

# Cortical Synchronization Mechanism for “Pop-Out” of Salient Image Contours

Shih-Cheng Yen and Leif H. Finkel

Department of Bioengineering and Institute of Neurological Sciences  
University of Pennsylvania  
Philadelphia, PA 19104, U. S. A.  
syen@jupiter.seas.upenn.edu  
leif@jupiter.seas.upenn.edu

## ABSTRACT

We present a model based on long-range intra-cortical connections which computes the salience of contours in a visual scene. The model accounts for a number of psychophysical and physiological results on contour salience, and provides a mechanism for several of the Gestalt laws of perceptual organization. In the model, cells lying on smooth contours facilitate each other, and strongly facilitated cells enter a “bursting” model. Horizontal connections allow bursting cells to synchronize, and perceptual salience is defined by the level of synchronized activity. In particular, we propose that the intrinsic properties of synchronization account for the increased salience of smooth, closed contours

## 1. INTRODUCTION

It has been suggested that cells in the supragranular layers of visual cortex with long-range horizontal connections might play a role in extracting salient features in a scene (Gilbert, 1992; Field *et al.*, 1993; Kovács and Julesz, 1993). These cells have been shown to be sensitive to stimuli outside the classical receptive field, allowing contextual information to influence the response of the cell (Nelson and Frost, 1985; Kapadia *et al.*, 1995). Similar cells in the supragranular layers of striate cortex have been observed to burst rapidly (Gray and McCormick, 1996) and could be involved in the temporal binding of contour elements. We present a model for computing the perceptual salience of contours that incorporates these two findings and is able to account for a number of physiological and psychophysical results (Polat and Sagi, 1993, 1994; Kapadia *et al.*, 1995; Field *et al.*, 1993; Kovács and Julesz, 1993, 1994; Kovács *et al.*, 1996).

## 2. MODEL ARCHITECTURE

Linear quadrature steerable filter pyramids (Freeman and Adelson, 1991) are used to model the response characteristics of cells in primary visual cortex. Steerable filters are computationally efficient as they allow the energy at any orientation and spatial frequency to be calculated from the responses of a set of basis filters. The fourth derivative of a Gaussian and its Hilbert transform were used as the filter kernels to approximate the shape of the receptive fields of simple cells.

Model cells are interconnected by long-range horizontal connections in a pattern similar to the co-circular connectivity pattern of Parent and Zucker (1989), as well as the “association field” proposed by Field *et al.* (1993). For a cell of orientation  $\theta_A$  at location “A”, there is a “preferred” orientation at location “B”,  $\theta_B$ , 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,  $\theta_A$ , at “A”. The connection weights between the cell with orientation  $\theta_A$ , at “A” and the oriented cells at “B” peak at  $\theta_B$ , and decrease with increasing angular difference between the two orientations. These 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 orientation (trans-axial). There is physiological and anatomical evidence consistent with the existence of both sets of connections (Rockland and Lund, 1983; Lund *et al.*, 1985; Nelson and Frost, 1985; Kapadia *et al.*, 1995; Fitzpatrick, 1996). The connection field is shown in Figure 1a. As observed physiologically, these excitatory connections only facilitate cells that receive local supra-threshold input. If the local orientation activity distribution at “B” peaks at  $\theta_B$ , the cell with orientation  $\theta_A$  at “A” will be strongly facilitated. As the local orientation at “B” deviates from  $\theta_B$ , the degree of facilitation decreases. The “preferred” orientation at “B” can thus be thought of as providing “support” for the orientation,  $\theta_A$ , at

“A”. Connection weights decrease for positions with increasing angular deviation from the orientation axis of the cell, as well as positions with increasing distance, in agreement with the physiological and psychophysical findings.

It is estimated that 20% of the horizontal connections target inhibitory cortical cells (McGuire *et al.*, 1991), and inhibition also occurs through di-synaptic pathways. Intra-cellular recordings have shown that stimulation of the horizontal connections produces short-latency EPSPs followed by longer-latency IPSPs (Weliky *et al.*, 1995). Since the mechanisms underlying long-distance inhibition are less well documented, we have chosen a simple descriptive mechanism. Each facilitated cell inhibits all other cells within the range of its long-distance connections such that only the elements that are strongly facilitated survive. This is consistent with neurophysiological evidence from Kapadia *et al.* (1995).

