A neural model of contour integration in the primary visual cortex[1]

Zhaoping Li
Hong Kong University of Science and Technology
Clear Water Bay, Hong Kong

Published in *Neural Computation* **10** 903-940, 1998.

Abstract

Experimental observations suggest that contour integration may take place in V1. However, there has yet to be a model of contour integration that only uses known V1 elements, operations, and connection patterns. This paper introduces such a model, using orientation selective cells, local cortical circuits, and horizontal intra-cortical connections. The model is composed of recurrently connected excitatory neurons and inhibitory interneurons, receiving visual input via oriented receptive fields resembling those found in primary visual cortex. Intracortical interactions modify initial activity patterns from input, selectively amplifying the activities of edges that form smooth contours in the image. The neural activities produced by such interactions are oscillatory and edge segments within a contour oscillate in synchrony. It is shown analytically and empirically that the extent of contour enhancement and neural synchrony increases with the smoothness, length, and closure of contours, as observed in experiments on some of these phenomena. In addition, the model incorporates a feedback mechanism which allows higher visual centers to selectively enhance or suppress sensitivities to given contours, effectively segmenting one from another. The model makes the testable prediction that the horizontal cortical connections are more likely to target excitatory (or inhibitory) cells when the two linked cells have their preferred orientation aligned with (or orthogonal to) their relative receptive field center displacements.

[1]I thank Jochen Braun for introducing me to the topic and helping me to appreciate the difficulties for dynamic stability in models of cortical circuits. I also thank Peter Dayan for many helpful conversations, discussions, and comments on the manuscript, and Yair Weiss, Shimon Ullman, Mriganka Sur, Geoffrey Hinton, and William Freeman for discussions and help on references. This work was supported by the Hong Kong Research Grant Council and the Center for Biological and Computational Learning at MIT.

1. Introduction

In early stages of the visual system, individual neurons are responsive only to stimuli in their classical receptive fields (RFs), which are only large enough to signal a small edge or contrast segment in the input (Hubel and Wiesel, 1962). The visual system must group separate local input elements into meaningful global features to infer the visual objects in the scene. Sometimes local features group together into regions, as in texture segmentation; at other times they group into contours which may represent boundaries of underlying objects. Although much is known about the early visual processing steps that extract local features such as oriented edges, it is still unclear how the brain groups local features into global and more meaningful features. In this study, we model the neural mechanisms underlying the first stages of grouping of edge elements into contours — namely contour integration.

One of the first problems for contour grouping is the abundance of candidate "edges" produced by the simple edge detection mechanism that is believed to operate in V1 (Marr 1982). Many of these "edges" are simply image contrast "noise" and are unlikely to belong to any significant or relevant object contour. It is desirable to influence the response of edge detectors by contextual information from the surround to enhance the sensitivity to more relevant edges. This could be the first step towards perceptual contour grouping. Indeed, V1 cells are observed to change their responses or sensitivities depending on the surround stimulus (Knierim and Van Essen 1992, Kapadia, Ito, Gilbert and Westheimer 1995), cells are more responsive if they are stimulated by edges that are aligned with other edge elements outside their RFs (Kapadia et al 1995). These observations correspond well with psychophysical observations that human sensitivity to edge segments is also higher when they are aligned with other edges (Polat and Sagi 1994, Kapadia et al 1995). Horizontal cortical connections linking cells of non-overlapping RFs have been observed and hypothesized as the underlying neural substrate (Rockland and Lund 1983, Gilbert 1992). These findings suggest that simple and local neural interactions even in V1 could contribute to primitive visual perceptual grouping as in contour integration, although V1 cells have been observed to change their sensitivities by visual attention (Motter 1993).

It is computationally desirable to understand how V1, with its small size receptive fields and finite range neural interactions, can nevertheless enhance sensitivities to globally defined contours. Models of visual algorithms have been proposed for similar purposes. For example, Zucker, Dobbins, and Iverson (1989) used relaxation labeling techniques (Hummel and Zucker 1983) to perform global curve detection or inference using only local compatibility interactions between curve segment labels (see discussion). However, modeling the contour enhancement using known V1 neural RFs and interactions is difficult and has yet to be accomplished. It is thus not clear if all the experimentally observed contour enhancement (Kapadia et al 1995) has to be attributed to feedback from higher visual areas. The difficulty in the modeling largely stems from the dynamic stability problems associated with recurrently connected neural networks — the system has to be sensitive enough to enhance any possible smooth contours the input may provide, but selective and stable enough so that noise is not significantly enhanced. To overcome this difficulty, previous neural models of contour enhancement have assumed some non-neural or biologically questionable operations (see discussion for more details) such as dipole fields and the requirement of feedback from higher centers (Grossberg and Mingolla 1985), non-linear rule-based interactions (Zucker et al 1989, Heitger and von der Heydt 1993), input gating (Braun Niebur, Schuster and Koch 1994), and global activity normalizations (Yen and Finkel 1996). In fact, modeling contour enhancement is difficult even without the constraints of the neural hardware. Many computer vision models of edge linking (e.g., Kass, Witkin, and Terzopoulos 1988) need user intervention, and many more autonomous models (e.g., Shashua and Ullman 1988, Guy and Medioni 1993, Williams and Jacobs 1996) all suffer from one or another problem (see discussion for more details). It is thus desirable to find out whether contour enhancement can actually be modeled using just V1 neural elements and operations, or whether contour enhancement in V1 has to be totally attributed to top-down feedback.

