Pergamon

PII: S0042-6989(97)00197-1

Vision Res., Vol. 38, No. 5, pp. 719–741, 1998
© 1998 Elsevier Science Ltd. All rights reserved
Printed in Great Britain
0042-6989/98 $19.00 + 0.00

Extraction of Perceptually Salient Contours by Striate Cortical Networks

SHIH-CHENG YEN,*† LEIF H. FINKEL†

Received 1 November 1996; in final form 16 June 1997

We present a cortical-based model for computing the perceptual salience of contours embedded in noisy images. It has been suggested that horizontal intra-cortical connections in primary visual cortex may modulate contrast detection thresholds and pre-attentive "pop-out". In our model, horizontal connections mediate context-dependent facilitatory and inhibitory interactions among oriented cells. Strongly facilitated cells undergo temporal synchronization; and perceptual salience is determined by the level of synchronized activity. The model accounts for a range of reported psychophysical and physiological effects of contour salience. In particular, the model proposes that intrinsic properties of synchronization account for the increased salience of smooth, closed contours. Application of the model to real images is demonstrated. © **1998 Elsevier Science Ltd. All rights reserved.**

Closure Contour Salience Striate cortex Synchronization

INTRODUCTION

The idea that perceptual salience depends upon context originated with the Gestalt psychologists. The Gestalt laws describe the influence of global context on the perception of local features. Elements tend to be perceptually grouped and made salient if they are close to each other (proximity), similar to one another (similarity), form a continuous contour (good continuation), form a closed contour (closure), or move together in the same direction (common fate) (Rock & Palmer, 1990). This tradition has motivated a number of recent psychophysical and physiological studies which have investigated context-dependent modification of contrast sensitivity thresholds, pre-attentive "pop-out", and the extraction of smooth contours and coherent motion trajectories from cluttered backgrounds (Field, Hayes, & Hess, 1993; Kovács & Julesz, 1993, 1994; Pettet, McKee, & Grzywacz, 1996; Newsome, Britten, & Movshon, 1989; Watamaniuk & Sekuler, 1992). Recent anatomical and physiological studies (Rockland & Lund, 1982; Rockland & Lund, 1983; Kapadia, Ito, Gilbert, & Westheimer, 1995; Singer & Gray, 1995; Gilbert, Das, Ito, Kapadia & Westheimer, 1996; Fitzpatrick, 1996) have suggested mechanisms by which these context-dependent effects may be carried out in striate cortex. It has been proposed that long-range horizontal connections may provide the means of modulating cell responses based on the structure of the distant surround (Gilbert *et al.*, 1996). Evidence suggests that the same cortical cells which are interconnected by long-range horizontal connections are also involved in synchronization (see Singer & Gray, 1995 for a review; Gray & McCormick, 1996). We test the ability of a model composed of units modeled after striate cortical cells, embedded in an anatomical network of long-range connections, and capable of temporal synchronization, to account for reported psychophysical results on contour salience.

The salience of a stimulus depends, in part, on its contrast relative to that of surrounding stimuli. However, contrast sensitivity can be modulated by the structure of stimulation in the surround. Polat and Sagi (1993, 1994) measured changes in detection thresholds for a low contrast Gabor patch when two high contrast Gabor patches of the same orientation were placed on either side of it. When the flanking stimuli are positioned at distances beyond 2λ (where λ is the standard deviation of the gaussian window of the Gabor function), local contrast sensitivity is increased. This facilitatory effect peaks at a separation of 3λ and remains above baseline out to 12λ. At the lowest spatial frequency tested ($\lambda = 0.3$ deg), this corresponds to a maximal range of 3.6 deg for the facilitation. At separations less than 2λ, the flanking stimuli decrease contrast sensitivity, presumably, as suggested by Polat & Sagi (1993), as a result of encroachment on the inhibitory surround of the classical receptive field (CRF) [DeAngelis, Freeman, and Ohzawa (1994) provide supporting physiological evidence for this effect.] All effects scale with λ. When the three Gabor patches are oriented orthogonal to the orientation axis, similar effects are obtained, although

*To whom all correspondence should be addressed [*Tel*: +1 215 898 0822; *Fax*: +1 215 573 2071; *Email* syen@seas.upenn.edu].

†Department of Bioengineering, University of Pennsylvania, 3320 Smith Walk, Philadelphia, PA 19104, U.S.A.

extending over a shorter spatial range (returning to baseline around 3λ, which is 0.225 deg for $\lambda =$ 0.075 deg).

In more complex images, these same spatial interactions between oriented elements are thought to mediate perceptual "pop-out". Using a 2AFC paradigm, Field *et al.* (1993) tested the ability of subjects to detect a target embedded in a large field of randomly oriented Gabor function elements. The target was composed of a small set of the oriented Gabor functions aligned to form a smooth contour. They found that target detectability depended on several factors, most importantly, the relative orientations of the elements. Targets could be detected if the orientations of adjacent contour elements differed by as much as 60 deg. A similar but weaker effect on detectability was observed when the elements were oriented perpendicular to the contour. All effects were robust over a range of element densities and presentation times.

Using a similar experimental paradigm, Kovács and Julesz (1993, 1994) demonstrated that closed contours are more salient than open contours. They reported that on a closed contour, the maximum inter-element separation, Δ_c, (defined to be at 75% correct performance for detection), is larger than the maximum detectable spacing for elements on an open contour, Δ_o.* They also found a striking difference in the effect of adding additional contour elements at threshold separations to closed versus open contours. The saliency of open contours increased monotonically as elements spaced at Δ_o were added; but for elements spaced at Δ_c, the saliency remained very low and dramatically increased only when the contour was closed. Pettet *et al.* (1996) further showed that the smoothness of the contour plays a large role in determining saliency. Contours with sharp curves have decreased salience whether they are open or closed.

More recent experiments by Kovács, Polat, and Norcia (1996) suggest that the critical factor in detecting a contour embedded in noise is the *relative* distance between elements of the contour and elements of the background. If target and background elements are indistinguishable, e.g., identical small circular dots, then a target contour can only be detected when its elements are spaced more closely than the spacing between background elements (Kovács *et al.*, 1996). However, when the elements are oriented, they can be separated up to 1.5-times the average separation of background (oriented) elements, and the contour can still be detected. Kovács and colleagues found that at threshold, the ratio between the separation of background elements and the separation of contour elements, φ, was approximately 0.65. These experiments suggest that background elements contribute a certain amount of "noise", which sets a lower limit for signal detection. Within some broad anatomical limit, the distance over which elements can be grouped into contours is therefore totally context-dependent.

The orientation-specific nature of these effects, as well as the observation that contours can be integrated stereoptically across depth planes (Hess & Field, 1995), suggest that these effects are mediated at the cortical level. Given that the influence of surrounding elements extends over several degrees of visual angle, it has been suggested that long-range horizontal connections in striate cortex may underlie these effects (Field *et al.*, 1993; Kovács & Julesz, 1993, 1994; Kapadia *et al.*, 1995; Gilbert *et al.*, 1996).† Horizontal connections have been observed to spread over 5–8 mm of cortex (Gilbert & Wiesel, 1979, 1983, 1989; Rockland & Lund, 1982, 1983; Martin & Whitteridge, 1984; Ts'o, Gilbert, & Wiesel, 1986; Ts'o & Gilbert, 1988). Considering that there is no overlap in the receptive fields for cells separated by 1.5 mm (Hubel & Wiesel, 1974), this represents connections between cells separated by distances 4–5-times their receptive field sizes. Depending on visual eccentricity, receptive field sizes vary from less than 1 deg to several degrees (Gilbert & Wiesel, 1979, 1983; DeAngelis *et al.*, 1994). This suggests that connections between cells separated by more than 10 deg are highly plausible over most of primary visual cortex. Indeed, interactions spanning 15 deg of visual cortex have been reported in the literature (Gilbert & Wiesel, 1979; Fitzpatrick, 1996). Thus, the range of these connections is well within the range of the interactions described in the psychophysical experiments (Polat & Sagi: 0.225–3.6 deg, Field *et al.*: 0.25–0.9 deg, Kovács & Julesz: 0.4–0.72 deg). Cross-correlation studies (Toyama, Kimura, & Tanaka, 1981a,b; Gilbert & Wiesel, 1989; Hata, Tsumoto, Sato, Hagihara, & Tamura, 1993) suggest that these horizontal connections are primarily between cells of like orientation tuning. They also match the topography of axonal projections recently identified using intracellular labels (Fitzpatrick, 1996).

Several physiological studies have reported changes in cell activity that may be mediated by these horizontal connections. Nelson and Frost (1985) found that when recording from striate cortex of anesthetized cats, the response to an optimally oriented bar presented inside the CRF could be suppressed by a drifting grating outside the CRF. They noted that the response showed periodic relief from inhibition at those times when collinear regions outside the receptive field were stimulated. These effects extended out to 5 deg of visual space from the center of the receptive field. They proposed that the flanking regions were contributing excitatory inputs that periodically countered the inhibitory inputs from the surround. Kapadia *et al.* (1995) systematically investigated these effects in awake behaving macaque monkeys by recording the responses of complex cells in the superficial layers of V1 to elongated bars outside CRF. They found significant facilitation of cell responses to a low contrast bar when a collinear high contrast bar was placed outside the CRF. This facilitation varied as a function of

*If the target elements differ in another feature, such as color, they can be spaced farther apart.

†These connections have also been implicated in texture segmentation (Knierim & Van Essen, 1992) and context-dependent receptive field reorganization (Pettet & Gilbert, 1992).

separation, alignment, and relative orientation between the two bars (see Fig. 4). They found a close correspondence between these physiological effects and psychophysical responses in human observers.

In striate cortex, 80% of the long-distance connections synapse onto excitatory cells and the other 20% target inhibitory interneurons (McGuire, Gilbert, Rivlin, & Wiesel, 1991). Using optical recordings, Weliky, Kandler, Fitzpatrick, & Katz (1995) observed short-latency excitation followed by a longer latency, presumably disynaptic, inhibition following stimulation of cells separated by up to 1300 μm (Weliky *et al.*, 1995).* Evidence suggests that this inhibition is more broadly tuned in orientation compared with excitatory inputs (Nelson & Frost, 1978, 1985; DeAngelis, Robson, Ohzawa, & Freeman, 1992; DeAngelis *et al.*, 1994). The results also suggest that the magnitude of the inhibition is similar to that of long-distance excitation, but inhibition arises from regions that are more isotropically distributed with respect to the CRF. Kapadia *et al.* (1995) also reported inhibitory effects on a target bar surrounded by randomly oriented bars in the distant surround. This inhibition was converted to strong facilitation with the addition of collinear bars in the surround. Similar inhibitory effects from outside the classical receptive field have been observed with other experimental paradigms (Knierim & Van Essen, 1992; Sillito, Grieve, Jones, Cudeiro, & Davis, 1995).

These same long-distance horizontal connections have also been implicated in the context-dependent temporal synchronization of cells with non-overlapping receptive fields, sometimes even separated by several millimeters of cortex (Gray, König, Engel, & Singer, 1989; Engel, König, Gray, & Singer, 1990; Engel, Kreiter, König, & Singer, 1991; Livingstone, 1996). These investigators and others have suggested that temporal synchronization may be used to group related features together across the visual field, as well as in other domains of sensory input. The mechanism underlying the generation of these synchronized responses has been debated, but recent evidence suggests that a certain subtype of cortical cells in the supragranular layers of striate cortex, termed "chattering" cells, may participate in the synchronization of cortical responses (Gray & McCormick, 1996).

THE MODEL

These accumulated studies suggest that perceptual salience may arise from temporal synchronization of cortical cells, which in turn depends upon the spatial relationship between stimulus elements as mediated by long-range horizontal connections. Horizontal connections can carry out modulatory interactions between cells, but temporal mechanisms are required to represent global context, such as closure. Our model is composed of an array of units modeled after cortical cells, which extract orientation and spatial frequency information at each location in the scene. Units are interconnected by long-range horizontal connections which provide both facilitation and inhibition. Only cells which receive both local input (to the CRF) and support from stimuli in the surround can be facilitated. A longer-latency inhibition suppresses the responses of weakly facilitated cells, allowing only the strongly facilitated cells to remain active. We assume that facilitation "promotes" cells into a bursting state, similar to the "chattering" behavior observed in Gray and McCormick (1996). Cells in the bursting mode can then undergo synchronization. In the model, facilitated cells are modeled as neural oscillators (Kopell & Ermentrout, 1986; Baldi & Meir, 1990; Somers & Kopell, 1993; Terman & Wang, 1995). The strength of the coupling between oscillators determines which cells synchronize with one another, and also determines the degree of synchronization achieved among the cells. We propose that perceptual salience is directly related to the level of synchronized activity among a group of cells. Based on the coherence of the synchronization, the network can generate an estimate of the salience of all contours in the image. The network's estimate be compared with human psychophysics using the paradigms in the experiments discussed above.

Model cells

Quadrature pair linear steerable filter pyramids (Freeman & Adelson, 1991) are used to efficiently represent the responses of oriented cells in primary visual cortex. Steerable filters allow the energy at any orientation and spatial frequency to be efficiently calculated from the responses of a set of basis filters. Steerable filters also allow the direct determination of the dominant orientation (i.e., the orientation preference of the maximally responding cell) at each spatial frequency and position (Freeman & Adelson, 1991). Steerable filters are used here for computational efficiency—they differ in aspect ratio and bandwidth from cortical cells. Nonetheless, they provide an adequate model of local orientation extraction for our purposes and their computational advantage will become significant in the sections described below.

We used the G2 (second derivative of a gaussian) and H2 (Hilbert transform of the G2 filter) filters, which have been shown to be reasonable approximations to the shape of receptive fields in V1 (Young, 1987; Young & Lesperance, 1993). The squared responses of the cells at the dominant orientation at each position in the image is illustrated in Fig. 11(b).

Facilitation

Units are interconnected by long-range horizontal connections. The sign, magnitude and time course of the synaptic interactions depend upon the position and orientation of the target cell, creating separate spatial zones of excitation and inhibition. The connection field is shown in Fig. 1(b). Excitatory connections are confined to two regions, one flaring out along the axis of orientation of the cell (co-axial), and another confined to a narrow zone extending orthogonally to the axis of

*Direct inhibitory connections spread only 200–600 μm (Kisvárday & Eysel, 1992).