In the model, cells that are strongly facilitated are assumed to enter a “bursting” mode, similar to the behavior of “chattering” cells described by Gray and McCormick (1996). We assume that bursting cells are able to synchronize with other bursting cells. Cells that enter the “bursting” mode are modeled as homogeneous coupled neural oscillators with a common fundamental frequency but different phases (Kopell and Ermentrout, 1986; Baldi and Meir, 1990). The phase of each oscillator is modulated by the phase of the other oscillators to which it is coupled. Oscillators are coupled only if the corresponding cells have strong, reciprocal, connections. A set of coupled oscillators together represent a contour. Since oscillators on different contours are not generally inter-connected, each contour in the scene can synchronize independently, and cells representing different contours to be discriminated and grouped together. It has been postulated that gamma frequency (40 Hz) oscillations often observed in the cortex could be responsible for generating perceptual binding across different cortical regions (Gray and Singer, 1995). Recent studies have questioned the functional significance and even the existence of these oscillations (Ghose and Freeman, 1992; Bair *et al.*, 1994). We use neural oscillators only as a simple means of computing synchronization and assume that synchronization may be achieved through a different mechanism in cortex.

The salience of a contour can only be computed when all the oscillators on the contour are synchronized. Salience is 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. This allows us to compare the results of the model with a number of psychophysical findings. All simulations were conducted with the same parameter set.

### 3. RESULTS

#### 3.1. Co-Axial Connection Pattern (Kapadia *et al.*, 1995)

The experiments of Kapadia *et al.* (1995) provide a test of the connection architecture. The degree of facilitation of post-synaptic cells was measured as a function of the position and orientation preference of a horizontally connected pre-synaptic cell. Figure 1 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. The results show good qualitative agreement with the experimental findings. The results are also in qualitative agreement with those of Polat and Sagi (1993). Our simulations do show an overestimation of the activation for closely spaced elements (Figure 1b) 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. There is also some discrepancy between the model and the data at large off-axial misalignments. 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.

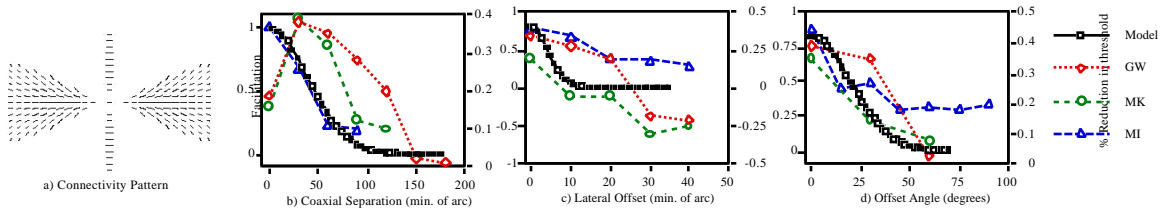


Figure 1. From left to right: Connectivity pattern of a horizontally oriented cell. Length of line indicates connection strength; The results are compared to the psychophysical data from 3 subjects (GW, MK, MI) reported in Kapadia *et al.* (1995).

### 3.2. Contrast Sensitivity Modulation (Kovács and Julesz, 1993, 1994)

The trans-axial connections in our model may underlie a surprising psychophysical observation 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 and elliptical contour. The contour itself was formed from aligned Gabor patches. They found a sharp 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 2 on each side of the contour, while the sensitivity on the contour itself was greatly decreased as compared to the sensitivity to the target in the absence of the contour. Figure 2b,c shows the contrast sensitivity maps from Kovács and Julesz (1993). Figure 2a shows a simplified "silhouette" of the connectivity pattern for a horizontally oriented cell in our model. The gray level represents the connection weights; dark regions are facilitatory, and white regions are inhibitory. Since the psychophysical experiments were carried out using low-contrast probes oriented parallel to the closest element on the contour, only the trans-axial connections are likely to be stimulated. We have thus left out 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 (Figure 2d,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 2, 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's (1994) data.

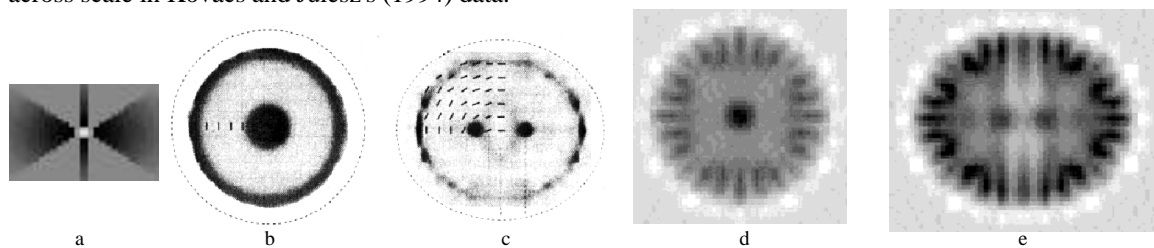


Figure 2. a) Simplified silhouette showing intra-filter inhibition and the facilitatory connections for a horizontally oriented cell. Dark regions are facilitatory, white regions are inhibitory and gray regions are neutral. b,c) Contrast sensitivity maps inside a circular and elliptical contour, from Kovács and Julesz (1994). d,e) Sensitivity maps from the model based on averaged connection fields of aligned Gabor units.