This paper introduces a model of contour enhancement using only V1 elements, based on experimental findings (detailed in the next section), such as orientation selective cells, local recurrent neural circuits, and finite range horizontal connections. The model is studied analytically and empirically to understand how sensitivity enhancement in long range contours is successfully carried out in a network of neurons with finite range interactions. The network dynamics are analyzed to reveal the temporal synchrony between cells within a contour, as observed in experiments (Gray and Singer 1989, Eckhorn et al 1988). Our analysis relates the extent of the contour enhancement and neural synchrony with the contour characteristics such as length, curvature, and closure. The model makes a testable prediction about the horizontal connection structure: the post-synaptic cell type is more likely to be excitatory (or inhibitory) if two cells linked by the horizontal connection prefer orientations that are aligned (or orthogonal) to the relative displacement between their RF centers. In addition, this model introduces a mechanism that allows higher visual centers selectively to enhance or suppress contour sensitivities, in addition to the contour enhancement performed within V1.

We like to emphasize that our work is mainly aimed at modeling the aspects of contour enhancement that are observed in V1. Contour integration is most likely completed by higher visual centers, which are absent in the present model. This is in contrast to many other models which aim to build a model with the best possible performance at contour integration rather than to understand how and where it is done in the brain. For instance, this model therefore does not address or define illusory contours, since V1 cells are not as evidently responsive to illusory contours as V2 cells (von der Heydt, Peterhans, Baumgartner 1984, Grosof, Shapley, and Hawken 1993), and T, L, X junction units, which are not known to exist in V1, are required to detect many types of illusory contours. (However, our model does help to fill in the gaps in incomplete contours — see section

3). Also, assuming that V1 does not address contours as global objects, the model merely enhances individual contour segments, without defining the saliency of a whole contour. Additionally, a mechanism of feedback control is provided by modeling the feedback signals and specifying their V1 target neurons, but we do not actually model how higher visual centers might respond to V1 outputs in order to construct the desired feedback.

The paper is organized as follows. First, we present the relevant experimental background. Then we describe the model and analyze it to show how contours are enhanced, how synchronization happens between contour segments, and how they depend on contour characteristics. The performance of the model is demonstrated by examples. Then we model the top-down feedback and demonstrate selective enhancement, suppression, and the effective segmentation by top-down control. Finally, we place the model in the context of experimental findings and other models, and discuss its limitations and possible extensions.

2. Experimental background

Primary cortical neurons respond to input edges only within their classical receptive fields, which are local regions in the visual field mostly too small to contain any visual object (Hubel and Wiesel 1962). RF centers are distributed visuotopically on the cortical surface and cells with overlapping RFs but different preferred edge orientations are grouped together into hypercolumns (Hubel and Wiesel 1962). Visual stimuli outside a cell's classical RF, in a region whose size is larger than the RF, can influence the responses of the cell (see review by Allman, Miezin, and Mcguinness 1985, Kinerim and Van Essen 1992). Generally, antagonistic suppression is observed when gratings or textures are presented in the surround (Allman et al 1995), although surround facilitation (Maffei and Fiorentini, 1976) and orientation contrast facilitation have also been observed (Sillito, Grieve, Jones, Cudeiro, and Davis 1995). By placing bars in the surround of the RF of a cell and roughly aligning them with a bar presented in the center in its preferred orientation, Kapadia et al (1995) demonstrated a significant increase in response to the central bar, even when there are additional random stimuli in the background. This enhancement of response decreases with increasing separation or misalignment between the central and surround bars, and is stronger when multiple bars in the surround are aligned with the central bar to generate a smooth contour (Kapadia et al 1995). Such contextual influences will be the mechanism used for the contour enhancement in our model. Qualitatively similar findings are observed psychophysically in humans under similar stimulus settings (Polat and Sagi 1994, Kapadia et al 1995). Human observers can also easily identify a smooth contour composed of individual or even disconnected edge segments among other random edge segments scattered in the background (Field, Hayes, and Hess 1993). The sensitivity to such contours is further enhanced when the contour closes on itself — this is called the closure effect (Kovacs and Julesz 1993). Also, responses of V1 cells are modulated by visual attention (Motter 1993), although earlier studies found these effects only in higher visual areas (Moran and Desimone 1985).

Primary visual cortex is composed of many excitatory and inhibitory neurons. Visual input is received mainly onto the dendrites of the excitatory cells, although inhibitory neurons also receive a very small amount of direct visual input (White 1989, Douglas and Martin 1990). The excitatory cells send outputs to higher visual areas (Douglas and Martin 1990) which in turn feed back to V1 (Valverde 1985, van Essen 1985). There are several layers and different groups of excitatory cells (and inhibitory cells), and they are likely to serve different functions, some cells are more concerned with receiving visual inputs whereas other cells process the signals and send outputs to higher visual centers (Salin and Bullier 1995). It is not yet clear what the target cell types are for the higher center feedback (Salin and Bullier 1995). Cortical neurons interact with each other locally and often reciprocally; the excitatory connections extend somewhat longer distances than the inhibitory ones (Douglas and Martin 1990, White 1989). These neural interactions typically link neurons with similar RF properties (White 1989).

The anatomical basis for the surround effect has been postulated to be the long-range horizontal connections linking cells up to 4 mm or more apart in the primary visual cortex (e.g., Kapadia et al 1995, Gilbert 1992, Allman et al 1985). These connections emanate from the excitatory pyramidal cells in upper layers and contact both the excitatory and inhibitory postsynaptic cells, enabling monosynaptic excitation and disynaptic inhibition from one cortical site to another (Gilbert 1992, McGuire, Gilbert, Rivlin and Wiesel 1991, Hirsch and Gilbert 1991, Weliky, Kandler, Fitzpatrick, and Katz 1995). The axonal fields of these connections are asymmetrical, extending for greater distances along one cortical axis than another (Rockland and Lund 1983, Gilbert and Wiesel 1983, Fitzpatrick 1996b). Cells preferring similar orientations tend to be linked (Ts'o, Gilbert, and Wiesel 1986, Gilbert and Wiesel 1989, Malach, Amir, Harel, and Grinvald 1993) whether or not the relative displacements of their receptive field centers are aligned with or orthogonal to their preferred orientations (Gilbert and Wiesel 1983).

