

Fig. 1. Structure of the model network. Each circle stands for a population of neurons with similar receptive fields. Edges and other local features of a figure (black parallelogram) activate edge cells (E) that project to border ownership cells (B) that have the same preferred orientation and position, but opposite side-of-figure preferences (in this example B_L on the left side of the figure and B_R on the right). B cells have reciprocal excitatory feedback connections with grouping cells (G) that integrate global contour information at multiple scales (only one scale is shown). The grouping cells bias border ownership cell activity according to the location of the figure. Because a border can be owned by only one figure, opposing border ownership cells compete directly via IE cells and indirectly via grouping cells. The B cells excite inhibitory border ownership cells (IB; again with the indexes L and R) of the same preferred side of the figure, which inhibit G cells in all directions except the preferred one. Thus, grouping cells that are activated by inconsistent figures mutually inhibit each other via B and IB cells and unstructured input creates only small activation of G cells. A weaker, broader, and less specific inhibitory interaction has been introduced in an extension of the model to explain the results presented in Fig. 5. This inhibition is realized by direct G to IB connections, depicted as dashed lines. Top-down attention is modeled as a broad and purely spatial input to the grouping cells (top). Blue and red connections are inhibitory and excitatory, respectively. High-contrast symbols indicate cells and connections activated by the shown figure. Blue, red, and orange connections are inhibitory, feedforward excitatory, and feedback excitatory, respectively.

throughout our model: Attentional modulation affects the activity of stimulated but not of nonstimulated neurons (Fig. S2). This quasi-multiplicative effect is due to the connectivity patterns and does not require complex neural properties (our model neurons are simple linear threshold units).

Fig. 2*B* shows a visual scene in which two identical stimuli (squares) are present. The attention effect is shown in Fig. 2*D*. Both E and B cells (columns 2 and 3) show substantial modulation, in agreement with neurophysiological findings (10). There is also an attentional enhancement in all grouping cell layers (columns 4–6), most prominently at the scale closest to that of the stimulus (column 5). Column 7 shows how G-cell responses are sharpened and scaled to the stimulus size.

Fig. 2 *E–H* shows other stimulus configurations; only the lower quadrant of the visual scene is shown with the stimulus (white lines) in the first column, together with the attentional input (superposed Gaussian shading), which is identical in all cases. As in Fig. 2 *C* and *D*, the following rows show the difference of responses to attended and unattended stimuli. In Fig. 2*E*, the stimulus is the same square as in Fig. 2*D* but displaced in space. The attentional input is unchanged and therefore no longer centered on this stimulus. Nevertheless, E and B cells are appropriately modulated by attention (columns 2 and 3) and the attention field is likewise repositioned at the G-cell level (columns 5 and 7). Fig. 2 *F* and *G* shows that the attentional modulation is also tuned to the object scale: Even though the stimulus is smaller (Fig. 2*F*) or larger (Fig. 2*G*) than that used in Fig. 2*B*, attention modulates the activities of all neuronal populations appropriately. Fig. S3 shows the activities

of inhibitory populations, and Fig. S4 shows the behavior of a model with direct between-object inhibition.

The attentional mechanism does not depend on grouping cells that are tailored to specific sensory stimuli. G cells are generically activated by a large variety of stimuli whose only common factor is that they are located in approximately the same region of the visual field. For instance, Fig. 2*H* shows that top-down attention can be directed to an isolated line (which has no equivalent in the grouping cell layer). Rather than creating a small focus of strong activation in the G-cell layer, as is the case for stimuli that correspond to the spatial scale of grouping cells (Fig. 2*D–G*), the line activates grouping cells of all sizes and on both of its sides. We also note an object-superiority effect: Activity of B cells is higher when an edge is part of an object than when the same edge is shown in isolation (compare Fig. 2*D* and *H*, column 3).