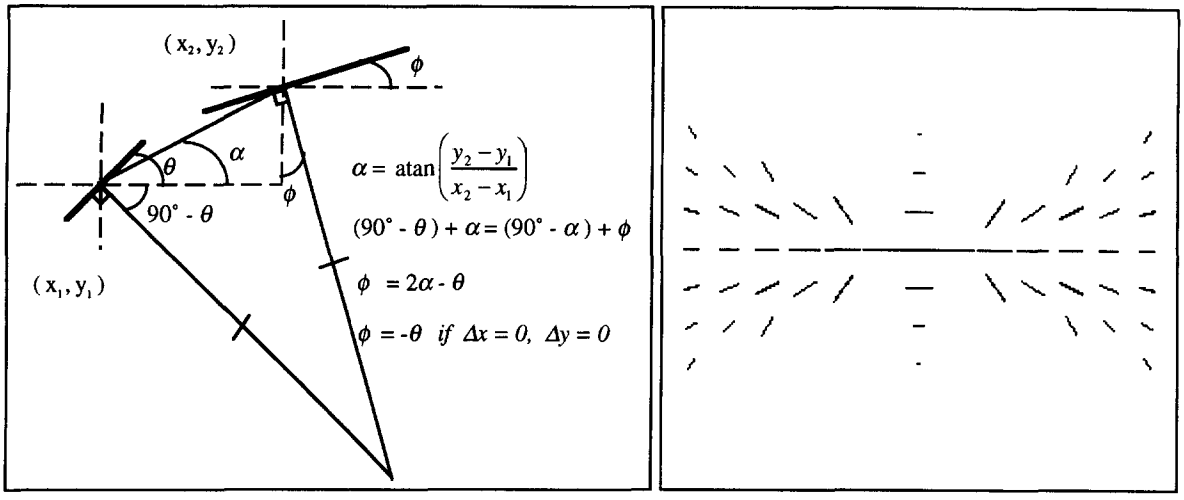

FIGURE 1. (a) Computation of the "preferred" orientation based on the co-circularity constraint. θ and ϕ are the tangents to the circle passing through (x_1, y_1) and (x_2, y_2). For a unit at (x_1, y_1) of orientation θ, the preferred orientation at (x_2, y_2) is ϕ. (b) Connectivity pattern of a horizontally oriented cell (located at the center of the image). The orientation of the lines represents the "preferred" orientation, while the length of the lines indicates connection strength.

orientation (trans-axial). The co-axial connections are similar to the "association field" proposed by Field *et al.* (1993), and are generated by a simple equation (see Appendix) modified from Parent and Zucker's (1989) "co-circular" connection scheme. For a cell of orientation θ_A at location "A", there is a "preferred" orientation at location "B", ϕ_B, given by the tangent to the unique circle which passes through both "A" and "B", and whose tangent at "A" agrees with the local orientation, θ_A, at "A" [see Fig. 1(a)].* If the local orientation activity distribution at "B" peaks at ϕ_B, the cell with orientation θ_A at "A" will be strongly facilitated. As the local orientation at "B" deviates from ϕ_B, the degree of facilitation decreases. The "preferred" orientation at "B" can thus be thought of as providing "support" for the orientation, θ_A, at "A". Connection weights also decrease for positions with increasing angular deviation from the orientation axis of the cell, reflecting a preference for straight lines and lines of low curvatures [see Field *et al.*, 1993; also Fig. 9(a)]. The connection weights also decrease with increasing distance (see Appendix for details).

A second set of trans-axial excitatory connections extends orthogonally from the orientation axis of the cell. Again, the strongest connections are to units at nearby positions with orientations parallel to that of the cell. This set of connections is more spatially focused, with the weights falling off in angle much more quickly than the first set of connections [Field *et al.*, 1993; also shown in Fig. 9(b)]. There is anatomical evidence consistent with the existence of these orthogonal connections (Rockland & Lund, 1982, 1983; Mitchison & Crick, 1982; Lund, Fitzpatrick, & Humphrey, 1985; Fitzpatrick, 1996). Psychophysical results (Field *et al.*, 1993; Polat & Sagi, 1994) also demonstrate facilitatory effects for contour elements arranged in a parallel fashion. These orthogonal connections will play a role in accounting for a number of the experimental results. The connection pattern for a cell preferring horizontal orientations is shown in Fig. 1(b).

The model has several additional properties that are motivated by experimental findings. As with the physiological results of Nelson & Frost (1978, 1985), and Kapadia *et al.* (1995), the facilitatory connections are only effective for cells receiving direct supra-threshold stimulation to the CRF. This prevents cells with weak, or no input from the visual field from being facilitated by the cells around it. This is also consistent with the results of von der Heydt and Peterhans (1989) showing that responses to certain types of illusory contours are only observed in V2 of the macaque and not in V1. Each cell receives a large number of inputs which may vary in magnitude as the activity of surrounding cells change with stimulus contrast. The facilitatory inputs are, therefore, normalized so that the performance of the model is largely independent of the average contrast of the stimuli (see Appendix). Finally, the two sets of connections compete, with the co-axial connections inactivating the trans-axial connections when the co-axial facilitation is stronger, and vice versa. This is broadly consistent with the Gestalt laws of grouping (e.g. the influence of proximity on binding of dot arrays, Koffka, 1935).

Inhibition

A major function of inhibition in the model is to distinguish signal from noise based on the degree of facilitation. Since elements in the background are randomly positioned and oriented, stray background elements may be optimally oriented to be facilitated by

*The "co-circular" connection pattern uses a circle as a model for all the possible smooth curves that could pass through both "A" and "B" and is not a circle or curvature detector.

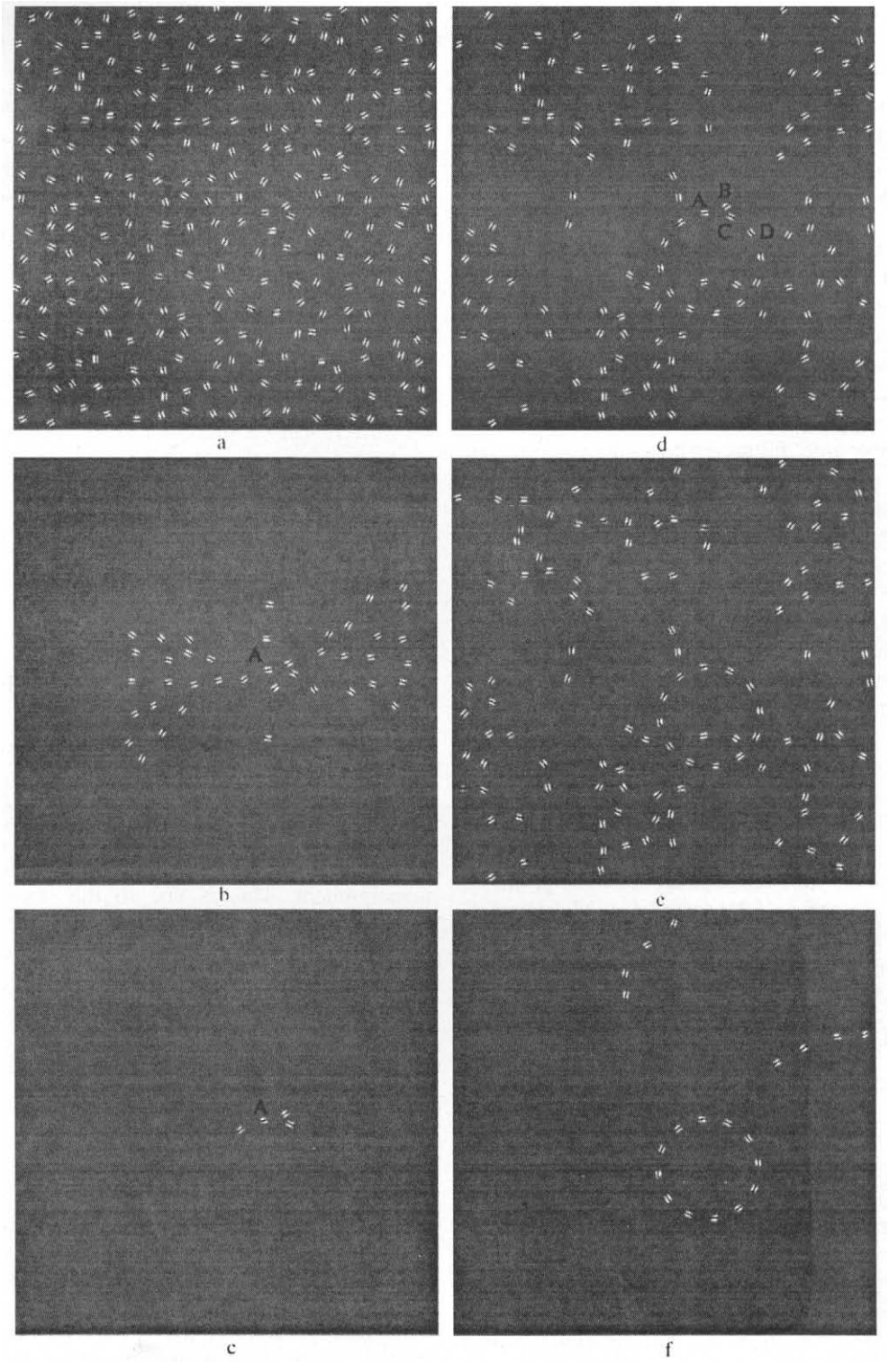

FIGURE 2. (a) Stimulus with embedded circular contour. (b) "Preferred" orientations for stimulus positions around element "A". (c) Elements with super-threshold inputs to element "A". (d) Plot of all elements that are facilitated. (e) Facilitated elements after inhibition stage. (f) Contours with salience values of at least 4 units.

elements on a contour. Inhibition suppresses the response of these "distractor" elements and prevents them from becoming attached to the "target" contour. Inhibition follows after facilitation and only the elements that are strongly facilitated persist. This allows the distractor elements in the background to be separated from the elements on the chain. The simulation results in Fig. 2 illustrate the role of inhibition in the model. The input stimulus, a closed contour made up of 12 elements, is displayed in Fig. 2(a). Figure 2(b) displays the "preferred" orientations for the positions around the element. The elements whose orientations are close to the "preferred" orientations provide strong facilitatory inputs to the post-synaptic cell. Figure 2(c) shows the elements

whose inputs to the element marked "A" are above the signal-to-noise ratio (see Appendix). The elements in the scene that are being facilitated are shown in Fig. 2(d). The element marked "B" represents a distractor element in the background that is being facilitated by the element on the contour, marked "A". The element "C" receives facilitation from "A" as well as from "D", which allows it to then inhibit "B". Figure 2(e) shows the result of this inhibition over the entire scene. Elements that are strongly facilitated are not affected by the inhibition. Figure 2(f) shows some of the more salient contours that have been extracted from the stimulus. Facilitation and inhibition operate in parallel over the scene and extract not only the target contour, but also other less salient contours in the scene that would be observable if given extended viewing time. Figure 3 shows the performance of the model on stimuli containing an open chain, a closed chain and a straight line containing elements oriented orthogonal to the contour.

Synchronization and salience

In the model, the cells that are strongly facilitated are assumed to enter a "bursting" mode, which allows them to synchronize with other similarly bursting cells. We use a simple, descriptive mechanism for temporal synchronization. Cells that enter the bursting mode are modeled as homogeneous coupled neural oscillators with a common fundamental frequency but different phases (Li & Hopfield, 1989; Kopell & Ermentrout, 1986; Kammen, Holmes, & Koch, 1989; Baldi & Meir, 1990; König & Schillen, 1991; Schillen & König, 1991; Grossberg & Somers, 1991; von der Malsburg & Buhmann, 1992; Terman & Wang, 1995). The phase of each oscillator is modulated by the phase of the oscillators to which it is coupled. Oscillators are coupled only to other oscillators with which they have strong, reciprocal, facilitated connections (see Appendix). A set of coupled oscillators together form a contour. Since oscillators on different contours are not generally interconnected, this allows each contour in the scene to synchronize independently. We use neural oscillators only as a simple functional means of computing synchronization and make no assumption regarding their possible functional role in cortex.

We propose that the saliency of a contour can only be computed when all the oscillators on the contour are synchronized. The salience of the contour is then represented by the sum of the activities of all the synchronized elements. The longer the chain of synchronized elements, the more perceptually salient it is. Synchronization occurs in parallel over the whole scene and the longest synchronized chain in the scene is identified as being the most salient, and the network selects it as its output.

RESULTS

All simulations were conducted with the same parameter set, which was chosen based on optimizing results for the stimulus in Fig. 2. Details of how the simulations were carried out are described in Appendix.

Experiment 1: co-axial connection pattern

Simulation of the psychophysical experiments of Kapadia *et al.* (1995) provides a test of the model's co-axial pattern of excitatory horizontal connections (see Methods section in Appendix). The response of a unit was determined as a function of the position and orientation preference of the pre-synaptic cell. Figure 4 shows the effect of varying the co-axial distance, off-axis misalignment, and angular orientation of the pre-synaptic cell. The pre-synaptic cell is assumed to have constant activity, and so the changes in facilitation are due only to the differences in the connection weights between the post-synaptic cell and pre-synaptic cells at different positions and orientations. Figure 4(a) shows the effect of the decrease in facilitatory connection weights with increasing co-axial separation between the pre- and post-synaptic cells. The results show a decreased post-synaptic response in qualitative agreement with the data. The results are also in qualitative agreement with those of Polat and Sagi (1993). Our simulations show an over-estimation of the activation for closely spaced elements [Fig. 4(a)] that would be compensated by local, short-range inhibition, which is not included in this simulation. Such inhibition would correspond to Polat and Sagi's (1993, 1994) observation of increased contrast detection thresholds at small separations.

The input also decreases as a function of off-axis misalignment. This is due to the smaller connection weights from off-axis locations, where the "preferred" orientation differs from that of the pre-synaptic cell. In the second simulation [Fig. 4(b)], the agreement between the model and the data on the effects of off-axial misalignment is good except at large lateral offsets. At large offsets the results of Kapadia *et al.* (1995) show that the influence of the surround becomes largely inhibitory. In the model, the facilitation at large lateral offsets becomes very weak and would be overwhelmed by the longer latency inhibition that is also not included in this particular simulation.

As the orientation of the pre-synaptic cell increasingly deviates from the "preferred" orientation, the weights decrease correspondingly. The results of the third simulation [Fig. 4(c)] show qualitative agreement with the data as the orientation of the pre-synaptic cell rotates over 70 deg. The co-axial connections are thus consistent with the reported psychophysical effects due to the structure of the surround.