The horizontal cortical connections are also implicated in the temporal synchrony of the 40-60 Hz oscillations of neural responses (Gray and Singer 1989, Eckhorn et al 1988). Take two neurons with non-overlapping RFs and aligned optimal orientations. The synchrony of their firing is negligible if two bars sweep over the two RFs independently, is significant when the bars sweep together, and is the strongest when a long single sweeping bar extends over both RFs (Singer and Gray 1995). Usually the degree of neural synchrony decreases with increasing separation between neurons (Singer and Gray 1995, Eckhorn 1994). The extent of the oscillatory neural activities is not completely certain (Singer and Gray 1995). It will be shown in our model that for inputs that contain contours, the strength of the neural oscillation depends on contour characteristics such as length and smoothness. Both the synchrony and the enhancement of responses for aligned edges have been postulated as mechanisms underlying feature linking (Gilbert 1992, Singer and Gray 1995, Eckhorn 1994).

3. The contour integration model

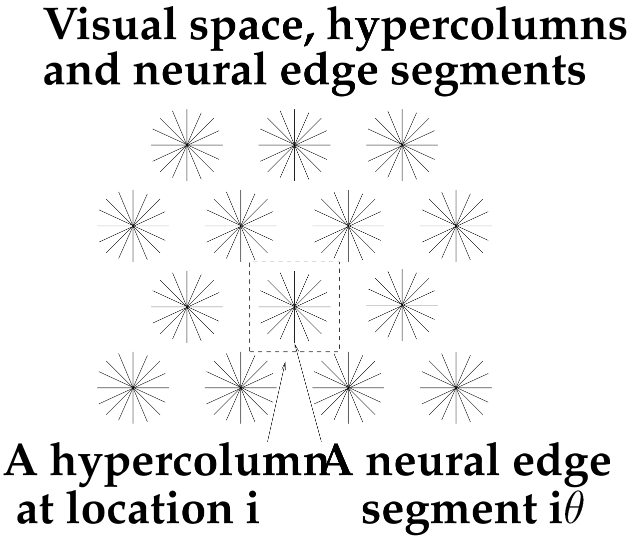

Figure 1: Model visual space, hypercolumns, and edge segments. The input space is a discrete hexagonal or Manhatten grid.

This section starts by outlining the model and its neural elements. Then the response of a single edge element is analyzed, and the enhancement of responses that results from interactions between the elements is analyzed and demonstrated. Next, the dynamics of the model are studied to reveal a tendency towards oscillations and the emergence of temporal coherence between contour elements. Lastly, we introduce and demonstrate the mechanism which allows top-down feedback control.

3.1 Model outline

Visual inputs are modeled as arriving at discrete spatial locations (Fig. 1). At each location i there is a model V1 hypercolumn composed of K neuron pairs. Each pair (i, θ) has RF center i and preferred orientation $\theta = k\pi/K$ for $k = 1, 2, ...K$, and is called (a neural representation of) an edge segment. Each edge segment consists of an excitatory and an inhibitory neuron that are connected with each other. The excitatory cell receives the visual input; its output quantifies the response or salience of the edge segment and projects to higher visual areas. The inhibitory cells are treated as interneurons. When an input image contains an edge at i with orientation β and input strength $\hat{I}_{i\beta}$, edge segment $i\theta$ receives input $I_{i\theta} = \hat{I}_{i\beta}\phi(\theta - \beta)$, where $\phi(\theta - \beta) = e^{-|\theta-\beta|/(\pi/8)}$ is the orientation tuning curve for the cell centered at β. Segments outside the hypercolumn i receive no input contribution from $\hat{I}_{i\beta}$.

The excitatory and inhibitory cells have membrane potentials $x_{i\theta}$ and $y_{i\theta}$, respectively, and their outputs are $g_x(x_{i\theta}) \geq 0$ and $g_y(y_{i\theta}) \geq 0$, analog numbers modeling firing rates. Both $g_x(.)$ and $g_y(.)$ are sigmoid-like non-linear and non-decreasing functions (Fig. 2). The cells for an edge segment send their outputs $g_x(x_{i\theta})$ and $g_y(y_{i\theta})$ to each other. The excitatory cell excites itself with synaptic strength J_o. Its output $g_x(x_{i\theta})$ is the edge's response to visual inputs. Edge segment $j\theta'$ at another location can excite edge $i\theta$ monosynaptical-