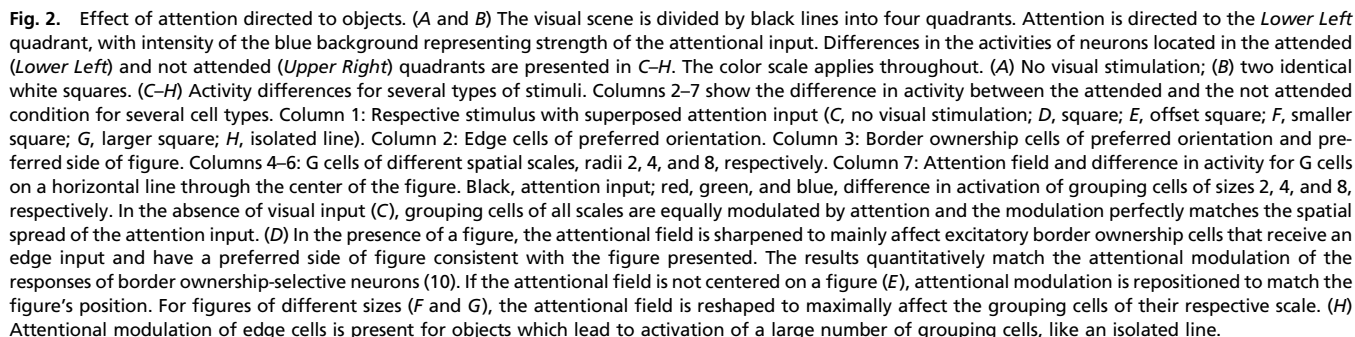
Attention needs to work in complex, cluttered environments where space alone is not sufficient to segregate foreground from background. Qiu et al. (10) studied responses in area V2 when two overlapping objects were presented and attention was directed either to the foreground or to the background object (Fig. 3). In the absence of attention, the figure-ground organization effect can be observed (Fig. 3, row 1). Of particular interest is the border ownership signal (difference between the firing rates of B cells of opposite directions) for the edge separating the foreground and the background figure. Global context and local features (T-like junctions) assign this edge to the foreground figure (see arrows in last column), in agreement with physiological results (12). This assignment is even stronger if attention is on the foreground figure (Fig. 3, row 2). However, if attention is directed toward the background figure, the border ownership modulation of the edge between the objects is nearly abolished (Fig. 3, row 3).

This result is not only in agreement with physiological evidence (10), it is also functionally desirable. When the background object is attended, the pattern-matching mechanisms downstream should receive *only* edge information from this object. In the occluded area, no such information is available and no information should therefore be sent to downstream processing areas; including edge information (from the foreground object) would lead them astray. Importantly, this result cannot be achieved by purely spatial attention because the occluding edge is closer to the center of the background object (the spotlight of attention) than to the foreground object. The ability to attend to objects in clutter is perhaps the most important function of the grouping mechanisms. Further tests of our model with figures embedded in noise show that the same broadly focused top-down attention can bring out figures from the background of noise (Fig. S5).

We performed a global sensitivity analysis (21) to determine the dependence of several attentional modulation indexes, border ownership modulation, and their interaction on several model parameters. Border ownership modulation was sensitive only to the strength of the excitatory feedback circuits (lateral connections between B cells and feedback from G to B cells). Attentional modulation of the B cells was sensitive also to the strength of the attention input. Attentional modulation of the nonpreferred B cells was sensitive to the strength of the inhibition caused by inhibitory edge (IE) cells as well. These dependencies allowed a sequential tuning of the parameters.

Quantitatively matching the observed border ownership modulation in the absence of attention constrains the strength of the positive feedback loop created by the mutual excitation between the cells coding for a proto-object and the cells coding for its local features. Subsequently, matching the experimentally observed strength of attentional modulation of B cells (10) constrains the strength of the attention input. We find that the maximal value of the attentional input to grouping cells needs to be 7% of the direct excitatory input that a grouping cell receives from a figure of its optimal size. To match the weak modulation of the opposite B cells, inhibition from IE cells needs to be just strong enough to prevent excitation by attention in the absence of edges.