Experiment 2: contrast sensitivity modulations

We believe the trans-axial connections in our model may be responsible for a surprising result of Kovács and Julesz (1993, 1994). They measured changes in contrast sensitivity to a low contrast Gabor target placed at various locations inside and outside a circular [Fig. 5(b)] and elliptical [Fig. 5(c)] contour. The contour itself was formed from aligned Gabor patches. They found a sharp

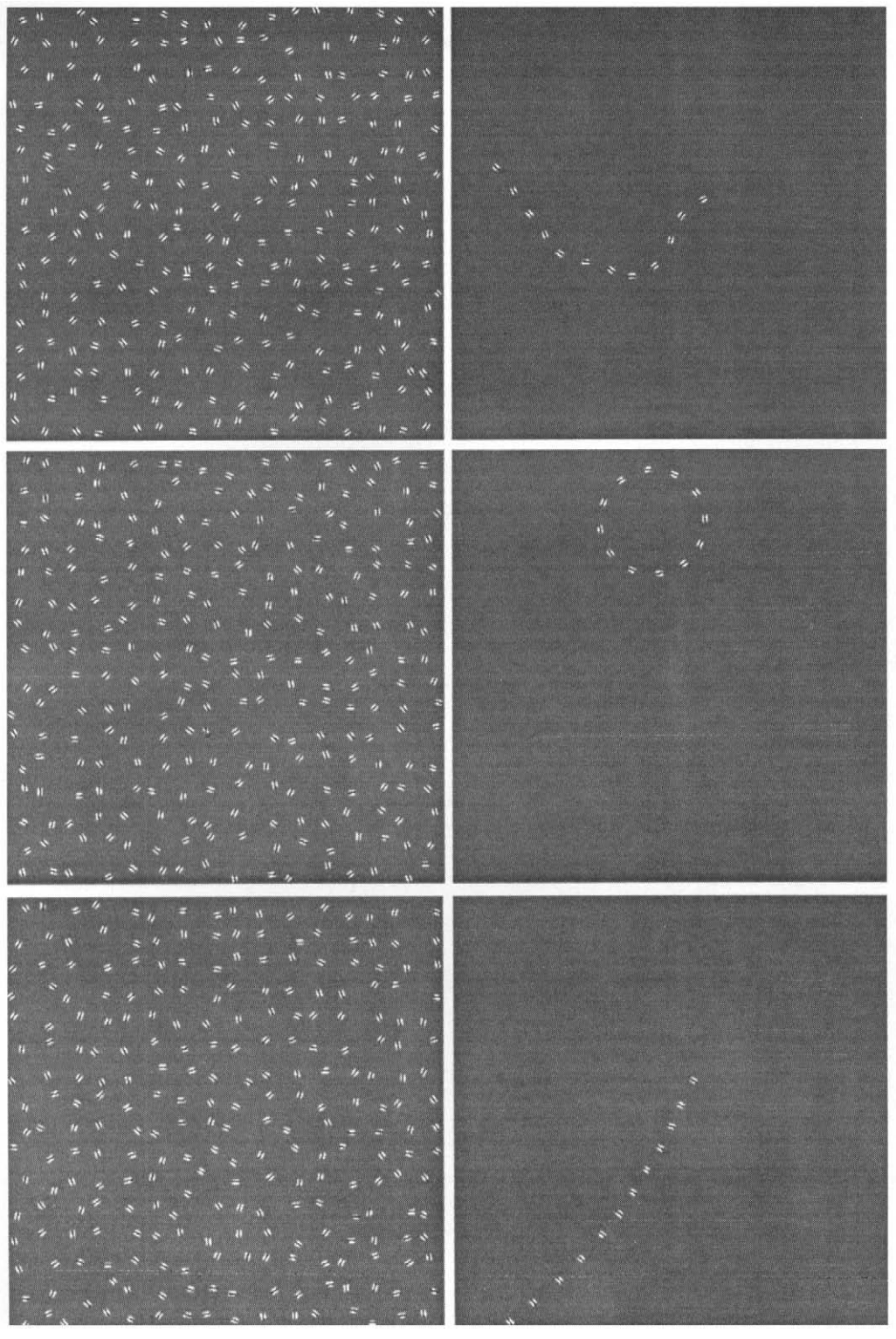

FIGURE 3. Salient contours extracted from stimuli containing an open contour (top), a closed contour (middle), and a straight line made up of orthogonal elements (bottom).

peak in contrast sensitivity at the center of the circle and at the two foci of the ellipse. In addition, contrast sensitivity was elevated at distances approximately 2λ on each side of the contour, while the sensitivity on the contour itself was greatly decreased as compared with the sensitivity to the target in the absence of the contour. Figure 5(b, c) shows the contrast sensitivity maps from Kovács and Julesz (1993). Figure 5(a) shows a simplified "silhouette" of the connectivity pattern for a horizontally oriented cell. The gray level represents the connection weights, with dark regions being strongly facilitatory, and white regions being inhibitory. Since the psychophysical experiments were carried out using low-contrast probes oriented parallel to the closest element on the

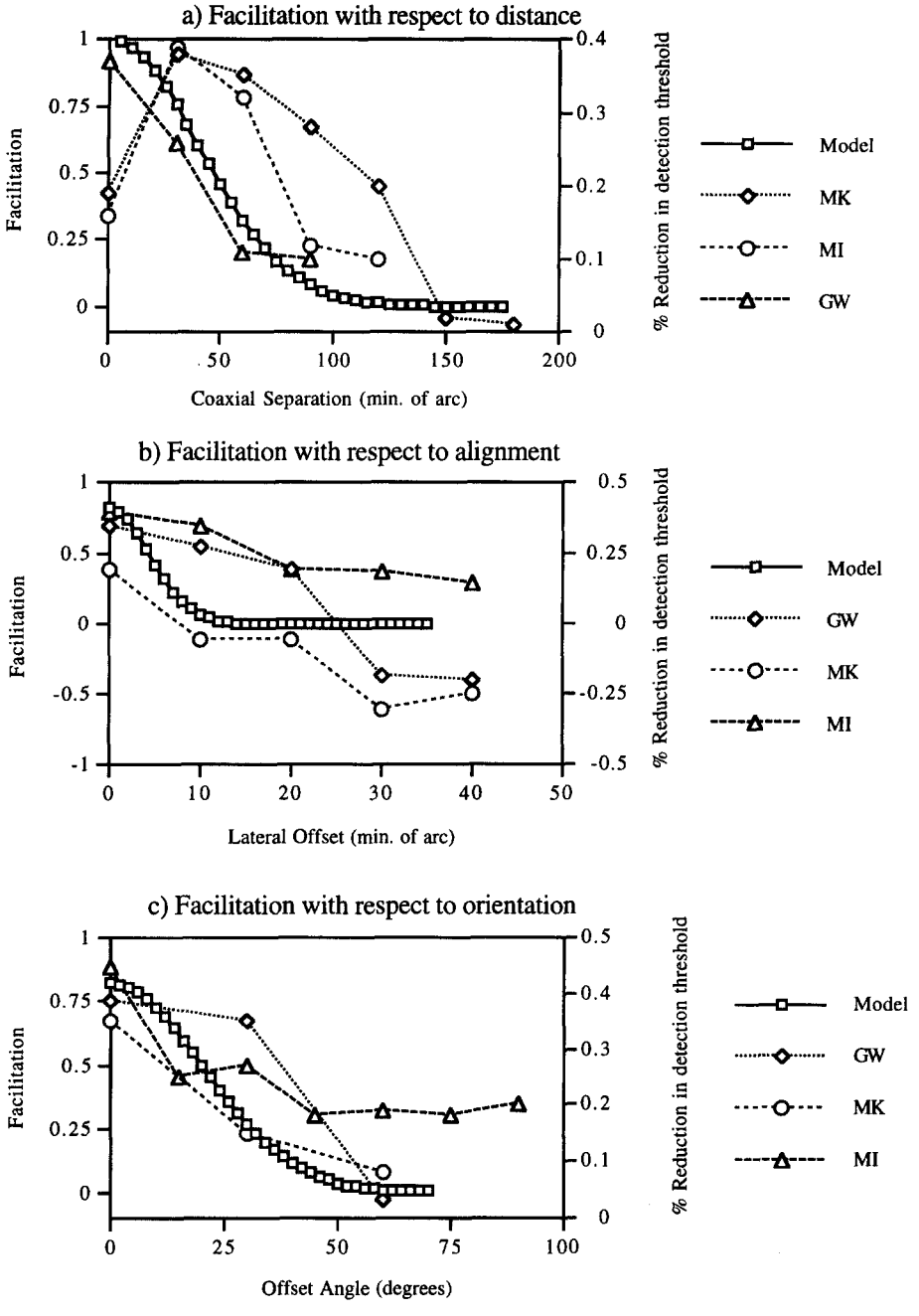

FIGURE 4. Simulations testing the co-axial connections in the model vs the psychophysical results from the three subjects (GW, MK, MI) reported in Kapadia *et al.* (1995). Changes in facilitation due to (a) co-axial distance between stimuli; (b) lateral offset; (c) angular difference between stimuli (see results of Experiment 1).

contour, only the trans-axial connections are likely to be stimulated. We have thus omitted the co-axial connections for simplicity. The cell is surrounded by an inhibitory region at very close distances, corresponding to the intra-filter inhibition observed by Polat and Sagi (1993, 1994), and Kapadia *et al.* (1995). If a number of these silhouettes are placed along a circular or elliptical contour, their excitatory regions superpose. The resulting map of facilitatory regions resembles the experimental findings [Fig. 5(d, e)]. Note especially the peak in the center of the circle and the two peaks in the ellipse due to the trans-axial facilitatory connections. The trans-axial connections are usually strong enough to be facilitatory only out to about 2λ, but the superposition of the subthreshold facilitatory connections combine at the center of the circle and at the two foci of the ellipse to become much stronger. Since the range of facilitation observed in Polat and Sagi's (1993, 1994) experiments scales with size of the Gabor elements, this would also explain the similarities across scale in Kovács and Julesz' (1994) data. Figure 6(a), reproduced from Kovács and Julesz (1993), shows a cross-section plot of the psychophysical sensitivity map inside a circular contour. Figure 6(b) shows the corresponding cross-section plot

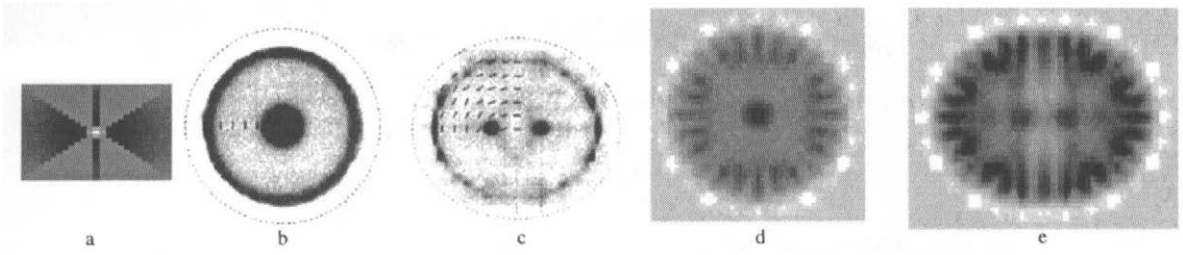

FIGURE 5. Contrast sensitivity at the center of a circle. (a) Simplified silhouette showing intra-filter inhibition and the trans-axial facilitatory connections for a horizontally oriented cell. Dark gray regions are facilitatory, white regions are inhibitory and neutral gray is neutral. (b, c) Contrast sensitivity maps inside a circular and elliptical contour, from Kovács and Julesz (1994). Reprinted with permission from Nature, 370, p. 645, Fig. 2) © (1994) Macmillan Magazines Ltd. (d, e) Sensitivity maps from the model based on averaged connection fields of aligned Gabor units.

from the model. The trough (in threshold elevation) at the center of the circle, and at 2λ away from the contour as well as the peak on the contour correspond well with the psychophysical data.

Experiment 3: extraction of salient contours

Using the same methods as Field *et al.* (1993), we generated stimulus arrays of 256 oriented Gabor elements. Pairs of stimulus arrays were presented to the network, one array contained a contour composed of 12 Gabor elements, the other contained only randomly oriented elements. For each stimulus, the network determines the "salience" of all contours, and selects the contour with the highest salience. Of the two stimuli in each pair presentation, the network "chooses" the stimulus containing the contour with the higher salience (see Appendix). Network performance was measured by computing the percentage of correct detection. The network was tested on a range of stimulus variables governing the target contour (see Fig. 7): (1) the angle, $\pm\beta$, between elements on a contour; (2) the angle, $\pm\beta$, between elements on a contour but with the elements aligned orthogonal to the contour passing through them; (3) the angle, $\pm\beta$, between elements with a random offset angle, $\pm\alpha$, with respect to the contour passing through them; and (4) average separation of the elements. Five hundred simulations were run at each data point. The results are shown in Fig. 9. When the elements are aligned, the performance of the network and human subjects both decrease with increasing β [Fig. 9(a)]. As β increases there is an increased likelihood that the connections between consecutive elements will fall outside the facilitatory zone (defined as κ, a ± 30 deg fan-out from the orientation axis in the model). This can be seen from Fig. 8. This also accounts for the much sharper drop in performance at 60 deg as compared with data from human psychophysics. A better approximation to the data would result from a gradual decrease in the connection strengths as positions deviate from the orientation axis of the post-synaptic cell. Interestingly, Fig. 8 shows that the discrepancy between the threshold orientation differences found by Field *et al.* (1993) (± 60 deg) and Sagi and Kovács (1993) (± 30 deg) may be due to a difference in how the angles were defined. While Field and colleagues measured performance as a function of β, Sagi and Kovács used Glass patterns to

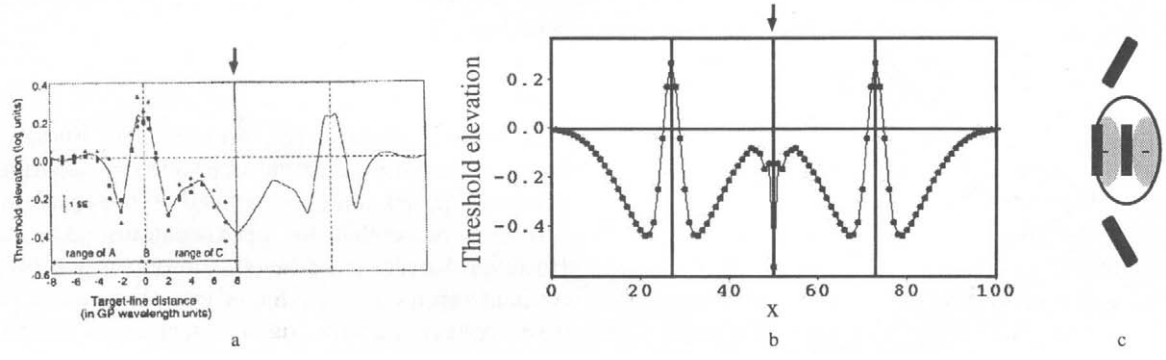

FIGURE 6. (a) A cross-section plot of the elevation in contrast increment thresholds obtained with a circular contour (reproduced with permission from Kovács and Julesz, 1993, Fig. 6). The ordinate axis represents positions along a horizontal axis from the center of the circle to positions outside the circle. Note the elevated thresholds on the contour, and the increased sensitivity at 2λ from the contour and at the center of the circle. The arrow marks the center of the circle. (b) Cross-section contrast sensitivity plot from the model. Threshold elevation is computed as the negative of the facilitation in the model. (c) Figure illustrating intra-filter inhibition (proposed by Polat and Sagi, 1993) within 2λ of a cell.