### 3.3. Extraction of Salient Contours (Field *et al.*, 1993)

Using the same methods as Field *et al.* (1993), we tested the model's ability to extract contours embedded in noise (see Figure 3). 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 saliency. Of the two stimuli in each pair presentation, the network “chooses” the stimulus containing the contour with the higher saliency. 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: 1) the angle,  $\theta$ , between elements on a contour, 2) the angle between elements on a contour but with the elements aligned orthogonal to the contour passing through them, 3) the angle between elements with a random offset angle,  $\pm \theta$ , with respect to the contour passing through them, and 4) average separation of the elements. 500 simulations were run at each data point. The results are shown in Figure 3. The model shows good qualitative agreement with the psychophysical data. When the elements are aligned, the performance of the network is mostly modulated by the co-axial connections, whereas when the elements are oriented orthogonal to the contour, the trans-axial connections mediate the performance of the network. The performance of both the model and human subjects are adversely affected as the weights between consecutive elements decrease in strength. This reduces the length of the contour and thus the saliency of the stimulus.

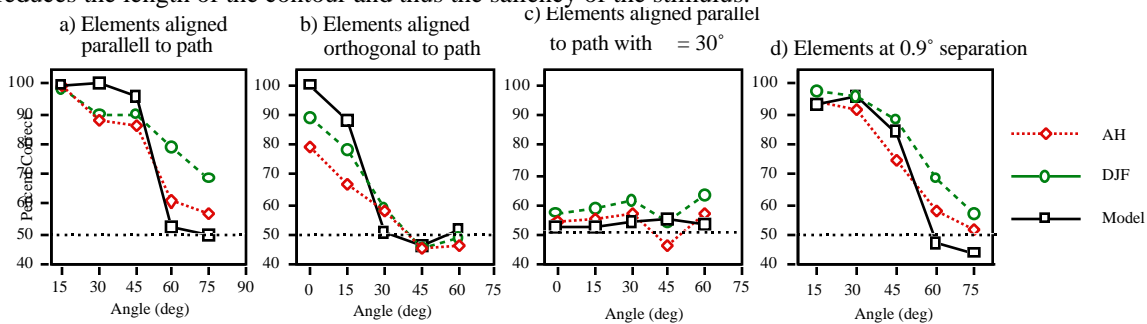


Figure 3. Simulation results are compared to the data from 2 subjects (AH, DJF) in Field *et al.* (1993). Stimuli consisted of 256 randomly oriented Gabor patches with 12 elements aligned to form a contour. Each data point represents results for 500 simulations.

### 3.3. Effects of Contour Closure (Kovács and Julesz, 1993, 1994)

In a series of experiments using similar stimuli to Field *et al.* (1993), Kovács and Julesz (1993) found that closed contours are much more salient than open contours. They found that when the inter-element spacing between all elements was gradually increased, the maximum inter-element separation for detecting closed contours ( $s_c$ , defined at 75% performance) is higher than that for open contours ( $s_o$ ). In addition, they showed that when elements spaced at  $s_o$  are added to a “jagged” (open) contour, the saliency of the contour increases monotonically but when elements spaced at  $s_c$  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  $s_c$ , the contour is not salient until it is closed, at which point it suddenly “pops-out” (see Figure 4c). 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 and Ermentrout, 1986; Ermentrout, 1985; Somers and Kopell, 1993). This property seems to be related 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 stronger the coupling, the better the synchronization, both in terms of speed and coherence (Somers and Kopell, 1993; Wang, 1995). We believe these findings may apply to the psychophysical results.