The pattern and strength of the lateral inhibition are key to obtaining sharpening of the attention field and quasi-multiplicative



are irrelevant. Egly, Driver, and Rafal (3) observed a significant reaction time difference in a detection task between two positions equidistant from the focus of attention when one of the positions was on the same perceptual object (outlined rectangle) as the focus of attention. Because our model predicts a reshaping of the attention field caused by figures, we used it to predict neuronal firing rate responses in the Egly et al. task. We assumed that higher neuronal activation lowers reaction times and we found that the

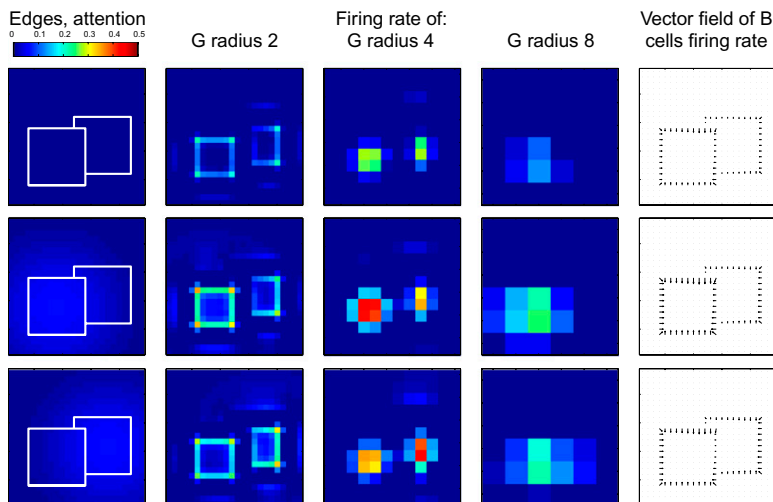


Fig. 3. Attention directed to one of two overlapping figures. First column: Sensory input and attentional modulation, same format as Fig. 2. Columns 2–4: Activity of grouping cells of different scales. Column 5: Vector field of border ownership cell activity where each border ownership cell is assigned a vector in the direction of the preferred figure side with a magnitude equal to its firing rate. At each location the vectors for all border ownership cells at this location are summed. In the absence of attention (*Top* row) or when attention is directed to the foreground (*Middle* row), the edge between the figures is correctly assigned to the foreground figure. If attention is directed toward the background figure (*Bottom* row), the border ownership signal of this edge is strongly reduced, to values consistent with experimental observations (10).

observed activities of the edge cells (target > cued > uncued) are, indeed, consistent with the observed reaction times (uncued > cued > target; Fig. 4). Importantly, these results are a genuine prediction: All parameter values were identical to those obtained from fitting the firing rates of the border ownership neurons; they were not retuned for predicting the Egly et al. data.

Kimchi et al. (11), studied stimulus-driven capture of attention by "objecthood." Several L-shapes of different orientations were presented and the task was to identify the color of one of them, which was precued to be the target. The elements could either be arranged randomly or be grouped such that four of them comprised the corners of a square, a perceptual object (Fig. 5A–C). It was irrelevant for the task whether the target shape was part of an object. Reaction times were fastest when the target was part of an object, slowest when the target was outside an object, and intermediate when no object was present. Using the same parameters as before, our model produced an activity enhancement of edge cells when the target was part of the object, consistent with the reaction time benefit. We did not, however, obtain a suppression when the target was outside the object (Fig. 5D). This empirically observed suppression was not reproduced by the model due to the absence of long-range inhibition between different G cells, which was not required by the constraints of the previous simulations. Competition between G cells occurred only in the presence of a common edge (via IE cells, Fig. 1). However, it is known that attention also involves competition between separate objects (e.g., refs 22–24). Accordingly, while keeping all other parameters constant, we added direct connections between G and inhibitory B (IB) cells (dashed lines in Fig. 1) that follow the same pattern as the excitatory G to B connections but with doubled radius. Through a weak but spatially broad activation of inhibitory border ownership cells (Fig. S4 D–I, column 6), grouping cells now compete with each other (Fig. S4D). This competition results in decreased activity for isolated edges (Fig. S3) and weak, broad suppression around objects. With this addition, edge modulation is weaker for items near objects (Fig. 5E) and our simulated results agree with all results observed by Kimchi et al. (11).