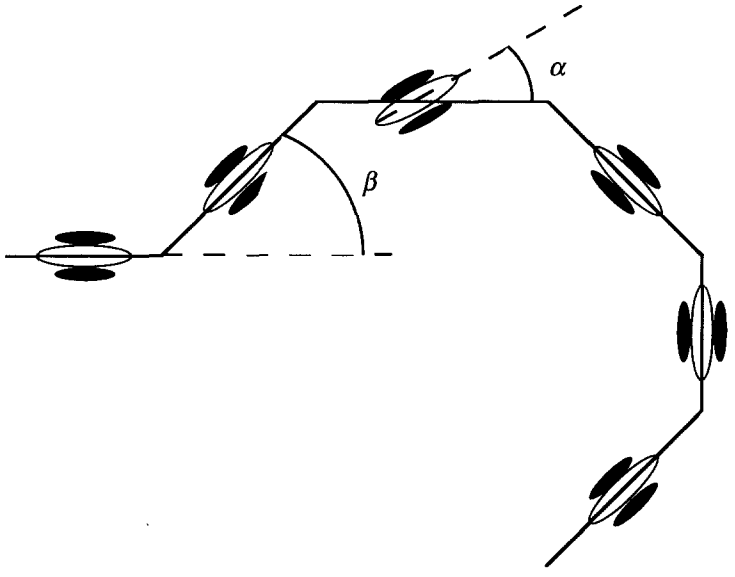

FIGURE 7. Figure illustrating the parameters used in generating the stimulus for Experiments 3 and 4. This figure is modified from Fig. 5 in Field *et al.* (1993).

measure subjects' performance as a function of the difference in angle between the orientation of the Gabor patch and the axis of displacement, which corresponds to angle κ in Fig. 8.

The model is also able to extract contours when the elements are aligned orthogonal to the contour. This ability is modulated by the trans-axial connections. The fan-out of the facilitatory connections are narrower ($\pm$10 deg) than the co-axial connections and as such the performance of the network falls off rather rapidly with increasing β, reaching chance performance at $\pm$30 deg. Again, the network's performance is comparable with the psychophysical data [Fig. 9(b)]. For complex stimuli, the trans-axial connections tend to attach background elements to the target contour but otherwise do not affect results. For this reason, in all other simulations reported in this paper, the trans-axial connections were inactivated. *In vivo*, this inactivation would require segregated inputs to a cell, or the involvement of separate populations of cells mediating co-axial and trans-axial connectivity.

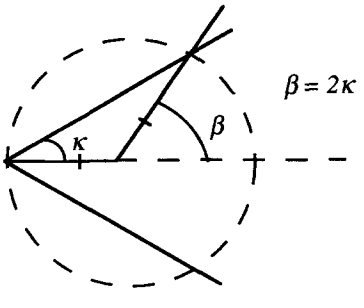

FIGURE 8. The maximum change in angle, β, that can be pre-attentively detected is dependent on the angle of the fan-out, κ, of the connections. In our simulations, κ for the co-axial connections was 30 deg which resulted in a sharp drop in performance at β = 60 deg, since there are no facilitatory connections beyond κ = 30 deg. For the trans-axial connections, κ was set to 10 deg which gives chance performance at β = 30 deg.

When the orientation of the elements are randomly offset with respect to the path of the contour, the pre-synaptic orientations deviate from the "preferred" orientations and thus reduce the strength of the inputs. This increases the likelihood that these inputs may fall below the threshold imposed by the background noise. This leads to "breakage" in the chains and to a decrease in the saliency of the contours. The model demonstrates the same qualitative behavior as the data for both α = 15 deg and α = 30 deg [Fig. 9(c, d)].

As the average separation between all elements increases, the degree of facilitation is decreased. However, since inputs from background elements also decrease, the signal-to-noise ratio is not altered. This allows the model to continue to detect the contour, regardless of the absolute separation between elements. This behavior agrees well with the psychophysical data [Fig. 9(e, f)].

Experiment 4: effects of contour closure

In their original paper, Kovács and Julesz (1993) reported that the maximum inter-element separation for detecting closed contours (defined at 75% performance) is nearly twice that for open contours (Δ_c = 1.8 Δ_o). However, Kovács *et al.* (1996) subsequently showed that elements on a circular contour can be moved up to 1.5-times further than the average separation of the background (φ = 0.65), as discussed in the Introduction section. If this finding applied to their earlier experiment, the maximum separation for open contours, Δ_o, would have to be smaller than the background separation, which is rather unlikely. It is possible that the average separation for the open and the closed contours were

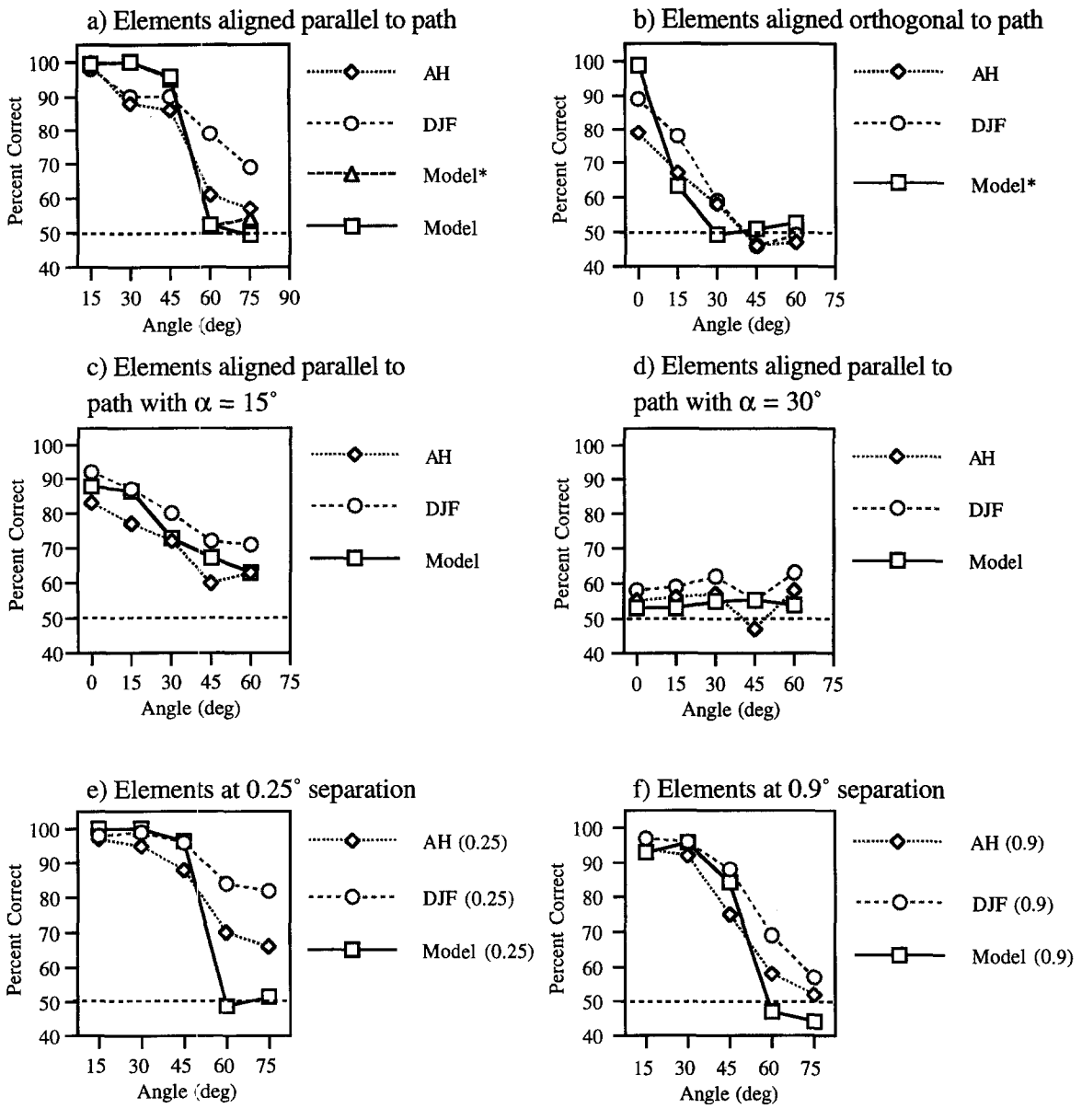

FIGURE 9. Simulations of the model compared with psychophysical results from two subjects (AH, DJF) in Field *et al.* (1993) (see results of Experiment 3) Model results labeled Model* were results of the full model with both co-axial and trans-axial connections, while those labeled Model have the trans-axial connections disabled.

different since they did not control precisely the average separation of the background elements (Kovács, personal communication). This would mean that if they were compared to the same background separation, the threshold separation of open and closed contours might not be $\Delta_c = 6\lambda$ and $\Delta_o = 3.3\lambda$, as found in their earlier paper, but really $\varphi_c = 0.65$ and $\varphi_o \approx 0.65$–1.0. (Δ is the threshold separation of contour elements at a particular ratio of background/contour element separation, φ.)

In addition, they showed that when elements spaced at Δ_o are added to a "jagged" (open) contour, the saliency of the contour increases monotonically, but when elements spaced at Δ_c are added to a circular contour, the saliency does not change until the last element is added and the contour becomes closed. In fact, at Δ_c, the contour is not salient until it is closed, at which point it suddenly "pops-out" [see Fig. 10(c)]. This finding places a strong constraint on the computation of saliency in visual perception.

Interestingly, it has been shown that synchronization in a chain of coupled neural oscillators is enhanced when the chain is closed (Kopell & Ermentrout, 1986; Ermentrout, 1985; Somers & Kopell, 1993). This property is due to the differences in boundary effects on synchronization between open and closed chains and appears to hold across different families of coupled oscillators. It has also been shown that synchronization is dependent on the coupling between oscillators—the

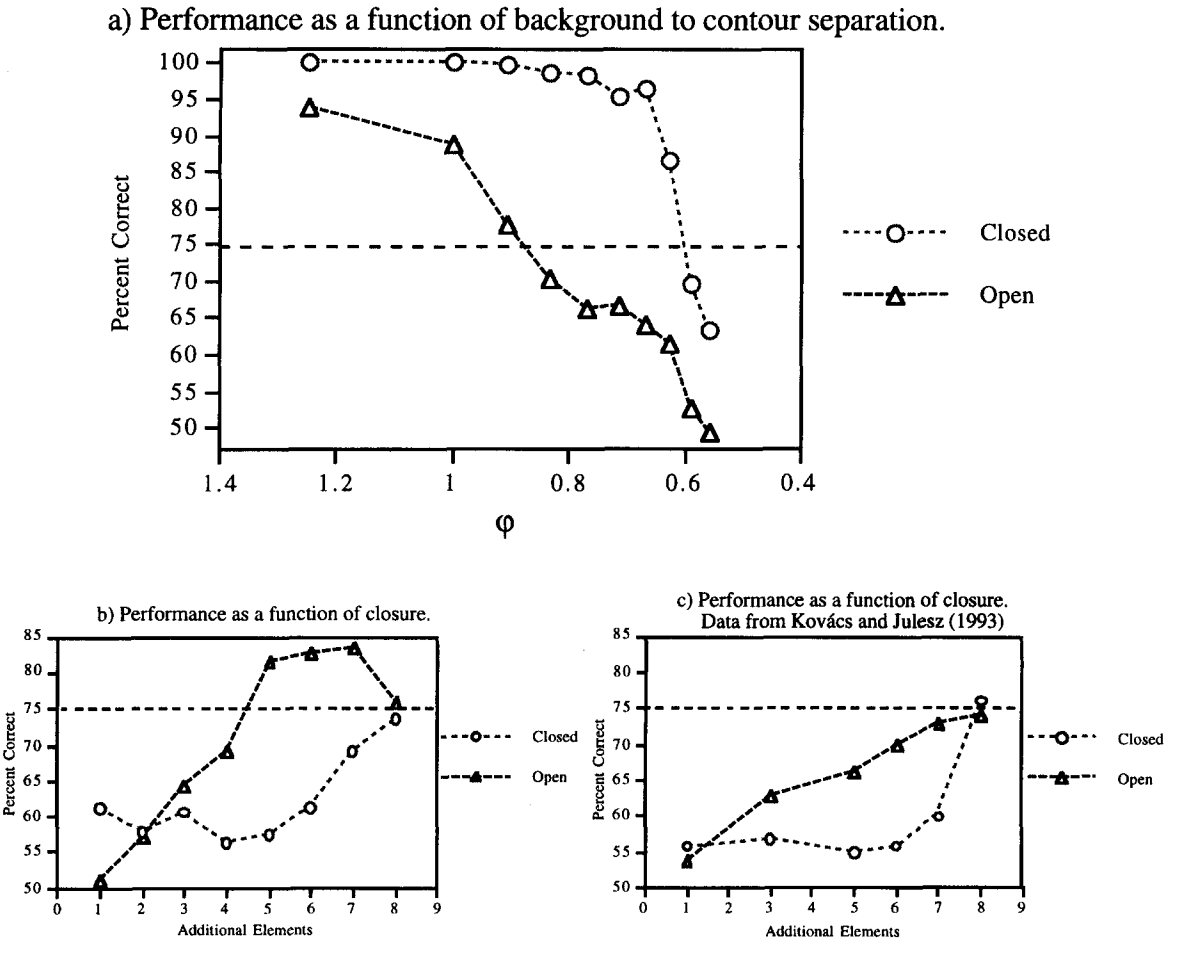

FIGURE 10. Simulation results of the model compared with the data from Kovács and Julesz (1993). (a) Difference in φ for open and closed contours. (b) Changes in salience with additional elements for open and closed contours. Open contours started with seven elements, while closed contours started with 17 elements. (c) The data from Kovács and Julesz (1993) are replotted for comparison.

stronger the coupling, the better the synchronization, both in terms of speed and coherence (Somers & Kopell, 1993; Terman & Wang, 1995). We believe these findings may apply to the psychophysical results. In the model, oscillators couple through their facilitatory connections to other oscillators. However, at the same time, each cell is also receiving noise from the background elements. The noise disrupts the coupling between the oscillators and imposes a threshold separation ratio, φ, beyond which coupling between elements is too weak to allow synchronization to take place. Since closed contours can synchronize with weaker coupling, this translates into a smaller φ_c as compared with φ_o. At φ_o, both open and closed contours are synchronized but at φ_c, elements are synchronized only when the chains are closed. If salience can only be computed for synchronized contours, then as additional elements are added to an open chain at φ_o, the salience would increase since the whole chain is synchronized. On the other hand, at φ_c, as long as the last element is missing, the chain is really an open chain, and since φ_c is smaller than φ_o, the elements on the chain will not be able to synchronize—and adding elements will have no effect on salience. Once the last element is added, the chain is immediately able to synchronize and the salience of the contour increases dramatically and causes the contour to "pop-out".

We simulated the experiments described above to illustrate this point. As in Kovács and Julesz (1993), we generated stimulus arrays containing 2025 elements. Contours were made up of 24 elements. Again the network was presented with two stimuli, one containing a contour and the other made up of all randomly oriented elements. The network picked the stimulus containing the synchronized contour with the higher salience. In separate trials, the contour elements were separated at increasingly greater spacings, while the background elements remained at the same separation (set to 25 pixels in our simulations). The results show that elements on a closed chain were able to synchronize at higher separations compared with the open chains. The threshold ratio of the contour separation to the background separation, defined at 75% accuracy, for open (φ_o) and closed (φ_c) contours were determined, as shown in Fig. 10(a). The results show φ_o to be approx. 0.9 while φ_c is approx. 0.6.

