemental Information for equations).

Each model area detects texture discontinuities through local center-surround interactions causing iso-orientation suppression. These center-surround interactions cause suppression in regions with a homogeneous orientation and a comparatively stronger response at the representation of the orientation boundaries in $V1_m$ and $V2_m$. In $V4_m$, the RFs are so large that the boundaries are not resolved so that entire figural region acts as a pop-out stimulus causing stronger activity for the figural orientation. This pop-out effect propagates via the feedback connections to neurons that respond to the same orientation in lower areas, causing a filling in of enhanced activity at the figure center.

To model the effect of attention, we varied the efficiency of the $V4_m$ boundary-detection process (see Supplemental Information) with stronger FGM in the figure-detection task. The feedback connections propagate this effect to $V2_m$ and $V1_m$ where the strength of the center modulation increases if the figure is attended.

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures, one table, and Supplemental Experimental Procedures and can be found with this article online at http://dx.doi.org/10.1016/j.neuron.2012.04.032.

ACKNOWLEDGMENTS

We thank Kor Brandsma, Dave Vleesenbeek, and Anneke Ditewig for biotechnical assistance. The research leading to these results has received funding from the European Union Sixth and Seventh Framework Programmes (EU IST Cognitive Systems, project 027198 "Decisions in Motion" and project 269921 "BrainScaleS") and a NWO-VICI grant awarded to P.R.R. F.R. is supported in part by CELEST, a National Science Foundation Science of Learning Center (NSF OMA-0835976) and the Office of Naval Research (ONR N00014-11-1-0535). V.A.F.L. is supported by an advanced investigator grant from the European Research Council. H.N. is supported by the Transregional Collaborative Research Centre SFB/TRR 62 "Companion-Technology for Cognitive Technical Systems" funded by the German Research Foundation (DFG).

Accepted: April 19, 2012
Published: July 11, 2012

REFERENCES

Allen, H.A., Humphreys, G.W., Colin, J., and Neumann, H. (2009). Ventral extra-striate cortical areas are required for human visual texture segmentation. J. Vis. *9*, 2–14, 1–14.

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.

Bhatt, R., Carpenter, G.A., and Grossberg, S. (2007). Texture segregation by visual cortex: perceptual grouping, attention, and learning. Vision Res. *47*, 3173–3211.

Burrows, B.E., and Moore, T. (2009). Influence and limitations of popout in the selection of salient visual stimuli by area V4 neurons. J. Neurosci. *29*, 15169–15177.

De Weerd, P., Sprague, J.M., Vandenbussche, E., and Orban, G.A. (1994). Two stages in visual texture segregation: a lesion study in the cat. J. Neurosci. *14*, 929–948.

Desimone, R., and Duncan, J. (1995). Neural mechanisms of selective visual attention. Annu. Rev. Neurosci. *18*, 193–222.

Donk, M., and van Zoest, W. (2008). Effects of salience are short-lived. Psychol. Sci. *19*, 733–739.

Driver, J., and Baylis, G.C. (1996). Edge-assignment and figure-ground segmentation in short-term visual matching. Cognit. Psychol. *31*, 248–306.

Driver, J., Davis, G., Russell, C., Turatto, M., and Freeman, E. (2001). Segmentation, attention and phenomenal visual objects. Cognition *80*, 61–95.

Duncan, J. (1984). Selective attention and the organization of visual information. J. Exp. Psychol. Gen. *113*, 501–517.

Fries, W., and Distel, H. (1983). Large layer VI neurons of monkey striate cortex (Meynert cells) project to the superior colliculus. Proc. R. Soc. Lond. B Biol. Sci. *219*, 53–59.

Gold, J.I., and Shadlen, M.N. (2007). The neural basis of decision making. Annu. Rev. Neurosci. *30*, 535–574.

Grossberg, S., and Mingolla, E. (1985). Neural dynamics of form perception: boundary completion, illusory figures, and neon color spreading. Psychol. Rev. *92*, 173–211.

Huang, X., and Paradiso, M.A. (2008). V1 response timing and surface filling-in. J. Neurophysiol. *100*, 539–547.