Discussion

In our model, attentional top-down influence is assumed to be simple: a broad Gaussian. However, the attentional modulation of the grouping cells, which pass down the attention signal to the early stages of the visual system, is no longer simple: It is focused on perceptual objects and tuned to their scale. The attentional modulation in the grouping cell layer is similar to a spotlight of attention that is shaped and positioned by bottom-up signals. This refinement of the attentional modulation results in rather even enhancement of contour signals irrespective of the size and

specific position of the attended object. The effective attentional spotlight is thus not imposed from central structures; these structures are unlikely to possess information about specific properties of the local features of objects. Instead, it is formed dynamically through the interplay of central structures (frontal and parietal cortex) that provide executive guidance of a general

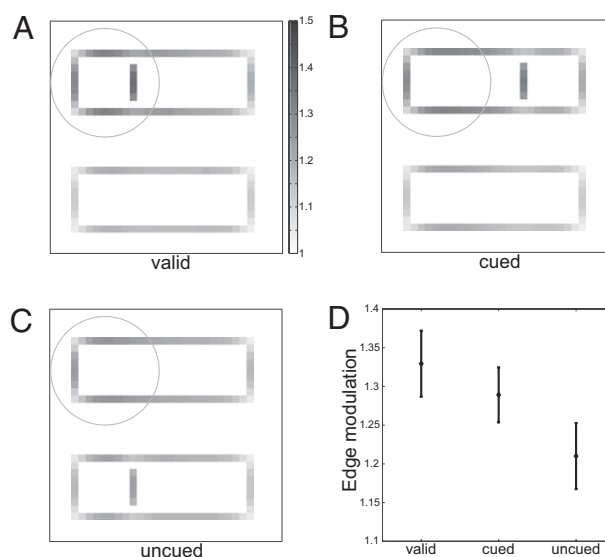


Fig. 4. Attention modulation of edge cells in a simulation of the Egly et al. (3) experiments. The outlines of two rectangles are presented. Top-down attention is directed toward the *Upper Left* corner. Attentional input is applied to grouping cells as a two-dimensional Gaussian (SD shown by circle). The target can be present at three positions relative to the cued location: (A) *valid* position, (B) *invalid* position but the target is in the *cued* rectangle, and (C) *invalid* position and the target is in the *uncued* rectangle. The distance between target and cue is the same in the invalid cases. The valid position produces the largest activity in the edge cells (A). Because in our model the attentional field is reshaped to match the figure, the activity of edge cells is higher if the target is in the cued rectangle (B) than if the target is as far away from the cue, but in the uncued rectangle (C). The depth of shading of every pixel represents the activity of the edge cells of preferred orientation at that location; see grayscale bar. (D) Activity of the edge cells representing the target for the valid, invalid cued, and invalid uncued conditions. The error bars represent the SD of the edge cells' activity. All parameters were identical to those obtained from fitting the responses of border ownership neurons. Relative neuronal activation is consistent with observed reaction times.