We also examined the changes in saliency due to the

addition of elements to open and closed contours. Stimuli containing 2025 elements were generated with contour elements spaced at φ_o and φ_c. The background separation was kept constant for all stimuli. The response of the network was measured as additional elements were added to an initial short contour of elements. The results are shown in Fig. 10(b). Under our simple synchronization mechanism, the more elements there are on a contour, the longer it takes to synchronize. For open contours, the addition of elements does not adversely affect the synchronization since the threshold was determined using the full number of elements. The additional elements do, however, increase the probability that the stimuli containing the target contour will be picked by the network since they add to the salience of the synchronized, target contour. For the closed contours, the initial chains of elements are unable to synchronize as they are separated further than the threshold for the open contours. This is despite the fact there are now fewer elements on the contour. In this case, the effect of closure is much stronger than the incremental benefit of having to synchronize fewer elements. Only when the last element is added and the contour is closed, is the network able to select the target contour as the contour with the highest salience. This matches the results of Kovács and Julesz (1993), where the saliency of the closed curve does not change significantly until the last element is added, thus "closing" the chain and causing it to "pop-out".

Pettet *et al.* (1996) recently showed that the smoothness of the contour was a stronger constraint for detection than closure. They used a slightly different paradigm which measured performance as a function of the number of background elements. They found that as more and more background elements were added to the scene, the detectability of jagged, closed contours decreased in the same manner as open contours. This was in contrast to smooth, closed contours that were still rather salient at high background densities. We believe our model may also be able to account for these results. The jagged contours contain elements that may be connected by relatively high curvatures, which in our model corresponds to much weaker connection weights. As the number of distractors increase, the signal-to-noise ratio of these connections quickly fall below the threshold required to form a connection. The uncoupling of these elements then removes the closure in the contour and the saliency of these contours then becomes equivalent to that of two separate open contours.

Experiment 5: real images

A more stringent test of the model's capabilities is the ability to extract perceptually salient contours in real images. Figure 11(a, b, c) shows a sample grayscale image, the output of the steerable filters, and the output of the model.* We wanted to identify all the salient contours in the image instead of just isolating the most salient contour, therefore, instead of using temporal binding to separately identify the salient contours in the scene, we modified the network to extract all the salient contours together as a group. These results thus illustrate the degree to which salient contours can be extracted by the non-temporal-based stages of the model. The network is able to extract some of the more salient contours and ignore other high-contrast edges detected by the steerable filters. Figure 11 is a good illustration of how camouflage attempts to re-order the salience of contours. Most of the edges of the plane are effectively invisible due to low contrast and similarity in texture to the background. The highest contrast edges correspond to the camouflage markings on the plane. Nonetheless, the network extracts the plane edges and chooses them as most salient due to their length and straightness. These simulations used filters at only one spatial scale and could be improved through interactions across multiple spatial frequencies. Nevertheless, the model shows promise for automated image processing applications.

*Additional results are shown in Yen & Finkel (1996a,b,c).

DISCUSSION

The model's estimation of salience depends upon a number of factors. The most critical of these are the spatial structure of the anatomical connections, and the balance of facilitatory and inhibitory inputs from contour vs background elements. The most distinguishing characteristic of our model, however, is the use of temporal synchronization to determine salience. In general, temporal mechanisms offer two great advantages—they provide a flexible representation for context-dependent groupings, and they allow effects of global properties to be represented locally (Singer & Gray, 1995; Eckhorn, 1994). In particular, we believe that the experimental results of Kovács and Julesz (1993, 1994) can be explained by the use of a temporal mechanism. Based on analytical results from Kopell and Ermentrout (1986), we have argued that closed contours can synchronize at greater element separations. We have demonstrated this effect in simulations, and shown that the absolute value of the maximum allowable separation depends upon the density of background elements. Closure affects the boundary conditions on synchronization, and in much the same way as standing waves are affected by open or closed boundaries, changing the topology of the oscillator chain leads to dramatically different results.

How might one attempt to account for Kovács and Julesz' results without resorting to temporal mechanisms? Salience could be defined solely on the basis of activity—cells responding to smooth contours would be strongly facilitated and after multiple iterations, contour elements could be differentiated from those of the background. In this case, the effects of facilitation might propagate along a chain of elements, and closed chains would develop a greater degree of facilitation due to the absence of weakly supported end elements. This mechanism assumes that each additional element increases the total facilitation that every element receives. In simulations of such an activity-based mechanism, we

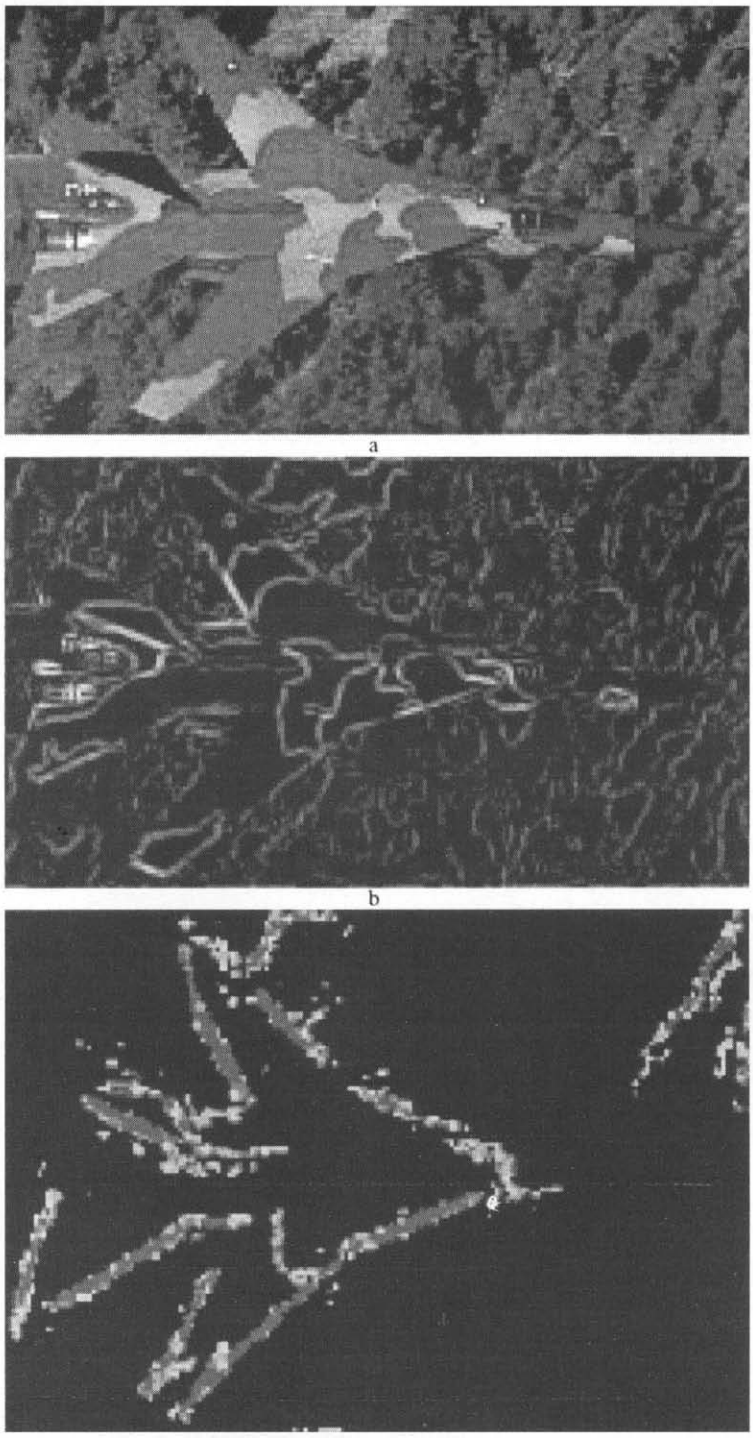

FIGURE 11. (a) Plane image. (b) Steerable filter response. (c) Result of model showing the most salient contours.

have found that the time required for activity levels to converge increases with the number of elements on the chain, unless there is total connectivity between elements. Activity-based representations also lead to a potential confusion between contrast and salience. If the cell is already maximally activated due to the use of maximum contrast stimuli, the cell may be unable to further increase its firing rate to represent increasing salience. The presence of distractors is also problematic for an activity-based representation. In a temporal representation, once the cells representing the contour synchronize, they exclusively facilitate each other, since the cells representing the distractor elements are not simultaneously active. This allows the representation of the contour to be maintained without interference from the background distractors. On the other hand, in an activity-based representation, as the contour elements increase in activity, distractor elements will receive increasing support and distinguishing them from the contour becomes difficult—particularly with respect to

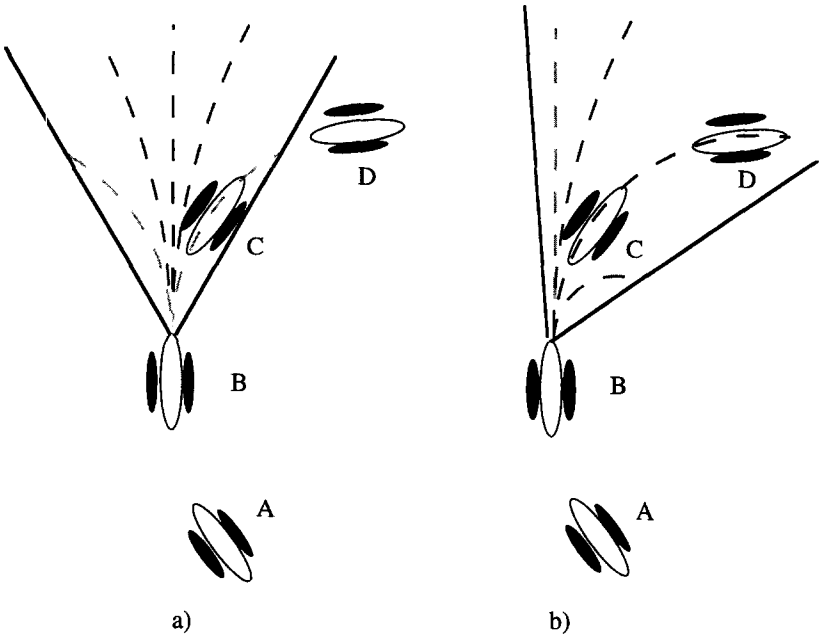

FIGURE 12. Figure illustrating the proposed contextual changes in the connection weights. The dotted lines signify positions with increasing angular deviation from the orientation axis of the post-synaptic cell. Darker lines represent stronger weights. The straight lines mark the fan-out of the facilitatory connections. (a) This figure illustrates the static connection pattern in the absence of supra-threshold stimuli in the surround. (b) This figure shows the proposed potentiation of the connection weights due to context. Not only would the strength of the connections change in terms of angular deviation from the orientation axis, the fan-out would also "steer" in the direction of the supra-threshold stimuli in the surround. This allows elements B and D to now become facilitatory.

elements at the ends of the contour, which may receive less support than the distractors. One possible solution might be to introduce inhibitory mechanisms that will allow strongly facilitated contour elements to inhibit other elements but we found that mechanisms of this sort cause contours that are close to each other to inhibit each other and thus disrupt the representation of multiple contours in nearby or overlapping regions. We also found that with an activity-based representation, the threshold separation between elements varies with the number of elements on the contour. This might be true of contours with small numbers of elements—contours containing two or three elements might appear salient if the elements are spaced much closer together than the background elements. However, the match with longer contours is significantly poorer: the results of Kovács and Julesz show that closed contours made up of 12 elements are just as salient as contours made up of 24 elements. Using a temporal representation, although shorter contours synchronize better, the probability that the salience of such a contour will exceed those found in a random stimulus array is rather low. The increase in salience with additional elements is thus linked to the increase in this probability. With longer chains, so long as the contour is able to synchronize, the salience is reliably higher than those from the random array. Thus, the perceptual threshold of the contour is determined only by the separation that allows synchronization of the elements to take place.

Most critically, we found that while an activity-based mechanism can generate incremental differences between open and closed contours, it cannot produce a *qualitative* change in salience, particularly at threshold (φ) separations. For a synchronization mechanism, the change from open to closed contours represents a transition between different "states". This results in differences in the threshold separation for open and closed contours, but more importantly, provides a mechanism to account for the dramatic difference in salience that accompanies closure. In the simulations of Fig. 10, addition of the last few elements results in a gradual increase in salience rather than a sharp transition. We attribute the gradual nature of the transition to the phase-coupled synchronization mechanism, and believe that a more realistic mechanism, such as relaxation oscillators (Somers & Kopell, 1993; Terman & Wang, 1995) would generate a sharper transition.

Activity-based models may be more applicable in accounting for changes in long-range interactions with stimulus contrast. Stemmler, Usher, and Niebur 1995) recently proposed that at low contrast levels, long-range interactions are primarily excitatory and result in contour completion; at high contrasts, inhibition dominates and the resulting suppression leads to pop-out. Recent experimental evidence has shown that facilitation and inhibition vary with the level of contrast (Mizobe, Polat, Kasamatsu, & Norcia, 1996; Weliky *et al.*, 1995). We did not incorporate this property into our model since the relevant psychophysical experiments were conducted with suprathreshold stimuli of equal contrast.

There may, in addition, be non-temporal mechanisms that contribute to the extraordinary salience of circles. Figure 12(a) shows the static connectivity pattern used in the current model—these connections favor straight lines, with weaker connections to cells off the orientation axis. Figure 12(b) shows a dynamic mechanism that could increase the salience of curved contours. The mechanism involves a context-dependent change in long-range connection strengths such as to optimally tune the cell's input to the structure of the surround. The elements labeled A–D form part of a circular contour. Element B may have facilitatory interactions from elements A and C, but not D, as the angular difference between B and D exceeds the fan-out, κ, of the connections. However, elements B and D may be indirectly joined by a smooth contour going through element C. This is illustrated in Fig. 12(b). The connectivity pattern may be dynamically potentiated by the surrounding elements in such a way as to "steer" the connections to suit the context. In this case, this would correspond to not only the strengthening of the connection between B and D, but also to change the relative weighting of the connections so that the connections corresponding to the curvature represented by elements B–C–D become the strongest connections. Although this mechanism does not explain the difference between open and closed curves (Kovács & Julesz, 1993), it might help account for the fact that circles are at least as salient as straight lines.

Relationship to cortical data

As first pointed out by Sha'ashua and Ullman (1988), the relative salience of different contours suggests something about the cortical mechanisms used to extract them. Sha'ashua and Ullman defined salience based on several heuristic features—contour length, curvature, discontinuities, and gap sizes. They showed that the salient contours in a noisy image can be identified by maximizing an objective function based on these features over all possible contours. Our model measures salience based on these same contour properties—but is implemented using biologically plausible mechanisms.