Hupé, J.M., James, A.C., Payne, B.R., Lomber, S.G., Girard, P., and Bullier, J. (1998). Cortical feedback improves discrimination between figure and background by V1, V2 and V3 neurons. Nature *394*, 784–787.

Hupé, J.M., James, A.C., Girard, P., and Bullier, J. (2001). Response modulations by static texture surround in area V1 of the macaque monkey do not depend on feedback connections from V2. J. Neurophysiol. *85*, 146–163.

Itti, L., and Koch, C. (2001). Computational modelling of visual attention. Nat. Rev. Neurosci. *2*, 194–203.

Joseph, J.S., and Optican, L.M. (1996). Involuntary attentional shifts due to orientation differences. Percept. Psychophys. *58*, 651–665.

Kanizsa, G., and Gerbino, W. (1976). Convexity and symmetry in figure-ground organization. In Vision and Artifact, M. Henle, ed. (New York: Springer), pp. 25–32.

Kastner, S., Nothdurft, H.C., and Pigarev, I.N. (1997). Neuronal correlates of pop-out in cat striate cortex. Vision Res. *37*, 371–376.

Kastner, S., De Weerd, P., and Ungerleider, L.G. (2000). Texture segregation in the human visual cortex: A functional MRI study. J. Neurophysiol. *83*, 2453–2457.

Knierim, J.J., and van Essen, D.C. (1992). Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. J. Neurophysiol. *67*, 961–980.

Koffka, K. (1935). Principles of Gestalt Psychology (New York: Harcourt, Brace and Company).

Lamme, V.A. (1995). The neurophysiology of figure-ground segregation in primary visual cortex. J. Neurosci. *15*, 1605–1615.

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

Lamme, V.A., Zipser, K., and Spekreijse, H. (1998b). Figure-ground activity in primary visual cortex is suppressed by anesthesia. Proc. Natl. Acad. Sci. USA *95*, 3263–3268.

Lamme, V.A., Rodriguez-Rodriguez, V., and Spekreijse, H. (1999). Separate processing dynamics for texture elements, boundaries and surfaces in primary visual cortex of the macaque monkey. Cereb. Cortex *9*, 406–413.

Landman, R., Spekreijse, H., and Lamme, V.A. (2003). Set size effects in the macaque striate cortex. J. Cogn. Neurosci. *15*, 873–882.

Lee, T.S., Yang, C.F., Romero, R.D., and Mumford, D. (2002). Neural activity in early visual cortex reflects behavioral experience and higher-order perceptual saliency. Nat. Neurosci. *5*, 589–597.

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

Li, Z. (1999). Contextual influences in V1 as a basis for pop out and asymmetry in visual search. Proc. Natl. Acad. Sci. USA *96*, 10530–10535.

Marcus, D.S., and Van Essen, D.C. (2002). Scene segmentation and attention in primate cortical areas V1 and V2. J. Neurophysiol. *88*, 2648–2658.

Melcher, D., and Kowler, E. (1999). Shapes, surfaces and saccades. Vision Res. *39*, 2929–2946.

Merigan, W.H. (1996). Basic visual capacities and shape discrimination after lesions of extrastriate area V4 in macaques. Vis. Neurosci. *13*, 51–60.

Moore, T. (1999). Shape representations and visual guidance of saccadic eye movements. Science *285*, 1914–1917.

Mumford, D. (1996). Commentary on "Banishing the homunculus" by H. Barlow. In Perception as Bayesian Inference, D.C. Knill and W. Richards, eds. (Cambridge, UK: Cambridge University Press).

Mumford, D., Kosslyn, S.M., Hillger, L.A., and Herrnstein, R.J. (1987). Discriminating figure from ground: the role of edge detection and region growing. Proc. Natl. Acad. Sci. USA *84*, 7354–7358.

Nothdurft, H.C., Gallant, J.L., and Van Essen, D.C. (1999). Response modulation by texture surround in primate area V1: correlates of "popout" under anesthesia. Vis. Neurosci. *16*, 15–34.

Nothdurft, H.C., Gallant, J.L., and Van Essen, D.C. (2000). Response profiles to texture border patterns in area V1. Vis. Neurosci. *17*, 421–436.