As in Kovács and Julesz (1993), we generated stimulus arrays containing 2025 elements. Contours were made up of 24 elements. Again the network is presented with two stimuli, one containing a contour and the other made up of all randomly oriented elements. The network picks the stimulus containing the synchronized contour with the higher saliency. In separate trials, the threshold separation for open and closed contours were determined. The ratio of the separation of the background elements to the that of elements on a closed curve,  $s_c/s_o$ , was found to be 0.6 (which is similar to the threshold of 0.65 recently reported by Kovács *et al.*, 1996), whereas the ratio for open contours,  $s_o/s_c$ , was found to be 0.9. ( $s_c$  is the

threshold separation of contour elements,  $\sigma$ , at a particular background separation). We then examined the changes in saliency for open and closed contours. The performance of the network was measured as additional elements were added to an initial short contour of elements. The results are shown in Figure 4b. At  $\sigma_c$ , both open and closed contours are synchronized but at  $\sigma_o$ , elements are synchronized only when the chains are closed. If saliency can only be computed for synchronized contours, then as additional elements are added to an open chain at  $\sigma_o$ , the saliency would increase since the whole chain is synchronized. On the other hand, at  $\sigma_c$ , as long as the last element is missing, the chain is really an open chain, and since  $\sigma_c$  is smaller than  $\sigma_o$ , the elements on the chain will not be able to synchronize and adding elements has no effect on saliency. Once the last element is added though, the chain is immediately able to synchronize and the saliency of the contour increases dramatically and causes the contour to “pop-out”.

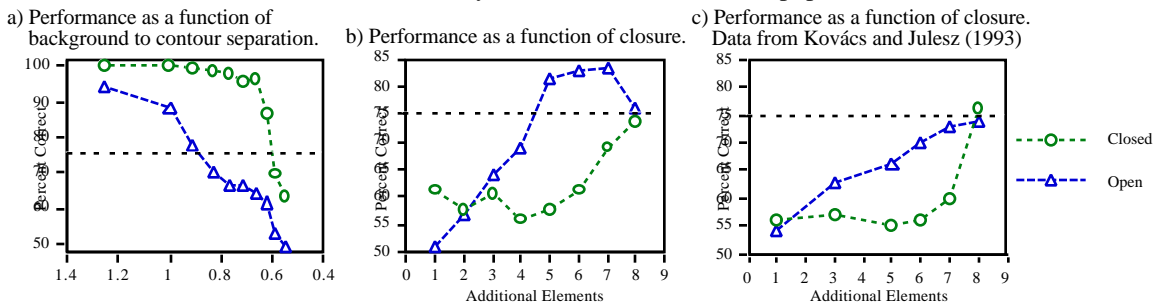


Figure 4. Simulation of the experiments of Kovács and Julesz (1993). Stimuli consisted of 2025 randomly oriented Gabor patches, with 24 elements aligned to form a contour. Each data point represents results from 500 trials. a) Plot of the performance of the model with respect to the ratio of the separation of the background elements to the contour elements. Results show closed contours are salient to a more salient than open contours. b) Changes in saliency as additional elements are added to open and closed contours. Results show that the saliency of open contours increase monotonically while the saliency of closed contours only change with the addition of the last element. Open contours were initially made up of 7 elements while closed contours were made up of 17 elements. c) The data from Kovács and Julesz (1993) are re-plotted for comparison.

#### 4. CONCLUSION

We have presented a cortically-based model that is able to identify perceptually salient contours in images containing high levels of noise. The model is based on the use of long distance intra-cortical connections that facilitate the responses of cells lying along smooth contours. Saliency is defined as the combined activity of the synchronized population of cells responding to a particular contour. The model qualitatively accounts for a range of physiological and psychophysical results and can be used in extracting salient contours in real images (Yen and Finkel, 1996).

#### ACKNOWLEDGMENTS

Supported by the Office of Naval Research (N00014-93-1-0681), The Whitaker Foundation, and the McDonnell-Pew Program in Cognitive Neuroscience.

#### 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.

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.
- Gilbert, C. D. (1992). Horizontal integration and cortical dynamics. *Neuron*, 9, 1-20.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- Parent, P. & Zucker, S. W. (1989). Trace inference, curvature consistency, and curve detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11, 823-839.
- 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.
- Rockland, K. S. & Lund, J. S. (1983). Intrinsic laminar lattice connections in primate visual cortex. *Journal of Comparative Neurology*, 216, 303-318.
- 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.
- Wang, D. (1995). Emergent synchrony in locally coupled neural oscillators. *IEEE Transactions on Neural Networks*, 6, 941-948.
- 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.
- Yen, S-C. & Finkel, L. H. (1996). Salient Contour Extraction by Temporal Binding in a Cortically-Based Network. In Touretzky, D. S., Mozer, M. C. & Hasselmo, M. E. (Eds), *Advances in Neural Information Processing Systems 9*. Massachusetts: MIT Press.