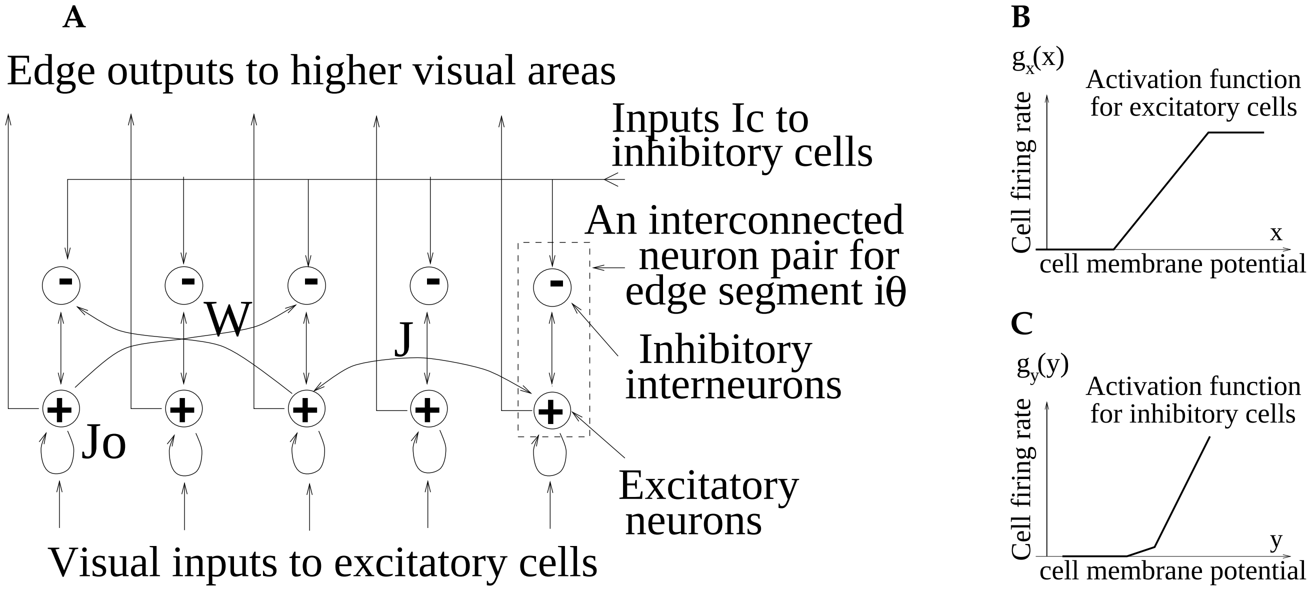

Figure 2: **A:** Model neural elements, edge elements, visual inputs, and neural connections. J_o: the self-excitatory connection, J the lateral excitatory connection between edge elements, W the lateral disynaptic inhibitory connection between edge elements, implemented as excitatory connections from the excitatory neuron of one edge element to the inhibitory neurons of the others. **B:** Activation function $g_x(x)$ for the excitatory cells and **C:** $g_y(y)$ for the inhibitory cells.

ly by sending the excitatory signal $J_{i\theta,j\theta'}g_x(x_{j\theta'})$ to the excitatory cell in edge $i\theta$, and/or inhibit the edge disynaptically by directing an excitatory signal $W_{i\theta,j\theta'}g_x(x_{j\theta'})$ to the inhibitory cell. Here $J_{i\theta,j\theta'}$ and $W_{i\theta,j\theta'}$ model the synaptic strengths of horizontal cortical connections. For any visual input pattern $I_{i\theta}$ for all i, θ, the neural dynamics evolve according to:

$$\dot{x}_{i\theta} = -\alpha_x x_{i\theta} - \sum_{\Delta\theta} \psi(\Delta\theta) g_y(y_{i,\theta+\Delta\theta}) + J_o g_x(x_{i\theta}) + \sum_{j\neq i,\theta'} J_{i\theta,j\theta'} g_x(x_{j\theta'}) + I_{i\theta} + I_o \quad (1)$$

$$\dot{y}_{i\theta} = -\alpha_y y_{i\theta} + g_x(x_{i\theta}) + \sum_{j\neq i,\theta'} W_{i\theta,j\theta'} g_x(x_{j\theta'}) + I_c \quad (2)$$

where $1/\alpha_x$ and $1/\alpha_y$ are the membrane time constants, I_c is the background input to the inhibitory cells which will later be used to model top-down control signal. I_o is the background input to the excitatory cells and includes a term that normalizes the activity — an inhibition that increases with the total activities in the local edge segments. Finally, $\psi(\Delta\theta)$ is an even function of $\Delta\theta$ modeling inhibition within a hypercolumn, and decreases with $|\Delta\theta|$. When $\psi(\Delta\theta) = 0$ for $\Delta\theta \neq 0$, the inhibitory cells couple only to the excitatory cell in the same edge segment, otherwise, an activated edge exerts some inhibition $\propto \psi(\Delta\theta)$ to other edges in the same hypercolumn (note that this interaction within an hypercolumn does not model the emergence of the cell orientation selectivity (Somers, Nelson, and Sur 1995)). Each neuron additionally receives some random noise input. The appendix lists the parameters used in this model. For ease of analysis, and without loss of generality, we use $\alpha_x = \alpha_y = 1$, and make $g_x()$ and $g_y()$ piecewise linear functions with threshold and saturation. Also, the excitatory cells have a unit gain $g'_x(x)$ in the operating range.

Given an input pattern $I_{i\theta}$, the network approaches a dynamic state after several membrane time constants and the response $g_x(x_{i\theta})$ gives a saliency map. When $g_x(x_{i\theta})$ at location i is a unimodal function of θ (identifying θ with $\theta + \pi$), the orientation $\bar{\theta}_i$ that would be perceived in higher centers can be modelled by $e^{i2\bar{\theta}_i} \propto \sum_\theta g_x(x_{i\theta})e^{i2\theta} / \sum_\theta g_x(x_{i\theta})$. Two edges of different orientations will be perceived to cross each other at location i if $g_x(x_{i\theta})$ is a bimodal function of θ.