Ogawa, T., and Komatsu, H. (2006). Neuronal dynamics of bottom-up and top-down processes in area V4 of macaque monkeys performing a visual search. Exp. Brain Res. *173*, 1–13.

Peterson, M.A., Harvey, E.M., and Weidenbacher, H.J. (1991). Shape recognition contributions to figure-ground reversal: which route counts? J. Exp. Psychol. Hum. Percept. Perform. *17*, 1075–1089.

Posner, M.I., Snyder, C.R., and Davidson, B.J. (1980). Attention and the detection of signals. J. Exp. Psychol. *109*, 160–174.

Qiu, F.T., Sugihara, T., and von der Heydt, R. (2007). Figure-ground mechanisms provide structure for selective attention. Nat. Neurosci. *10*, 1492–1499.

Reynolds, J.H., and Chelazzi, L. (2004). Attentional modulation of visual processing. Annu. Rev. Neurosci. *27*, 611–647.

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

Roelfsema, P.R., and Houtkamp, R. (2011). Incremental grouping of image elements in vision. Atten. Percept. Psychophys. *73*, 2542–2572.

Roelfsema, P.R., Lamme, V.A., Spekreijse, H., and Bosch, H. (2002). Figure-ground segregation in a recurrent network architecture. J. Cogn. Neurosci. *14*, 525–537.

Roelfsema, P.R., Tolboom, M., and Khayat, P.S. (2007). Different processing phases for features, figures, and selective attention in the primary visual cortex. Neuron *56*, 785–792.

Rossi, A.F., Desimone, R., and Ungerleider, L.G. (2001). Contextual modulation in primary visual cortex of macaques. J. Neurosci. *21*, 1698–1709.

Rubin, E. (1915). Visuell wahrgenommene figuren (Copenhagen: Glydendalske).

Schira, M.M., Fahle, M., Donner, T.H., Kraft, A., and Brandt, S.A. (2004). Differential contribution of early visual areas to the perceptual process of contour processing. J. Neurophysiol. *91*, 1716–1721.

Scholl, B.J. (2001). Objects and attention: the state of the art. Cognition *80*, 1–46.

Scholte, H.S., Jolij, J., Fahrenfort, J.J., and Lamme, V.A. (2008). Feedforward and recurrent processing in scene segmentation: electroencephalography and functional magnetic resonance imaging. J. Cogn. Neurosci. *20*, 2097–2109.

Supèr, H. (2006). Figure-ground activity in V1 and guidance of saccadic eye movements. J. Physiol. Paris *100*, 63–69.

Supèr, H., Spekreijse, H., and Lamme, V.A. (2001). Two distinct modes of sensory processing observed in monkey primary visual cortex (V1). Nat. Neurosci. *4*, 304–310.

Supèr, H., van der Togt, C., Spekreijse, H., and Lamme, V.A. (2004). Correspondence of presaccadic activity in the monkey primary visual cortex with saccadic eye movements. Proc. Natl. Acad. Sci. USA *101*, 3230–3235.

Theeuwes, J., Reimann, B., and Mortier, M. (2006). Visual search for featural singletons: No top-down modulation, only bottom-up priming. Vis. Cogn. *14*, 466–489.

Treue, S. (2001). Neural correlates of attention in primate visual cortex. Trends Neurosci. *24*, 295–300.

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

Vecera, S.P., Flevaris, A.V., and Filapek, J.C. (2004). Exogenous spatial attention influences figure-ground assignment. Psychol. Sci. *15*, 20–26.

Walther, D., and Koch, C. (2006). Modeling attention to salient proto-objects. Neural Netw. *19*, 1395–1407.

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

Wurtz, R.H., and Albano, J.E. (1980). Visual-motor function of the primate superior colliculus. Annu. Rev. Neurosci. *3*, 189–226.

Zhaoping, L. (2003). V1 mechanisms and some figure-ground and border effects. J. Physiol. Paris *97*, 503–515.

Zhou, H., Friedman, H.S., and von der Heydt, R. (2000). Coding of border ownership in monkey visual cortex. J. Neurosci. *20*, 6594–6611.

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