Although we have attempted to use experimental data to guide our selection of parameters, the model is only intended as a functional description of the operations occurring in visual cortex. For example, the model is not specific as to the type of cortical cell used to determine salience. Independent evidence supports the participation of both simple and complex cells. Kapadia *et al.* (1995) identified complex cells in the supragranular layers of striate cortex as undergoing modulation by stimuli outside the receptive field. Field *et al.* (1993) found that contour extraction was not significantly affected if the Gabor elements were randomly phase-shifted. This also suggests that complex cells, which are not sensitive to spatial phase, might be responsible. However, the complex cells would require rather small receptive fields in order to detect deviations from collinearity. In this regard, simple cells are better suited to detecting collinearity, since any misalignment would greatly reduce the activity of the cells involved. The "chattering" cells in the superficial layers of striate cortex have been found to have properties of simple cells (Gray & McCormick, 1996). Both simple cells and complex cells are found in the superficial layers of cortex where the majority of long-distance horizontal connections are located (Mangini & Pearlman, 1980; Mullikin, Jones, & Palmer, 1984; Gray & McCormick, 1996). It is therefore possible that both of these cell types could be involved or that the complex cells respond due to recruitment by the "chattering" simple cells. Livingstone (1996) recently reported correlated activity across layers of striate cortex in macaque monkeys which could be evidence in support of the latter.

We have assumed that only cells with strong local thalamic input can be facilitated by the long-range horizontal connections. We base this assumption on the observation that facilitatory effects from outside the classical receptive field must be coupled with direct stimulation within the classical receptive field (Nelson & Frost, 1978, 1985; Kapadia *et al.*, 1995). Facilitation could depend upon a voltage-dependent gating mechanism in which horizontal inputs alter the gain of the cell (De Weerd, Gattass, Desimone, & Ungerleider, 1995). Alternatively, horizontal inputs could provide subthreshold activation which requires direct thalamic input to exceed threshold. Psychophysical evidence shows that detection of subthreshold stimuli may be facilitated when superposed on other subthreshold stimuli or on illusory contours (Kulikowski & King-Smith, 1973; Dresp & Bonnet, 1995).* Optical recordings of striate cortex have also revealed broad regions of subthreshold activity consistent with the anatomy of the long-distance connections (Das & Gilbert, 1995). Differentiating between the two mechanisms remains an experimental question, and the model is not predisposed towards either mechanism.

Modulation of cell firing rates by horizontal inputs can result in an alteration of the perceived "brightness" of the contour. While this effect may be consistent with increased saliency, the orientation-specific nature of the long-range inputs may also lead to skewing of the local orientation representation (Westheimer, 1986; Gilbert & Wiesel, 1990). In our simulations, the ratio of the steerable filter responses always reflect the local orientation information. However, at locations without orientation information, facilitation from neighboring locations can "steer" the basis responses to the orientations that best complement the surround. For instance, if a location is flanked by elements with strong vertical information, the local vertical basis filter receives the greatest facilitation and the local orientation becomes "steered" to vertical. This subthreshold "steered" response could become supra-threshold in higher cortical areas—in much the same way as illusory contours are found to

*The facilitation reported by Dresp and Bonnet (1995) was independent of contrast polarity, which may be further evidence that complex cells are involved.

evoke responses in area V2 but not V1 (von der Heydt & Peterhans, 1989).

In the model, inhibition is used to suppress the responses of distractor elements. Although long-range inhibitory effects have been observed physiologically (Nelson & Frost, 1978; Kapadia *et al.*, 1995; Weliky *et al.*, 1995), they were not observed in the experiments of Polat and Sagi (1993, 1994). This may be due to the reduced nature of the stimulus display—the stimulus consisted of only three Gabor patches, as compared with the use of arrays of oriented elements and gratings in the physiological experiments. We speculate that inhibitory effects would emerge with more complex stimuli and would lead to a decrease in contrast sensitivity at locations in the "inhibitory field" of elements on a contour.

Based on reported anatomical and psychophysical studies, we have also included a set of long-distance trans-axial connections. These connections are responsible, in the model, for some of the effects observed by Field *et al.* (1993), as well as the increase in contrast sensitivity observed at the center of closed circles and at the foci of ellipses (Kovács & Julesz, 1994). This latter result strongly depends upon the narrow fan-out of the trans-axial connections, such that a significant intersection is achieved only at the foci. The strengths of these connections decrease rapidly with distance, and the superposition of many subthreshold facilitation is thus required to alter contrast sensitivity. Our model would thus predict that the peak will weaken as the number of elements on the contour is decreased. Misalignment or "wobbling" of the contour elements should also lead to a decrease in the effect. Owing to the orientation dependence of the trans-axial facilitation, only elements on a limited arc of the circle contribute to the change in contrast sensitivity. The effect of these elements may be enhanced by the synchronized nature of their inputs. We would predict that two isolated antipodal arcs that synchronize independently would not lead to changes in contrast sensitivity. Use of trans-axial connections is compatible with the "brush-fire" representation suggested by Kovács and Julesz (1993), and might also be involved in the effects of figure–ground segregation on the responses of striate cortical cells observed by Lamme (1995), Lee, Mumford, and Schiller (1995) and Zipser, Lamme, and Schiller (1996). These connections may also be involved in the detection of symmetry properties.

Synchronization and salience

The phase-coupled synchronization mechanism represents the simplest possible model of a temporal-based binding mechanism. It most likely does not reflect how synchronization is achieved in cortex. We have used phase coupled oscillators following the work by Kopell and Ermentrout (1986), Kammen *et al.* (1989) and Baldi and Meir (1990). Recent work has shown that relaxation oscillators converge much more rapidly (Somers & Kopell, 1993; Terman & Wang, 1995). There is evidence both supporting (Singer & Gray, 1995; Livingstone, 1996; Gray & McCormick, 1996) and arguing against (Ghose & Freeman, 1992; Bair, Koch, Newsome, & Britten, 1994) the significance of actual oscillations in visual cortex. A number of more biologically plausible synchronization mechanisms exist, some depending solely on bursting properties of cells (Traub, Whittington, Stanford, & Jefferys, 1996). These may serve as a means of distinguishing different contours and equally importantly, as a mechanism for determining saliency and other global properties of the stimulus (for example, closure).

The link between synchronization and salience can be tested experimentally. Synchronization could be measured between two cells in striate cortex with widely displaced, non-overlapping receptive fields. An optimally oriented element is positioned in each receptive field, and randomly positioned elements fill the background as in Field *et al.* (1993) and Kovács and Julesz (1993, 1994). If the optimal orientations are dissimilar, the cells should be desynchronized, but when the orientations of intervening elements are rotated to form a smooth contour between the receptive fields, the activity should become synchronized.* As the intervening contour elements are spaced farther apart synchronization should abruptly cease at the threshold separation, φ, which will depend upon the density of background elements.

We have employed the simplest possible algorithm for computing salience—the unweighted linear sum of activities. Since salience is relative (a long contour might look salient against a field of short contours, but not against longer contours) a normalized salience measure might be more appropriate. The sum of the activities of a synchronized populations should really be compared to the average length of the background contours to provide an accurate measure of salience. The absolute magnitude of the activities should also be taken into account, thus, a statistical measure such as the χ^2 value, as suggested by Grzywacz, Watamaniuk, and McKee (1995), might be appropriate. We believe the results of the model discussed in the paper will remain unaffected by such a change in the salience measure.

One crucial requirement of any synchronization mechanism is to prevent different contours from merging and becoming synchronized. In addition, cells should be able to simultaneously synchronize with different populations, representing multiple contours, for example, at a "Y"-junction. This requires an explicit mechanism for desynchronization. Previous studies have used global inhibition to separate synchronized populations (von der Malsburg & Buhmann, 1992; Terman & Wang, 1995; Campbell & Wang, 1996) or time-delayed inhibition (Schillen & König, 1991). These mechanisms work for populations that are synchronized independently and only need to be distinct in phase from each other. However, the synchronization mechanism must also involve a *local* desynchronization that allows separate contours to be

*A related prediction was made by Hummel and Biederman (1992).

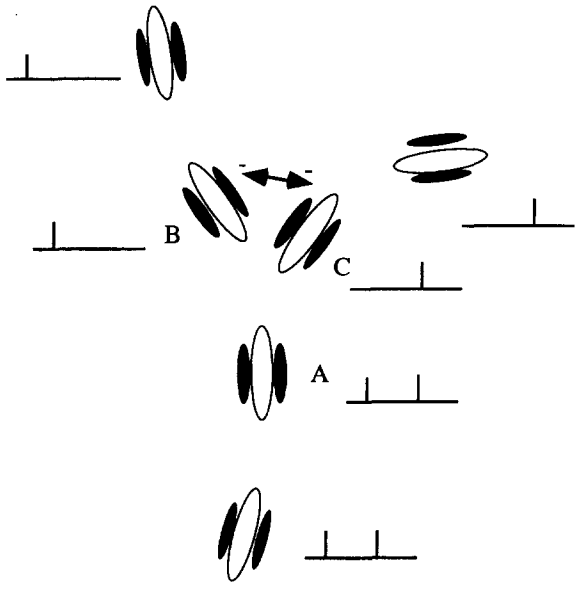

FIGURE 13. Figure illustrating the use of temporal codes in representing multiple contours at a single location. At element "A", there are two equally plausible continuations of the contour, one going through "B", and another going through "C". Owing to the desynchronization between "B" and "C", "A" could take on multiple tags to remain consistent with both "B" and "C". This would allow multiple contours to be represented at "A". This scheme could be implemented by cells capable of synchronizing with multiple different populations or through the synchronization of subpopulations of cells with different populations.

segregated from each other. One possible mechanism might involve separately synchronized subpopulations within an orientation column. Figure 13 illustrates an alternative mechanism in which a single cell can synchronize with multiple different populations. In the figure, elements A, B, C lie along two different contours bifurcating at element A. Elements A and B would normally synchronize since they are interconnected, and so would elements A and C. However, elements B and C are incompatible with each other and would actively desynchronize. This incompatibility between B and C thus causes A, and the cells already synchronized to A, to represent both the phases at B and at C separately. This allows two contours to be stably represented at A and could also have implications for the generation of bistable percepts, like the Necker cube.

The model has focused on feature-based influences on salience. In our model, these "bottom-up" processes take place in parallel and provide a salience ranking for all the contours in the image. A selective attentional process could identify each contour based on its temporal pattern, and sequentially visit contours in rank order (Koch & Ullman, 1985). Salience can also be modulated in a top-down manner by attention. Visual search has been modeled as a competition between targets and non-targets for access to short-term memory (Duncan & Humphreys, 1989; Wolfe, 1994), where the rejection of one distractor leads to the suppression of all similar distractors. Temporal mechanisms offer a means of implementing such a process. In addition, temporal mechanisms might allow high-level conjunctions and disjunctions of features to be represented. One possible physiological mechanism for these attentional effects may involve cholinergic inputs which are known to alter the bursting properties of cortical cells (McCormick, 1993; Sillito, 1993; Cox, Metherate, & Ashe, 1994; Gray & McCormick, 1996). Salience may thus ultimately reflect the contributions of several classes of cortical connections: ascending, descending, horizontal, and neuromodulatory.

REFERENCES

Bair, W., Koch, C., Newsome, W. & Britten, K. (1994). Power spectrum analysis of bursting cells in area MT in the behaving monkey. *Journal of Neuroscience, 14*, 2870–2892.

Baldi, P. & Meir, R. (1990). Computing with arrays of coupled oscillators: an application to preattentive texture discrimination. *Neural Computation, 2*, 458–471.

Campbell, S. & Wang, D. L. (1996). Synchronization and desynchronization in a network of locally coupled Wilson–Cowan oscillators. *IEEE Transactions on Neural Networks, 7*, 541–554.

Cox, C. L., Metherate, R. & Ashe, J. H. (1994). Modulation of cellular excitability in neocortex: muscarinic receptor and second messenger-mediated actions of acetylcholine. *Synapse, 16*, 123–136.

Das, A. & Gilbert, C. D. (1995). Long-range horizontal connections and their role in cortical reorganization revealed by optical recording of cat primary visual cortex. *Nature, 375*, 780–784.

DeAngelis, G. C., Robson, J. G., Ohzawa, I. & Freeman, R. D. (1992). Organization of suppression in receptive fields of neurons in cat visual cortex. *Journal of Neurophysiology, 68*, 144–163.

DeAngelis, G. C., Freeman, R. D. & Ohazawa, I. (1994). Length and width tuning of neurons in the cat's primary visual cortex. *Journal of Neurophysiology, 71*, 347–374.

De Weerd, P., Gattass, R., Desimone, R. & Ungerleider, L. G. (1995). Responses of cells in monkey visual cortex during perceptual filling-in of an artificial scotoma. *Nature, 377*, 731–734.

Dresp, B. & Bonnet, C. (1995). Subthreshold summation with illusory contours. *Vision Research, 35*, 1071–1078.

Duncan, J. & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review, 96*, 433–458.

Eckhorn, R. (1994). Oscillatory and non-oscillatory synchronizations in the visual cortex and their possible roles in associations of visual features. *Progress in Brain Research, 102*, 405–426.

Engel, A. K., König, P., Gray, C. M. & Singer, W. (1990). Stimulus-dependent neuronal oscillations in cat visual cortex: intercolumnar interaction as determined by cross-correlational analysis. *European Journal of Neuroscience, 2*, 558–606.

Engel, A. K., Kreiter, A. K., König, P. & Singer, W. (1991). Synchronization of oscillatory neuronal responses between striate and extrastriate visual cortical areas of the cat. *Proceedings of the National Academy of Sciences USA, 88*, 6048–6052.

Ermentrout, G. B. (1985). The behavior of rings of coupled oscillators. *Journal of Mathematical Biology, 23*, 55–74.

Field, D. J., Hayes, A. & Hess, R. F. (1993). Contour integration by the human visual system: evidence for a local Association Field. *Vision Research, 33*, 173–193.

Fitzpatrick, D. (1996). The functional-organization of local circuits in visual cortex—insights from the study of tree shrew striate cortex. *Cerebral Cortex, 6*, 329–341.

Freeman, W. T. & Adelson, E. H. (1991). The design and use of steerable filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 13*, 891–906.

Ghose, G. M. & Freeman, R. D. (1992). Oscillatory discharge in the visual system: does it have a functional role? *Journal of Neurophysiology, 68*, 1558–1574.

Gilbert, C. D. & Wiesel, T. N. (1979). Morphology and intracortical

projections of functionally identified neurons in cat visual cortex. *Nature, 280*, 120–125.

Gilbert, C. D. & Wiesel, T. N. (1983). Clustered intrinsic connections in cat visual cortex. *Journal of Neuroscience, 3*, 1116–1133.

Gilbert, C. D. & Wiesel, T. N. (1989). Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex. *Journal of Neuroscience, 9*, 2432–2442.