3.2 A single edge element

Before studying the contextual interactions between edges, we first analyze the input response properties of a single edge segment $i\theta$, ignoring the other edges. For simplicity, we omit the subscripts i and θ, and denote $I = I_{i\theta} + I_o$:

$$\dot{x} = -x - g_y(y) + J_o g_x(x) + I \quad (3)$$

$$\dot{y} = -y + g_x(x) + I_c \quad (4)$$

The average neural activity is determined by the equilibrium point $E = (\bar{x}, \bar{y})$, which is the intersection of the two curves on which $\dot{x} = 0$ and $\dot{y} = 0$, respectively (Fig. 3B):

$$\dot{\bar{x}} = 0 = -\bar{x} - g_y(\bar{y}) + J_o g_x(\bar{x}) + I \quad (5)$$

$$\dot{\bar{y}} = 0 = -\bar{y} + g_x(\bar{x}) + I_c \quad (6)$$

Increasing I or I_c raises or lowers the output $g_x(\bar{x})$. The input sensitivities, determined by solving the linearized version of the above equations, are

$$\delta g_x(\bar{x})/\delta I = \frac{g'_x(\bar{x})}{1 + g'_y(\bar{y})g'_x(\bar{x}) - J_o g'_x(\bar{x})} \tag{7}$$

$$\delta g_x(\bar{x})/\delta I_c = \frac{-g'_y(\bar{y})g'_x(\bar{x})}{1 + g'_y(\bar{y})g'_x(\bar{x}) - J_o g'_x(\bar{x})} = -g'_y(\bar{y}) \cdot [\delta g_x(\bar{x})/\delta I] \tag{8}$$

In cases that interest us, the output $g_x(\bar{x})$ increases continuously with visual input I (Fig. 3A), i.e., $\delta g_x(\bar{x})/\delta I \geq 0$ and this requires $g'_y(\bar{y})g'_x(\bar{x}) > J_o g'_x(\bar{x}) - 1$. The visual input gain $\delta g_x(\bar{x})/\delta I$ is zero below input threshold where $g'_x = 0$, becomes non-zero after threshold where $g'_x = 1$, and can decrease beyond threshold when the inhibitory gain $g'_y(\bar{y})$ increases. At high I the activity $g_x(\bar{x})$ saturates when $g'_x = 0$ again. Qualitatively this sensitivity curve corresponds to physiological observations[2]. Note that threshold input value I and the edge input response curve in Fig. 3A change with I_c. This is because, as shown in Equation (8), increasing I_c decreases the output $g_x(\bar{x})$.

According to equations (7) and (8), increasing I and I_c simultaneously increases the output $g_x(\bar{x})$ if $\Delta I/\Delta I_c > g'_y(y)$, and decreases it otherwise. This leads to the following consequences. First, the visual input could be directed to both the excitatory (increasing I) and inhibitory (increasing I_c) cells as experimentally observed (White 1989). Their net effect will be to increase the edge response $g_x(\bar{x})$ as long as the visual input partition to the two cell types is appropriate. Second, the effect of input from other edges via horizontal connections can be seen as increasing I and/or I_c. Therefore, in general, the net contextual influence on the edge can be facilitatory or suppressive, depending on relative recruitment of horizontal fibers, as experimentally observed (Hirsch and Gilbert 1991). Furthermore, since the gain $g'_y(y)$ increases with input level, such contextual influence is more likely to be inhibitory at higher input levels. This is also experimentally observed (Hirsch and Gilbert 1991, Sengpiel, Baddeley, Freeman, Harrad, and Blakemore 1995, Weliky et al 1995). In our model, for simplicity, the visual input is directed solely to the excitatory cells, but this could easily be generalized. The horizontal connections in the model are specified (in section 3.3) such that the net contextual influence from appropriately aligned edges is facilitatory at all contrast levels, although the contextual influence from less aligned edges (these edges can still prefer similar orientations) can depend on stimulus levels. Hence, in this model the change in the dominance between facilitatory and suppressive contextual influences occurs only for input patterns that are not contours, and therefore is not going to be discussed further in this paper.

3.3 Interactions between edge segments for contour integration

[2]When the segment has too strong a self-excitation — large $J_o g'_x(\bar{x})$ — and not enough inhibition — small $g'_y(\bar{y})$, such that $g'_y(\bar{y})g'_x(\bar{x}) < J_o g'_x(\bar{x}) - 1$ — the system is unstable and the output $g_x(\bar{x})$ jumps discontinuously with input I. Such cases are not considered here as we restrict to outputs as continuous functions of inputs for most input values.