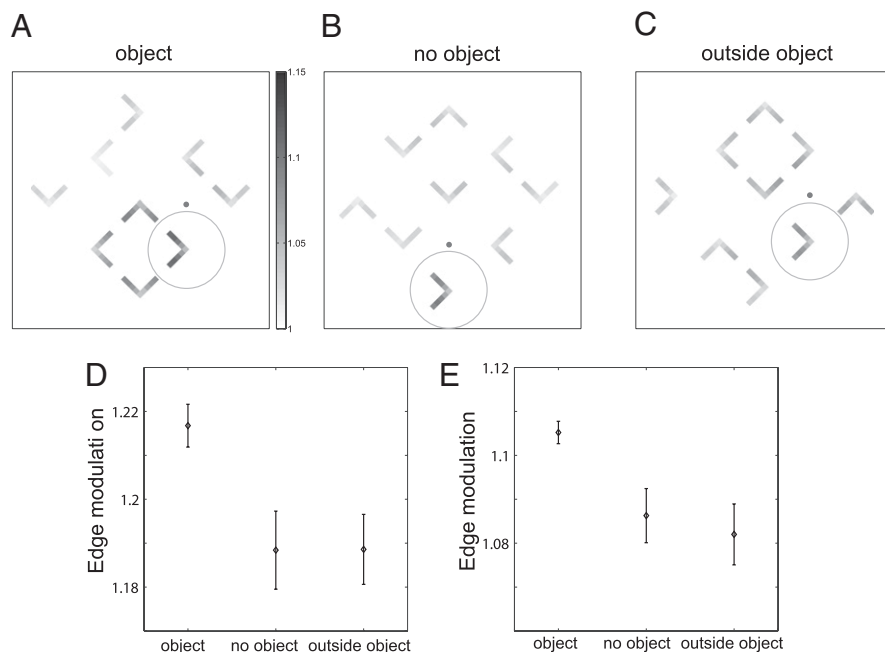


Fig. 5. Attentional modulation of edge cells in a simulation of the Kimchi et al. (11) experiments. An array of 3×3 L-shapes is presented and top-down attention (represented by a circle) is drawn to one of them (the target) by providing a cue (solid dot) and a direction relative to the cue (in all three examples it is downward). The task, identification of the color of the target, is independent of its orientation or of whether it is part of an object. The target can be part of an object (A), not part of an object when no object is present anywhere (B), or not part of an object that is present elsewhere (C). Shading of a pixel represents the activity of edge cells at the location of the pixel. (D) Mean edge modulation of the tips of the target (bars: SDs) for 200 simulations of random L-shape orientations in each category. The model and parameters are identical to those obtained from fitting the border ownership data. Consistent with the observed reaction times, being part of an object enhances edge activation. However, presence or absence of an object elsewhere did not make a consistent difference. (E) The same as D but after a weak and spatially broad inhibition between grouping cells is introduced (dashed lines in Fig. 1). The inhibition results in a spatially broad and nonspecific activity decrease of all cells outside an object. The smaller activity of the edge cells outside an object compared with the case in which no object is present is consistent with the observed reaction times (11). The additional inhibitory mechanism does not qualitatively change any of the previous results (Fig. S4).

nature, with peripheral (early cortical) areas that contribute the representation of detailed object and feature attributes.

The model presented here explains a large number of experimental observations. Attention has a quasi-multiplicative effect in early visual areas, in that the difference between the attended and the nonattended conditions is nonzero if the neuron is activated by the visual input and near zero for neurons not driven by the visual input. We show that this can be understood from the local circuitry that is characterized by specific excitatory connections and nonspecific inhibitory ones.

Although in the model attention acts at the level of grouping cells, it produces attentional modulation also for isolated lines that are not part of a larger object. This modulation is the result of weak activation of a large number of grouping cells. This activity represents potential proto-objects that would be consistent with the presented edge. Grouping cells diffusely inhibit each other (via IB cells), resulting in weaker attentional modulation when grouping cells in a large area are activated compared with focal activation. This inhibition explains why there is a general advantage in perception for features that are part of an extended object compared with isolated features, a phenomenon known as “object superiority effect” (25). Specifically, the model predicts that attentive enhancement should be stronger for the edge of a square than for an isolated edge (Fig. 2). An object superiority effect of this kind has been shown for attentional capture (11).