Gilbert, C. D. & Wiesel, T. N. (1990). The influence of contextual stimuli on the orientation selectivity of cells in primary visual cortex of the cat. *Vision Research, 30*, 1689–1701.

Gilbert, C. D., Das, A., Ito, M., Kapadia, M. & Westheimer, G. (1996). Spatial integration and cortical dynamics. *Proceedings of the National Academy of Sciences USA, 93*, 615–622.

Gray, C. M., König, P., Engel, A. K. & Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature, 338*, 334–337.

Gray, C. M. & McCormick, D. A. (1996). Chattering cells-superficial pyramidal neurons contributing to the generation of synchronous oscillations in the visual-cortex. *Science, 274*, 109–113.

Grossberg, S. & Somers, D. (1991). Synchronized oscillations during cooperative feature linking in a cortical model of visual perception. *Neural Networks, 4*, 453–466.

Grzywacz, N. M., Watamaniuk, S. N. & McKee, S. P. (1995). Temporal coherence theory for the detection and measurement of visual motion. *Vision Research, 35*, 3183–3203.

Hata, Y., Tsumoto, T., Sato, H., Hagihara, K. & Tamura, H. (1993). Development of local horizontal interactions in cat visual cortex studied by cross-correlation analysis. *Journal of Neurophysiology, 69*, 40–56.

Heeger, D. J. (1992). Normalization of cell responses in cat striate cortex. *Visual Neuroscience, 9*, 181–197.

Hess, R. F. & Field, D. J. (1995). Contour integration across depth. *Vision Research, 35*, 1699–1711.

Hubel, D. H. & Wiesel, T. N. (1974). Uniformity of monkey striate cortex: a parallel relationship between field size, scatter and magnification factor. *Journal of Comparative Neurology, 158*, 295–306.

Hummel, J. E. & Biederman, I. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review, 99*, 480–517.

Kammen, D. M., Holmes, P. J. & Koch, C. (1989). Origin of oscillations in visual cortex: feedback versus local coupling. In R. M. J. Cotterill (Ed.), *Models of brain functions* (pp. 273–284). Cambridge: Cambridge University Press.

Kapadia, M. K., Ito, M., Gilbert, C. D. & Westheimer, G. (1995). Improvement in visual sensitivity by changes in local context: parallel studies in human observers and in V1 of alert monkeys. *Neuron, 15*, 843–856.

Kisvárday, Z. F. & Eysel, U. T. (1992). Cellular organization of reciprocal patchy networks in layer III of cat visual cortex (area 17). *Neuroscience, 46*, 275–286.

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

Koch, C. & Ullman, S. (1985). Shifts in selective visual attention: towards the underlying neural circuitry. *Human Neurobiology, 4*, 219–227.

Koffka, K. (1935). *Principles of Gestalt psychology*. New York: Harcourt Brace.

König, P. & Schillen, T. B. (1991). Stimulus-dependent assembly formation of oscillatory responses: I. Synchronization. *Neural Computation, 3*, 155–166.

Kopell, N. & Ermentrout, G. B. (1986). Symmetry and phaselocking in chains of weakly coupled oscillators. *Communications on Pure and Applied Mathematics, 39*, 623–660.

Kovács, I. & Julesz, B. (1993). A closed curve is much more than an incomplete one: effect of closure in figure–ground segmentation. *Proceedings of National Academy of Sciences USA, 90*, 7495–7497.

Kovács, I. & Julesz, B. (1994). Perceptual sensitivity maps within globally defined visual shapes. *Nature, 370*, 644–646.

Kovács, I., Polat, U. & Norcia, A. M. (1996). Breakdown of binding mechanisms in amblyopia. *Investigative Ophthalmology & Visual Science, 37*, 3078.

Kulikowski, J. J. & King-Smith, P. E. (1973). Spatial arrangement of line, edge and grating detectors revealed by subthreshold summation. *Vision Research, 13*, 1455–1478.

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

Lee, T. S., Mumford, D. & Schiller, P. H. (1995). Neural correlates of boundary and medial axis representations in primate striate cortex. *Investigative Ophthalmology & Visual Science, 36*, 477.

Li, Z. & Hopfield, J. J. (1989). Modeling the olfactory bulb and its neural oscillatory processings. *Biological Cybernetics, 61*, 379–392.

Livingstone, M. S. (1996). Oscillatory firing and interneuronal correlations in squirrel-monkey striate cortex. *Journal of Neurophysiology, 75*, 2467–2485.

Lund, J., Fitzpatrick, D. & Humphrey, A. L. (1985). The striate visual cortex of the tree shrew. In Jones, E. G. & Peters, A. (Eds), *Cerebral cortex* (pp. 157–205). New York: Plenum.

Mangini, N. J. & Pearlman, A. L. (1980). Laminar distribution of receptive field properties in the primary visual cortex of the mouse. *Journal of Comparative Neurology, 193*, 203–222.

Martin, K. A. & Whitteridge, D. (1984). Form, function and intracortical projections of spiny neurones in the striate visual cortex of the cat. *Journal of Physiology, 353*, 463–504.

McCormick, D. A. (1993). Actions of acetylcholine in the cerebral cortex and thalamus and implications for function. *Progress in Brain Research, 98*, 303–308.

McGuire, B. A., Gilbert, C. D., Rivlin, P. K. & Wiesel, T. N. (1991). Targets of horizontal connections in macaque primary visual cortex. *Journal of Comparative Neurology, 305*, 370–392.

Mitchison, G. & Crick, F. (1982). Long axons within the striate cortex: their distribution, orientation, and patterns of connection. *Proceedings of National Academy of Sciences USA, 79*, 3661–3665.

Mizobe, K., Polat, U., Kasamatsu, T. & Norcia, A. M. (1996). Lateral masking reveals facilitation and suppression from the same single cells in cat area-17. *Investigative Ophthalmology & Visual Science, 37*, 2188.

Mullikin, W. H., Jones, J. P. & Palmer, L. A. (1984). Receptive-field properties and laminar distribution of X-like and Y-like simple cells in cat area 17. *Journal of Neurophysiology, 52*, 350–371.

Nelson, J. I. & Frost, B. J. (1978). Orientation-selective inhibition from beyond the classic visual receptive field. *Brain Research, 139*, 359–365.

Nelson, J. I. & Frost, B. J. (1985). Intracortical facilitation among co-oriented, co-axially aligned simple cells in cat striate cortex. *Experimental Brain Research, 61*, 54–61.

Newsome, W. T., Britten, K. H. & Movshon, J. A. (1989). Neuronal correlates of a perceptual decision. *Nature, 341*, 52–54.

Parent, P. & Zucker, S. W. (1989). Trace inference, curvature consistency, and curve detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 11*, 823–839.

Pettet, M. W. & Gilbert, C. D. (1992). Dynamic changes in receptive field size in cat primary visual cortex. *Proceedings of the National Academy of Science USA, 89*, 8366–8370.

Pettet, M. W., McKee, S. P. & Grzywacz, N. M. (1996). Smoothness constrains long-range interactions mediating contour-detection. *Investigative Ophthalmology and Visual Science, 37*, 4368.

Polat, U. & Sagi, D. (1993). Lateral interactions between spatial channels: suppression and facilitation revealed by lateral masking experiments. *Vision Research, 33*, 993–999.

Polat, U. & Sagi, D. (1994). The architecture of perceptual spatial interactions. *Vision Research, 34*, 73–78.

Rock, I. & Palmer, S. (1990). The legacy of Gestalt psychology. *Scientific American, 263*, 84–90.

Rockland, K. S. & Lund, J. S. (1982). Widespread periodic intrinsic connections in the tree shrew visual cortex. *Science, 215*, 532–534.

Rockland, K. S. & Lund, J. S. (1983). Intrinsic laminar lattice connections in primate visual cortex. *Journal of Comparative Neurology, 216*, 303–318.

Sagi, D. & Kovács, I. (1993). Long range processes involved in the

perception of Glass patterns. *Investigative Ophthalmology and Visual Science, 34*, 2106.

Schillen, T. B. & König, R. (1991). Stimulus-dependent assembly formation of oscillatory responses: II Desynchronization. *Neural Computation, 3*, 167–177.

Sha'ashua, A. & Ullman, S. (1988). Structural saliency: the detection of globally salient structures using a locally connected network. *Proceedings of the second international conference on computer vision*, Dec. 1988, Tampa, FL, USA, (pp. 321–327).

Sillito, A. M. (1993). The cholinergic neuromodulatory system: an evaluation of its functional roles. *Progress in Brain Research, 98*, 371–378.

Sillito, A. M., Grieve, K. L., Jones, H. E., Cudeiro, J. & Davis, J. (1995). Visual cortical mechanisms detecting focal orientation discontinuities. *Nature, 378*, 492–496.

Singer, W. & Gray, C. M. (1995). Visual feature integration and the temporal correlation hypothesis. *Annual Review of Neuroscience, 18*, 555–586.

Somers, D. & Kopell, N. (1993). Rapid synchronization through fast threshold modulation. *Biological Cybernetics, 68*, 393–407.

Stemmler, M., Usher, M. & Niebur, E. (1995). Lateral interactions in primary visual cortex: a model bridging physiology and psychophysics. *Science, 269*, 1877–1880.

Terman, D. & Wang, D. L. (1995). Global competition and local cooperation in a network of neural oscillators. *Physica D, 81*, 148–176.

Toyama, K., Kimura, M. & Tanaka, K. (1981b) Organization of cat visual cortex as investigated by cross-correlation techniques. *Journal of Neurophysiology, 46*, 202–214.

Traub, R. D., Whittington, M. A., Stanford, I. M. & Jefferys, J. G. R. (1996). A mechanism for generation of long-range synchronous fast oscillations in the cortex. *Nature, 383*, 621–624.

Ts'o, D. & Gilbert, C. D. (1988). The organization of chromatic and spatial interactions in the primate striate cortex. *Journal of Neuroscience, 8*, 1712–1727.

Ts'o, D., Gilbert, C. D. & Wiesel, T. N. (1986). Relationships between horizontal and functional architecture in cat striate cortex as revealed by cross-correlation analysis. *Journal of Neuroscience, 6*, 1160–1170.

von der Heydt, R. & Peterhans, E. (1989). Mechanisms of contour perception in monkey visual cortex. I. Lines of pattern discontinuity. *Journal of Neuroscience, 9*, 1731–1748.

von der Malsburg, C. & Buhmann, J. (1992). Sensory segmentation with coupled neural oscillators. *Biological Cybernetics, 67*, 233–242.

Watamaniuk, S. N. & Sekuler, R. (1992). Temporal and spatial integration in dynamic random-dot stimuli. *Vision Research, 32*, 2341–2347.

Weliky, M., Kandler, K., Fitzpatrick, D. & Katz, L. C. (1995). Patterns of excitation and inhibition evoked by horizontal connections in visual cortex share a common relationship to orientation columns. *Neuron, 15*, 541–552.

Westheimer, G. (1986). Spatial interaction in the domain of disparity signals in human stereoscopic vision. *Journal of Physiology, 370*, 619–629.

Wolfe, J. M. (1994). Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin and Review, 1*, 202–238.

Yen, S-C. & Finkel, L. H. (1996a) Pop-out of salient contours in a network based on striate cortical connectivity. *Investigative Ophthalmology and Visual Science, 37*, 1348.

Yen, S-C. & Finkel, L. H. (1996b). Cortical synchronization mechanisms for "pop-out" of salient image contours. In Bower, J. M. (Ed.), *The Neurobiology of computation*, Massachusetts: Kluwer Academic Publisher.

Yen, S-C. & Finkel, L. H. (1996c). Salient contour extraction by temporal binding in a cortically-based Network. In Jordan, M. I., Mozer, M. C. & Petsche, T. (Eds), *Advances in neural information processing systems 9*. Massachusetts: MIT Press.

Young, R. A. (1987). The gaussian derivative model for spatial vision: I Retinal mechanisms. *Spatial Vision, 2*, 273–293.

Young, R. A. & Lesperance, R. M. (1993). A physiological model of motion analysis for machine vision. *Proceedings of the SPIE—The International Society for Optical Engineering, 1913*, 48–123.

Zipser, K., Lamme, V. A. F. & Schiller, P. H. (1996). Contextual modulation in primary visual-cortex. *Journal of Neuroscience, 1622*, 7376–7389.

Acknowledgements—We thank Ilona Kovács, Mitesh Kapadia, and Nancy Kopell for helpful discussions of this work. Supported by the Office of Naval Research N00014-93-1-0681 and the Whitaker Foundation.

APPENDIX

Model cells. The cells in the model are represented by linearly separable G2 and H2 steerable filters. The responses of the filters are represented as:

$$A(x, y, \theta) = \int\int f(x, y, \theta) I(x, y) dx dy$$

where $f(x, y, \theta)$ is the kernel of a steerable filter oriented at θ, and $I(x, y)$ is the input image.

The dominant orientation at each position is computed directly from the basis responses using the method proposed in Freeman and Adelson (1991). The facilitation is then computed at only the dominant orientation as a simplification. The responses at the dominant orientation are also squared to allow interactions between cells of opposite contrast polarity. This is consistent with results from Dresp and Bonnet (1995), who showed that the facilitation of contrast sensitivity to a subthreshold bar does not depend on the direction of contrast.

Facilitation. Each cell in the model receives weighted inputs from other oriented cells in its surround. The connection weights are dependent on position as well as orientation. The facilitation that a post-synaptic cell of orientation θ receives may be represented by:

$$F(x, y, \theta) = \int_{(i,j)} \int\int_{\in N} W(\theta, i, j, \psi) A(x + i, y + j, \psi) d\psi di dj,$$

where ψ is the orientation of the pre-synaptic cell, i and j represent positions in a neighborhood N, around the post-synaptic cell. $W(\theta, i, j, \psi)$ is the connection weight between the post-synaptic cell oriented at θ, and a pre-synaptic cell of orientation ψ, and $A(x+i, y+j, \psi)$ is the activity of the pre-synaptic cell.

The connection weight between the pre-synaptic and post-synaptic cell depends on their relative locations, as well as their respective orientations. The co-circularity rule imposes the constraint that for two points lying on a circle, the average of their tangent orientations equals the slope of the line between them. Given the orientation, θ, of the post-synaptic cell, the "preferred" orientation, ϕ, at the position (i, j) of the pre-synaptic cell is specified by:

$$\phi(\theta, i, j) = 2 \tan^{-1}\left(\frac{j}{i}\right) - \theta$$

where θ is the orientation of the post-synaptic cell, and i,j are positions relative to the post-synaptic cell [see Fig. 1(a)]. The connection weights peak at ϕ and fall off as a gaussian function of the difference in the acute angle ($||_{\text{acute}}$) between ϕ and ψ, with half-width at half-height, σ^c_ψ, for the co-axial connections:

$$B(\theta, i, j, \psi) = G(|\psi - \phi(\theta, i, j)|_{\text{acute}}, \sigma^c_\psi).$$

The connection weights also fall off as a gaussian function of distance, similar to the effects observed in psychophysical studies by Polat and Sagi (1993), as well as in cortical cells by Kapadia *et al.* (1995):

$$D(i, j) = G\left(\sqrt{i^2 + j^2}, \sigma^c_d\right)$$

where σ^c_d represents the half-width at half-height of the gaussian function. Similarly, the trans-axial connections are governed by σ^t_ψ and σ^t_d.