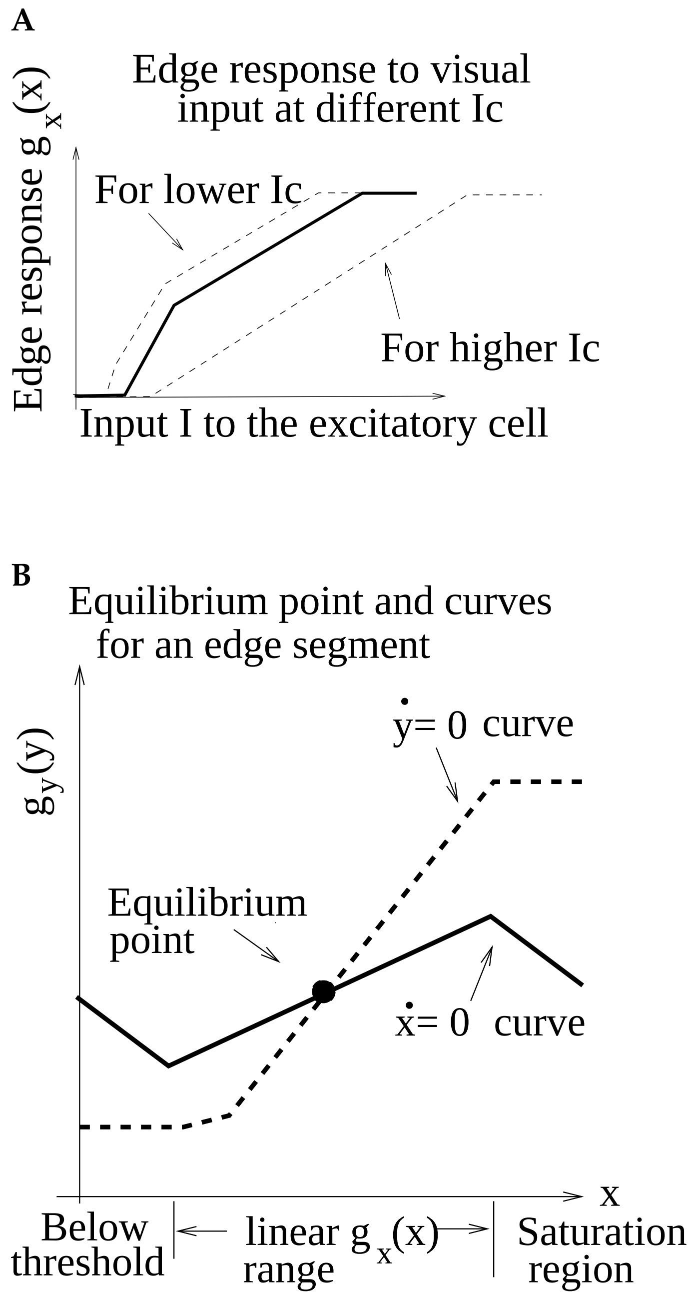

Figure 3: **A:** Edge response $g_x(\bar{x})$ as a function of visual input. Three response curves, one solid and two dashed ones, are plotted for three different cortical inputs I_c to the inhibitory cell. **B:** Equilibrium point and curves for an edge element. The solid curve is the $\dot{x} = 0$ curve, and the dashed curve is the $\dot{y} = 0$ curve. Increasing I raises the $\dot{x} = 0$ curve, and increasing I_c raises, and somewhat deforms, the (monotonously increasing) $\dot{y} = 0$ curve, thus changing the equilibrium point $(\bar{x}, \bar{y})$. The neurons may approach the equilibrium point after some transient, or oscillate around it, as discussed in section (3.6).

Edge element $(j\theta')$ excites or inhibits the edge element $(i\theta)$ by sending an excitatory-to-excitatory output $J_{i\theta,j\theta'}g_x(x_{j\theta'})$ or an excitatory-to-inhibitory output $W_{i\theta,j\theta'}g_x(x_{j\theta'})$. The goal for the connection structure $J_{i\theta,j\theta'}$ and $W_{i\theta,j\theta'}$ is that edge elements within a smooth contour should enhance each other's activities, and that the isolated edge elements caused by noisy inputs should be suppressed, or at least not enhanced. Hence:

- The connection $J_{i\theta,j\theta'}$ will be large if one can find a a smooth or small curvature contour to connect $(i\theta)$ and $(j\theta')$, and generally decreases with increasing curvature of the contour.

- Edge elements will inhibit each other via $W_{i\theta,j\theta'}$ when they are alternative choices in the route of a smooth contour.

- Both connection types will decrease with increasing distances between the edge segments, and become zero for large distances.

- The connections have translation, rotation, and reflection invariance. This means the following. Let $i - j$ be the line connecting the centers of two edges $(i\theta)$ and $(j\theta')$, which form angles θ_1 and θ_2 with this connecting line. The connections $J_{i\theta,j\theta'}$ and $W_{i\theta,j\theta'}$ depend only on $|i-j|$, θ_1, and θ_2, and satisfy $J_{i\theta,j\theta'} = J_{j\theta',i\theta}$ and $W_{i\theta,j\theta'} = W_{j\theta',i\theta}$.

Given these requirements, connections $J_{i\theta,j\theta'}$ and $W_{i\theta,j\theta'}$ both link cells that prefer similar orientations, as observed in experiments (Gilbert 1992, Weliky et al 1995) (Fig. 4). In addition, when the preferred orientations of two linked cells are aligned with the relative displacement of their RF centers, the post-synaptic cell type is more likely excitatory (the connection $J_{i\theta,j\theta'}$); while when the preferred orientations are more or less orthogonal to the relative displacement of the RF centers, the post-synaptic cell type is more likely inhibitory (the connection $W_{i\theta,j\theta'}$). This provides a computational explanation to the puzzling experimental finding (Gilbert and Wiesel 1983) that some horizontal connections link cells whose preferred orientations and relative RF center displacement do not align, but instead are roughly orthogonal to each other. These connections can serve to establish competition between alternative routes of a single contour by contacting inhibitory post-synaptic cells. This prediction (see Appendix for its derivation) of the model about the correlation between post-synaptic cell types and the degree of alignment between the two linked RFs has not been systematically investigated in experiments, a test is thus desirable. This excitatory and inhibitory edge interaction pattern is qualitatively similar with the edge compatibility function in Zucker et al (1989). Altogether, an edge in a smooth contour mostly receives facilitatory inputs $J_{i\theta,j\theta'}g_x(x_{j\theta'})$, and few, if any, inhibitory inputs $W_{i\theta,j\theta'}g_x(x_{j\theta'})$, from other edges in the contour. This helps to enhance the response of edges in a contour at any contrast level. Fig. 5 demonstrates the contour enhancement and noise suppression using such connections. The higher salience of the contours emerges rather quickly after the onset of the stimulus. For the example in Fig. 5A where