Another consequence of attentional input acting at the level of grouping cells is that attentional modulation is spread across the representation of an object: attention to one part of a perceptual object enhances processing of other parts of the same object (3, 5, 7). Our model predicts that such spreading of attentional modulation takes place in areas where local features of the object are represented, in our case, V2. This prediction can be directly tested in neurophysiological experiments. The model can

be extended to include additional cells coding for local features (e.g., local curvature) and additional searchable features for the grouping cells (e.g., color).

The model presented here quantitatively reproduces the side-of-figure-dependent influence of attention in border ownership neurons (10) and it qualitatively reproduces reaction time effects of task-irrelevant perceptual objects (3, 11). Important for its function are reciprocal connections between cells coding for the local features and those integrating them and inhibition between inconsistent representations. As a result, attentional modulation is focused to the scale of the object and repositioned to the center of the object.

Materials and Methods

Network Description. The network is described as a system of ordinary differential equations, with the dynamics of each neuron,

$$\tau f'(t) = -f + \left[\sum W \right]_+, \quad [1]$$

where f represents the neuron's activity level and τ its time constant, chosen as $\tau = 10^{-2}$ s for all neurons, W is the neuron's inputs, and $[\]_+$ means rectification. The stationary output of each neuron is thus linear in its inputs, except for a rectifying threshold that is set at 0.

The sensory input to the network is an edge map, similar to the model described by ref. 13. Perceptual organization over large spatial scales is established through specialized grouping cells that integrate over space and provide context information to border ownership selective cells (B cells, Fig. 1). A second input to the model is the top-down attention input, which is assumed to act on grouping cells. The attention input is purely spatial and, on the basis of psychophysical findings (26), it is considered to have lower spatial resolution than the sensory input. Excluding a set of simulations presented in Fig. S2, attentional input is considered to have a SD eight times the size of a border ownership cell receptive field (the median receptive field size for border ownership cells is 0.7° at 2° eccentricity) (12). The strength of the

attention input is a tuned parameter; the electrophysiologically measured attentional modulation of B neurons (preferred vs. antipreferred side of figure) (10) is obtained for a relatively weak value, 0.07 of the strength of the sensory input.

At the level of V2, E cells activate two border ownership (B)-cell populations with opposing border ownership preferences. Parallel to its preferred orientation, a border ownership cell activates IE cells and B cells of the same preferred orientation. IE cells inhibit all neighboring B and IE cells nonspecifically. All these connections are assumed to have a Gaussian distribution pattern, with a SD twice the receptive field size. Orthogonal to the preferred orientation, these connections create a "Mexican hat" profile. The strength of the nonspecific lateral inhibition was tuned to match the measured attention modulation of B neurons. The strength of the specific lateral excitation was tuned to match the border ownership modulation in the absence of attention (10).

B cells activate grouping cells of multiple scales whose integration fields are on the border ownership cells' preferred figure side. The connection patterns are presented in Fig. S1 and similar to those used in previous models of perceptual organization (13). The G cells are sensitive to cocircular arrangement of edges (cf. ref. 16), which combines features of the Gestalt laws of good continuation, convexity of contour, and compact shape (27, 28). The inputs to G cells of different preferred sizes are scaled such that a straight line produces equal excitation in G cells of different scales. Border ownership cells also excite IB cells that inhibit grouping cells whose activation is inconsistent with a figure edge in the receptive field of the border ownership cell. Grouping cells have reciprocal connections with border ownership cells, but with a different scaling: Each border ownership cell is equally affected by grouping cells of different spatial scales. The grouping cells also connect with the same pattern and with equal strength to all of the edge cells that coactivate with the border ownership cells that they receive input from. The feedback from each G cell is reciprocal, thus activating the same neurons that led to its activation and also the neurons that are coactivated with it. The scaling of these connections was chosen such that images of different scales are processed in a similar manner. The strength of the excitatory feedback loop from border ownership cells to grouping cells and back was tuned to match the border ownership modulation in the absence of attention (10).