Connection fan-out is limited to low curvature deviations from the

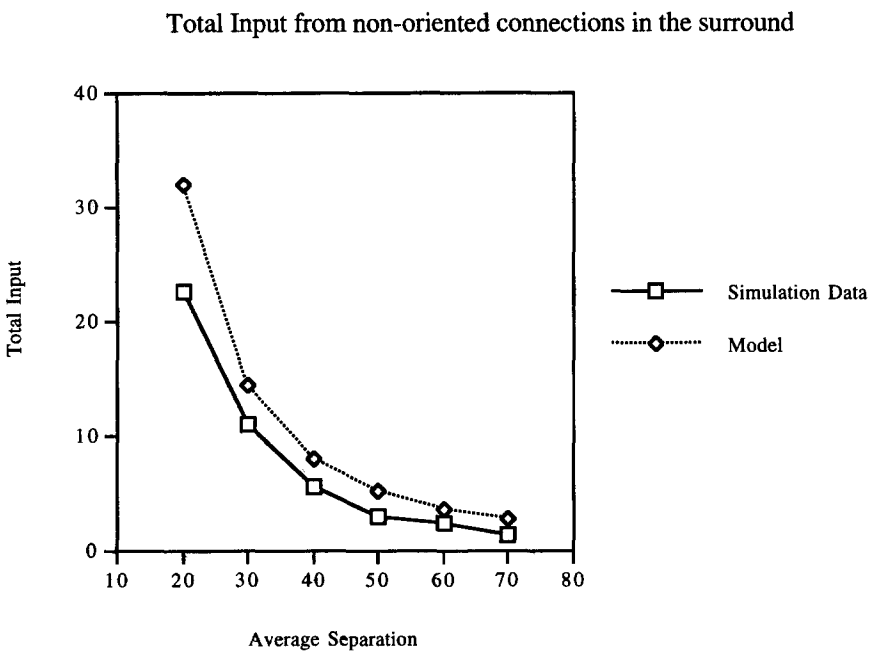

FIGURE A1. Match between the quantitative model for the prediction of average separation of background elements and the actual values obtained in our simulations with stimuli at different separations. The model predicts the average separation using the sum of the inputs from non-oriented cells. The average separation is then used to impose a threshold, which inputs have to exceed in order to be facilitatory.

orientation axis and to a narrow region extending orthogonal to the cell's orientation axis. In addition, horizontal connections have been observed to be reciprocal (Kisvárday & Eysel, 1992). These constraints may be expressed as:

$$\Gamma(\theta,i,j,\psi) = \begin{cases} 1, \textit{ if } \tan^{-1}\left(\frac{j}{i}\right) - \theta < \kappa^c, \\ \quad \textit{and } \tan^{-1}\left(\frac{-j}{-i}\right) - \psi < \kappa^c, \\ 1, \textit{ if } \tan^{-1}\left(\frac{j}{i}\right) - \theta = \frac{\pi}{2} \pm \kappa^t, \\ \quad \textit{and } \tan^{-1}\left(\frac{-j}{-i}\right) - \psi = \frac{\pi}{2} \pm \kappa^t, \\ 0, \textit{ otherwise} \end{cases}$$

where κ^c represents the maximum angular deviation of the co-axial connections and κ^t represents the maximum angular deviation of the trans-axial connections. The terms containing $(-i, -j)$ represent the reciprocal fan-out constraints imposed by the pre-synaptic cell.

Thus, the facilitation for a cell of orientation θ, located at position (x, y), can be represented by:

$$F(x,y,\theta) = \iiint_{(i,j)\in N} \Gamma(\theta,i,j,\psi)D(i,j)B(\theta,i,j,\psi) A(x+i,y+j,\psi)d\psi\, di\, dj \quad \text{(A1)}$$

Fig. 1(b) illustrates the connectivity pattern for a horizontally oriented cell.

Dynamic threshold. Each cell receives a large number of inputs from its surround and the facilitated inputs must be significantly larger than the noise contributed by background elements in order to be effective. We used the following mechanism: each cell estimates the average input from its immediate neighbors, and any input larger than 1.5-times the average input becomes facilitatory; any input lower than this threshold has no excitatory effect. We believe the threshold computation is carried out by connections from a separate set of non-oriented cells (similar to those reported in Fitzpatrick, 1996), which, in effect, have the same activity at different orientations. The activity of the oriented cells described thus far in the paper depend on both distance and orientation and would thus not be suitable for carrying out a computation estimating distance. Assuming an uniform distribution of elements on a square grid, each cell is surrounded by eight immediate neighbors, 16 next-nearest neighbors, 24 next-next-nearest neighbors, etc. As the elements become separated by a larger and larger distance, the facilitation from each cell decreases as a function of distance. If the elements were separated by an average distance of s, the total input to a cell, I, from the surround might be expressed as:

$$I = 8e^{\frac{s^2}{2(\sigma_d)^2}} + 16e^{\frac{(2s)^2}{2(\sigma_d)^2}} + \ldots,$$

where σ_d is the half-width at half-height of the gaussian fall-off in connection weights with respect to distance. This may be approximated as:

$$I = \int_0^\infty 8xe^{\frac{(sx)^2}{2(\sigma_d)^2}}dx.$$

s can thus be solved as a function of I:

$$s = \sqrt{\frac{8(\sigma_d)^2}{I}}.$$

The threshold was then set to G $(1.5{*}s,\sigma_d)$. This means that only elements that are located less than 1.5-times the average separation of the background elements are able to facilitate the post-synaptic cell. Cells that are close but do not have the right orientation will also not be able to facilitate the cell, since the fall-off in the weights with respect to orientation is rather sharp (e.g. $\sigma_\psi = 20$ deg).

To verify this analysis, we simulated this computation by creating a network of the non-orientation specific cells and estimated the average input that the post-synaptic cell would receive with different average separation of the elements ($\sigma_d = 40$). The data are shown in Fig. A1. The model slightly overestimates the separation of the elements but otherwise provides a good match.

Competitive inactivation of connections. The co-axial and trans-axial inputs that exceed the dynamic threshold for each cell then compete for dominance. The facilitatory inputs are segregated into four regions around the cell—two regions for the co-axial connections, one on each end of the cell, and two regions for the trans-axial connections, one on either side of the cell. The suprathreshold inputs in each region are averaged and the sum of the two averaged inputs of the co-axial connections are compared with those from the trans-axial connections. The stronger of the two inhibits the other set of connections such that only one set of connections is active at any one time. This mechanism allows the stimuli to modulate the dominance of one set of connections over the other.

Inhibition. The first stage of the model extracts the cells that are receiving facilitation above the level of noise from the background elements. A longer-latency inhibition forms the next stage of the model. We modeled the inhibition to originate from outside the facilitatory zones, as suggested by the results of Nelson and Frost (1978, 1985) and Kapadia *et al.* (1995):

$$I(x,y,\theta) = \begin{cases} 1, \text{ if } \tan^{-1}\left(\frac{i}{j}\right) - \theta > \kappa^c, \\ 1, \text{ if } \tan^{-1}\left(\frac{j}{i}\right) - \theta > \frac{\pi}{2} \pm \kappa^t, \\ 0, \text{ otherwise.} \end{cases}$$

The magnitude of this inhibition is set such that it is strong enough to suppress cells with weak support but not sufficient to suppress a cell with strong facilitation (Nelson and Frost, 1978, 1985; Kapadia *et al.*, 1995). To implement the inhibition, we assume that each element divides its support equally among the cells which provide it with suprathreshold pre-synaptic inputs. Each of these elements may also receive support from other cells. Inhibition suppresses any cell whose total support is less than a fixed threshold (in all simulations reported, set to 0.5).*

Synchronization and salience. Synchronization is modeled using neural oscillators to represent the cells that enter the bursting mode. The coupling between oscillators is given by the same set of weights that govern facilitation. Oscillators are coupled to other oscillators that have strong, reciprocal, facilitated connections, and are also within a threshold separation, τ, of each other. This threshold is determined by the background separation ($\tau = 2.8\ s$ in all our simulations). The oscillators are initialized with random phases from 0 to 360 deg. The oscillators synchronize using a phase averaging rule:

$$\Theta_i(t) = \frac{\sum \Theta_j(t-1)}{n}$$

where Θ represents the phase of the oscillator and n represents the number of oscillators affecting the phase of oscillator i.

The oscillators synchronize iteratively and synchronization is defined as the following condition:

$$|\Theta_i(t) - \Theta)_j(t)|_{\text{acute}} < \delta,\ i,j, \in C,\ t < t_{\max}$$

where C represents all the coupled oscillators on the same contour, δ represents the maximum phase difference between oscillators, and $t_{\max}$ represents the maximum number of time steps the oscillators are allowed to synchronize. Only if the chain synchronizes does the chain become reliably and coherently represented. The salience, S_c, of the chain is represented by the sum of the activities of all the synchronized elements in the group, C:

$$S_c = \Sigma A_i,\ i \in C$$

A summary of the parameters used in our simulations is presented in the following table:

σ_d^c (distance of the co-axial connections)	40
σ_ψ^c (angle of the co-axial connections)	20 deg
κ^c (fan-out of the co-axial connections)	30 deg
σ_d^t (distance of the trans-axial connections)	30
σ_ψ^t (angle of the trans-axial connections)	20 deg
κ^t (fan-out of the trans-axial connections)	10 deg
σ_d (with distance of the non-oriented connections)	$\sigma_D^c(40)$
Facilitatory threshold	1.5
Threshold separation for coupling	2.8 *s*
Maximum time steps for synchronization	100
Maximum phase difference for synchronization	10 deg

Methods

Experiment 1

In our simulations of Experiment 1, we compared the facilitation in the model to the psychophysical results from Kapadia *et al.* (1995). We computed the facilitation that a vertically oriented cell would receive from a pre-synaptic cell as we varied the pre-synaptic cell's orientation and position. Since there is only one pre-synaptic cell, whose contrast is equal to unity, A(1) reduces to:

$$F(x_{center}, y_{center}, \theta_{center}) = \Gamma(\theta_{centre}, x - x_{center}, y - y_{center}, \psi) \times G\left(\sqrt{(x - x_{center})^2 + (y - y_{center})^2}, \sigma_d^c\right) \times G\left(\psi - \left(2\tan^{-1}\left(\frac{y - y_{center}}{x - x_{center}}\right) - \theta_{center}\right), \sigma_\psi^c\right).$$

The three sets of simulations included:

1. $0 \le y - y_{\text{center}} \le 175$ arcmin
2. $0 \le x - x_{\text{center}} \le 35$ arcmin
3. 90 deg $\le \psi \le$ 160 deg

Our model computes distances based on pixel separation so in order to compare the results with the Kapadia *et al.* (1995) results, we assumed that 10 pixels in the model, which represents the size of a Gabor element, corresponded to 0.12 deg in visual angle. This is consistent with the sizes used in the psychophysical experiments of Field *et al.* (1993) and Kovaćs and Julesz (1993, 1994).

Experiments 3 and 4

In order to concentrate on the problem of modeling the interactions between cortical cells, we assumed that the responses from the filtering stage have been optimized and simplified such that only the cell representing the orientation in the image remains active. We thus used an input that consisted only of the locations and orientations of the Gabor patches present in the image. The Gabor patches are all assumed to be of equal contrast. Since the only orientation with non-zero activity would be the dominant orientation, A(1) simplifies to:

$$F(x,y,\theta) = \iint_{(i,j)\in N} \Gamma(\theta, i, j, \psi) G\left(\sqrt{i^2 + j^2}, \sigma_d\right) G(\psi - \phi(\theta, i, j), \sigma_\psi) di\, dj,$$

where ψ is the orientation of the pre-synaptic cell at position (i,j) relative to the post-synaptic cell at position (x,y) and ϕ is the post-synaptic cell's "preferred" orientation at the same position.

As in Field *et al.* (1993), we generated stimulus arrays consisting of 256 elements in Experiment 3. The contours, if present, were made up of 12 aligned elements, using the same method as Field *et al.* (1993) (see Fig. 7). The remaining stimuli consisted of all randomly oriented elements. Each stimulus was presented in turn to the network and the network selected the contour with the highest salience, i.e., the longest synchronized chain. The stimulus containing the contour with the higher salience was selected as the target. If the contours in both stimuli were equally salient, one of the two was randomly picked to be the target. Five hundred trials were simulated for each data point. The stimuli were spaced at an average distance of 32 pixels, except for the simulation, where we tested the effect of average separation. The separations for that simulation were set to 16, 32 and 58, corresponding approximately to 0.25, 0.5 and 0.9 deg used in the psychophysical experiments. In order to simplify the simulations, most of our simulations were run with only the co-axial connections enabled. Figure 9(a) shows that the results do not change when both sets of connections are active (the curves labeled Model* are from simulations with both sets of connections active, while the curves labeled Model are with only the co-axial connections enabled).

For Experiment 4, we used stimuli similar to Kovács and Julesz (1993, 1994). The stimulus arrays contained 2025 elements each. Contour elements were made up of 24 Gabor patches. Open contours were generated with relative orientations chosen randomly from the range of ±30 deg. Closed contours were generated with elements placed on a circle at every 15 deg±5 deg. To determine φ, the background elements were spaced at 25 pixels, while the separation of the contour elements was in the range of 20–45 pixels, corresponding

*It is possible to express the effects of facilitation and inhibition in a form consistent with the steering equation. This would allow both local and long-range inputs to be computed analytically (E. Simoncelli, personal communication).

to φ in the range of 1.25–0.556. The threshold separations for open and closed contours were defined to be at 75% correct identification. To simulate the changes in salience, additional elements were added to a short open contour and an incomplete closed contour, spaced at φ_0 and φ_c, respectively. The performance of the network was then computed for each element added. Again, 500 trials were simulated for each data point.

Experiment 5

Images were convolved with H2 steerable filters at one spatial frequency to extract edge information. The dominant orientation at each location was then computed directly from the responses of the filters, as in Freeman and Adelson (1991). This allowed us to generate a map of dominant orientations, as well as the activity at the dominant orientation. This activity was used to threshold the image such that facilitation is only computed for the image locations with supra-threshold filter responses. Each supra-threshold filter then sums the facilitation from each of its neighbors by computing the activity at the "preferred" orientation. This is done by steering the steerable filters at each of the neighboring locations to the "preferred" orientation. In order to preserve the orientation information, the facilitation acts by scaling the basis responses proportionally. The facilitation is then followed by a global normalization to re-scale the responses of all the filters into the range of 0.0–1.0. This normalization is similar to Heeger (1992), where each cell's response is divided by the average of the cell activities in the scene. The activity after every cycle is also thresholded so the facilitation is computed only for the cells with supra-threshold activity. We allowed the process to iterate repeatedly until the network stabilizes.