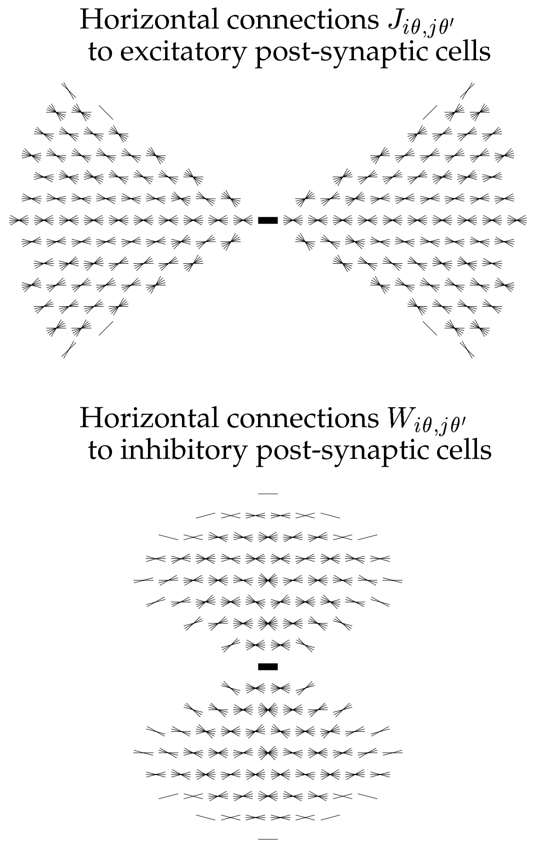

Figure 4: Model connections in an hexagonal grid visual space. Top and bottom graphs depict respectively the connections $J_{i\theta,j\theta'}$ and $W_{i\theta,j\theta'}$ from the center (thick) horizontal edge to other edges, denoted by their edge locations and orientations in the visual field. The quantitative values of the connection strengths $J_{i\theta,j\theta'}$ and $W_{i\theta,j\theta'}$ are described in the appendix. Each edge has the same connection pattern, suitably translated and rotated.

the visual input is weak, it takes a duration of 3 membrane time constants just for individual edges to integrate enough visual input for non-zero responses, but only another single time constant duration for the contextual input to drive the saliencies of the edges in the contour significantly higher than the background noise. The strengths and weakness of the model are further demonstrated in Fig. 5B, where the model is challenged by the difficulties in a natural image (see the figure caption and the discussion for details).

3.4 A straight line

We can understand the performance of the model by analyzing some examples. In the first example, the visual input is a horizontal line on the x axis

$$\hat{I}_{i\theta} = \begin{cases} I_{line} & \text{if } i \text{ is on the line and } \theta = 0 \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

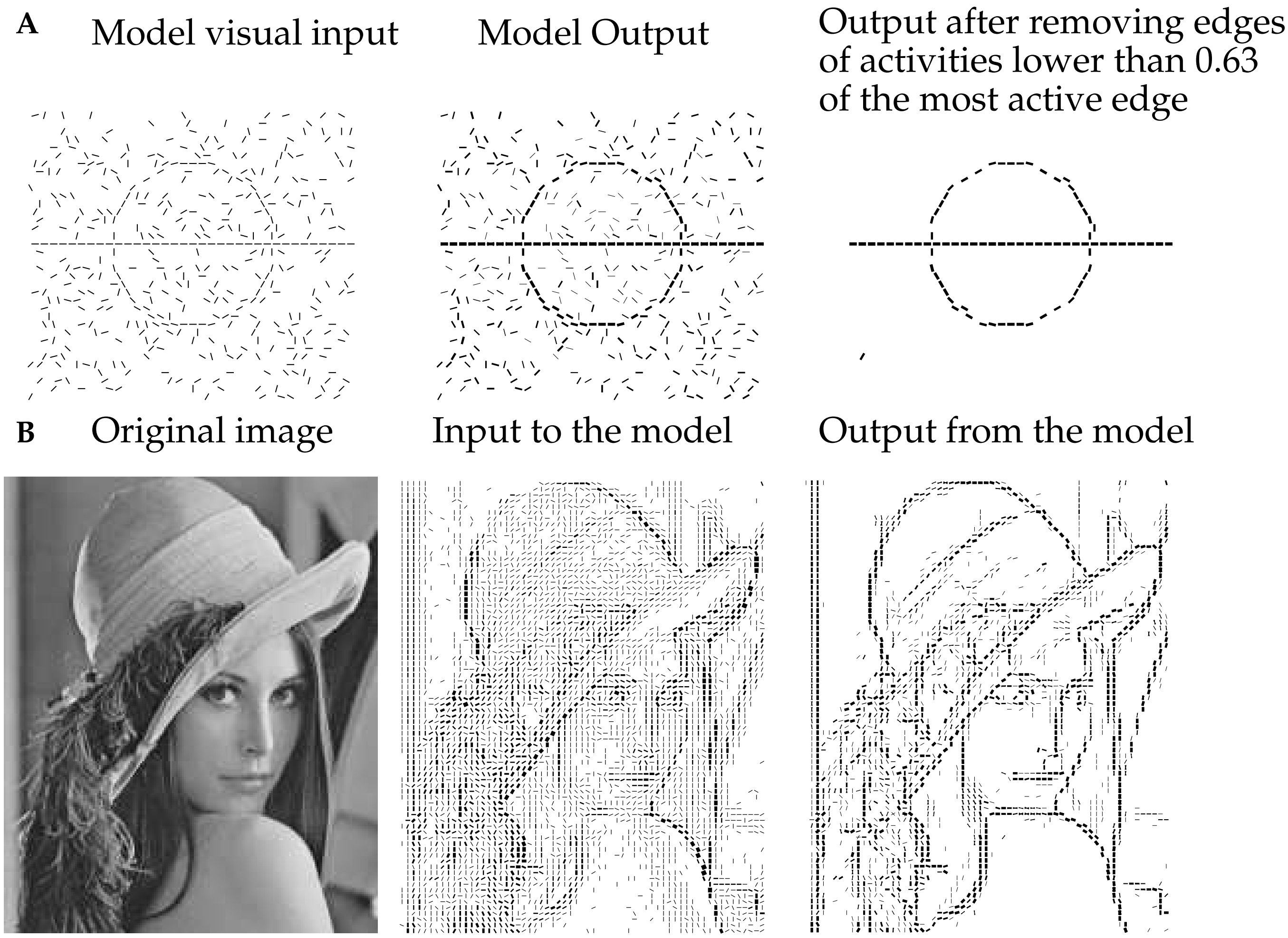
A
Model visual input
Model Output
Output after removing edges of activities lower than 0.63 of the most active edge
B
Original image
Input to the model
Output from the model