In an extension of the model (see discussion of Fig. 5), the feedback from the grouping cells additionally activates the inhibitory border ownership neurons, with a connection pattern similar to the feedback to excitatory border ownership neurons and the same connection strengths but twice their spatial spread.

To constrain the model parameters we used three dimensionless quantities that were determined from experimental data (10): the border ownership modulation index when attention is directed away from the object at the receptive field and the attention modulation indexes for the border ownership cells in preferred and antipreferred conditions (see ref. 10 for definitions). We performed a global sensitivity analysis on five synaptic weight constants that are not scaling parameters (Table S2). The strength of the

excitatory feedback loop between B and G cells and the strength of the excitatory lateral connections between B cells were chosen such that the correct border ownership modulation was obtained. After setting these parameters, the attention input was tuned such that the experimentally observed value of the attention modulation of the preferred border ownership cells is reproduced. The attention modulation of the nonpreferred border ownership cell was used to tune the strength of the lateral inhibition produced by IE cells. The obtained value also reproduces the observed minimal effect of attention in the absence of figures (Fig. S2). From the available data, we could not determine how specific the lateral inhibition produced by IE cells is, and we assumed this inhibition to be nonspecific. See *SI Materials and Methods* for a complete specification of the network.

Effects of Local Circuitry. We consider first attentional modulation in the absence of edge inputs (Fig. 2 A and C) and we always assume that top-down attentional influence at the G-cell level is spatially broad (Fig. 1). The grouping cells provide excitatory input to a large population of border ownership (B) and IE neurons. In the absence of sensory input, the direct excitatory input from grouping cells to a B cell is compensated by the sum of all inhibitory input from the IE cells; attention therefore does not modify B-cell activities that do not receive sensory input. We note that if top-down attentional influence at the G-cell level is highly spatially localized, attention can produce a small activity increase at the B-cell level (Fig. S2), consistent with experimental observations in which the attentional focus was spatially sharp (19).

If a sensory stimulus, e.g., an oriented line, is presented in the absence of attention (Fig. S3B, Upper Right quadrant) it provides, via edge cells, bottom-up input to those B cells that are selective for its orientation. Along the axis orthogonal to the preferred orientation of the cells, due to the Mexican hat pattern of excitation/inhibition, a localized sensory stimulus produces a localized activation of both B and IE cells, with the surrounding cells being inhibited below their threshold. Along the axis parallel to the preferred orientation, the signals related to the sensory stimulus propagate, and their analysis is discussed in *Results*.

When a stimulus is attended, both sensory and attentional inputs are present (Fig. S3B, Lower Left quadrant, with the activity difference between the attended and the unattended stimulus shown in Fig. S3C). B cells far from the sensory input receive balanced excitation and inhibition from the attention input, as described above, and thus do not change their firing rates. At the location of the stimulus, because nearby IE cells are inhibited below their threshold by the stimulus, B cells receive excitation from the attention input that is no longer balanced by inhibition. Attentional modulation thus affects the activity of stimulated neurons but not that of nonstimulated neurons.

ACKNOWLEDGMENTS. We thank Anne Martin for a critical reading of the manuscript. This work was supported by National Institutes of Health Grants 5R01EY016281-02 and 01EY02966 and Office of Naval Research Grant N000141010278.

- Mountcastle VB (1995) The parietal system and some higher brain functions. *Cereb Cortex* 5:377–390.
- Schall JD (1997) Visuomotor areas of the frontal lobe. *Cerebral Cortex*, eds Rockland KS, Peters A, Kaas J (Plenum, New York), Vol 4, pp 527–638.
- Egley R, Driver J, Rafal RD (1994) Shifting visual attention between objects and locations: Evidence from normal and parietal lesion subjects. *J Exp Psychol Gen* 123: 161–177.
- Cave KR, Bichot NP (1999) Visuospatial attention: Beyond a spotlight model. *Psychon Bull Rev* 6:204–223.
- Scholl BJ (2001) Objects and attention: The state of the art. *Cognition* 80:1–46.
- Matsukura M, Vecera SP (2006) The return of object-based attention: Selection of multiple-region objects. *Percept Psychophys* 68:1163–1175.
- He ZJ, Nakayama K (1995) Visual attention to surfaces in three-dimensional space. *Proc Natl Acad Sci USA* 92:11155–11159.
- Roelfsema PR, Lamme VAF, Spekreijse H (1998) Object-based attention in the primary visual cortex of the macaque monkey. *Nature* 395:376–381.
- Ito M, Gilbert CD (1999) Attention modulates contextual influences in the primary visual cortex of alert monkeys. *Neuron* 22:593–604.
- Qiu FT, Sugihara T, von der Heydt R (2007) Figure-ground mechanisms provide structure for selective attention. *Nat Neurosci* 10:1492–1499.
- Kimchi R, Yeshurun Y, Cohen-Savransky A (2007) Automatic, stimulus-driven attentional capture by objecthood. *Psychon Bull Rev* 14:166–172.
- Zhou H, Friedman HS, von der Heydt R (2000) Coding of border ownership in monkey visual cortex. *J Neurosci* 20:6594–6611.
- Craft E, Schütze H, Niebur H, von der Heydt R (2007) A neural model of figure-ground organization. *J Neurophysiol* 97:4310–4326.
- Reynolds JH, Heeger DJ (2009) The normalization model of attention. *Neuron* 61: 168–185.
- Tsotsos JK, et al. (1995) Modelling visual attention via selective tuning. *Artif Intell* 78: 507–545.
- Sigman M, Cecchi GA, Gilbert CD, Magnasco MO (2001) On a common circle: Natural scenes and Gestalt rules. *Proc Natl Acad Sci USA* 98:1935–1940.
- Chow CC, Jin DZ, Treves A (2002) Is the world full of circles? *J Vis* 2:571–576.
- Reynolds JH, Chelazzi L (2004) Attentional modulation of visual processing. *Annu Rev Neurosci* 27:611–647.
- Luck SJ, Chelazzi L, Hillyard SA, Desimone R (1997) Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *J Neurophysiol* 77:24–42.
- Ghose GM, Maunsell JHR (2002) Attentional modulation in visual cortex depends on task timing. *Nature* 419:616–620.
- Ziehnt T, Tomlin A (2009) GUI-HDMR: A software tool for global sensitivity analysis of complex models. *Environ Model Softw* 24:775–785.
- Desimone R (1998) Visual attention mediated by biased competition in extrastriate visual cortex. *Philos Trans R Soc Lond B Biol Sci* 353:1245–1255.
- Hopf JM, et al. (2006) Direct neurophysiological evidence for spatial suppression surrounding the focus of attention in vision. *Proc Natl Acad Sci USA* 103:1053–1058.
- Boehler CN, Tsotsos JK, Schoenfeld MA, Heinze HJ, Hopf JM (2009) The center-surround profile of the focus of attention arises from recurrent processing in visual cortex. *Cereb Cortex* 19:982–991.
- Williams A, Weisstein N (1978) Line segments are perceived better in a coherent context than alone: An object-line effect in visual perception. *Mem Cognit* 6:85–90.
- Intriligator J, Cavanagh P (2001) The spatial resolution of visual attention. *Cognit Psychol* 43:171–216.
- Koffka K (1935) *Principles of Gestalt psychology* (Harcourt-Brace, New York).
- Kanizsa G (1979) *Organization in Vision: Essays on Gestalt Perception* (Praeger, New York).