Figure 5: Contour enhancement and noise reduction. **A:** Performance for a synthetic image. **B:** Performance for an input obtained by edge detection from a natural photo. The input and output edge strengths are denoted proportionately by the thicknesses of the edges. The same format applies to other figures in this paper. The model outputs are the temporal averages of $g_x(x)$ over a period of 24 membrane time constants after input onset. In **A**, all visible edges have the same strength in the input, and are differentially enhanced or suppressed at the output. On average, the line and circle segments are roughly 2.5 times as salient as the "noise" segments. For demonstration, we display the outputs after thresholding out the weaker edges (right). Because of the discrete sampling grid, the apparent gaps in the contours actually sit on no grid point. Hence, no gaps actually exists for the contours and no filling-in is needed. Similar situations occur in other figures in this paper. In **B**, different edges have different input strength. The weak contour at the chin, at the brim of the hat above the eyes, and at the top of the hat (and many noisy edges), are in fact sub-threshold in input strength, however, they are more or less perserved at the model output. Since this model has only a single scale, finer edges (e.g., details around the eyes) in the photos can not be sampled or enhanced, a long gap in a contour (e.g., the long weak contour at the brim of the hat above the eyes at low input contrast) is difficult to fill. Aliasing problems (e.g., the left cheek and hair lines) and artificial contour gaps (e.g., in the brim of the hat above the eyes because no sampling point exist at the contour location) just like those in the circle in **A** happen because of a discrete sampling grid and a lack of a signal interpolation algorithm. In addition, two accidentally aligned but different contours (e.g., the contours for the hat, the hair line, and the cheek line near the right side of the image) can threaten to join each other because of the lack of T, L,and X junction signals that should prevent them. This model for contour enhancement also high-lights region boundaries and pops out novelties (Li 1997). Consequently, the edge segments for the hair pieces near the surrounds of the feather (even though the feather region is not very homogeneous), e.g., the single feather across the brim, are more likely enhanced than those inside the feather. See discussions in section 4.

The inputs I_c to the inhibitory cells of all edges are assumed to be the same. Let us consider the simplest case when edge elements outside the line (i.e., when i is not on the x axis or $\theta \neq 0$) are silent due to insufficient excitation. Then we can ignore all edges beyond the line, treat this system as one dimensional, and omit index θ. Let i denote the (one-dimensional) location of the line segment, then $W_{ij} = 0$, and $J_{ij} \geq 0$ for all $i, j \neq i$, and

$$\dot{x}_i = -x_i - g_y(y_i) + J_o g_x(x_i) + \sum_{j \neq i} J_{ij} g_x(x_j) + I_{line} + I_o \tag{10}$$

$$\dot{y}_i = -y_i + g_x(x_i) + I_c \tag{11}$$

If the line is infinite, then, by symmetry, each neuron pair will have the same equilibrium point[3] $E = (\bar{x}, \bar{y})$ determined by:

$$\dot{\bar{x}} = 0 = -\bar{x} - g_y(\bar{y}) + J_o g_x(\bar{x}) + (I_{line} + (\sum_{j \neq i} J_{ij}) g_x(\bar{x})) + I_o \tag{12}$$

$$= -\bar{x} - g_y(\bar{y}) + (J_o + \sum_{j \neq i} J_{ij}) g_x(\bar{x}) + I_{line} + I_o \tag{13}$$

$$\dot{\bar{y}} = 0 = -\bar{y} + g_x(\bar{x}) + I_c \tag{14}$$

This can be seen either as a single edge with extra external input $\Delta I = (\sum_{j \neq i} J_{ij}) g_x(\bar{x})$, or a giant single "edge" with a stronger self-excitatory connection $(J_o + \sum_{j \neq i} J_{ij})$ (Fig. 6). Either way, activity $g_x(\bar{x})$ for each edge element in the line is enhanced (Figs. 5 and 7).

The same minimum strength of input is required to excite a segment in a line or an isolated edge (Fig. 6A), if all segments in the line are equally excited. This is because the line segments will have zero output before they reach threshold and so can not excite each other. Therefore, they behave independently as isolated edges before threshold. However, when some line segments receive sub-threshold and others have super-threshold visual inputs, the former can give non-zero output under contextual excitation from the latter. This leads to sub-threshold activation or filling-in for the weaker or missing segments in a line (or contour, Fig. 9C).

3.5 Curvature, contour length, and contour closure

Enhancement in contours other than lines can be understood based on the special case of a contour of a constant non-zero curvature, namely a circle. It is apparent that the analysis for a straight line is also applicable here, assuming for simplicity that the corresponding circle for the contour has a diameter larger than the longest synaptic connection between cells. Index i again denotes the (one-dimensional) location of the segments along the (one-dimensional) contour, and $|i-j|$ the (one-dimensional) distance between the segments i and j. The activities of the elements along the contour are analogously enhanced.

[3]This equilibrium solution may or may not be stable, as studied in section (3.6). However, this does not matter, since this solution, stable or not, roughly gives the cells' average responses.

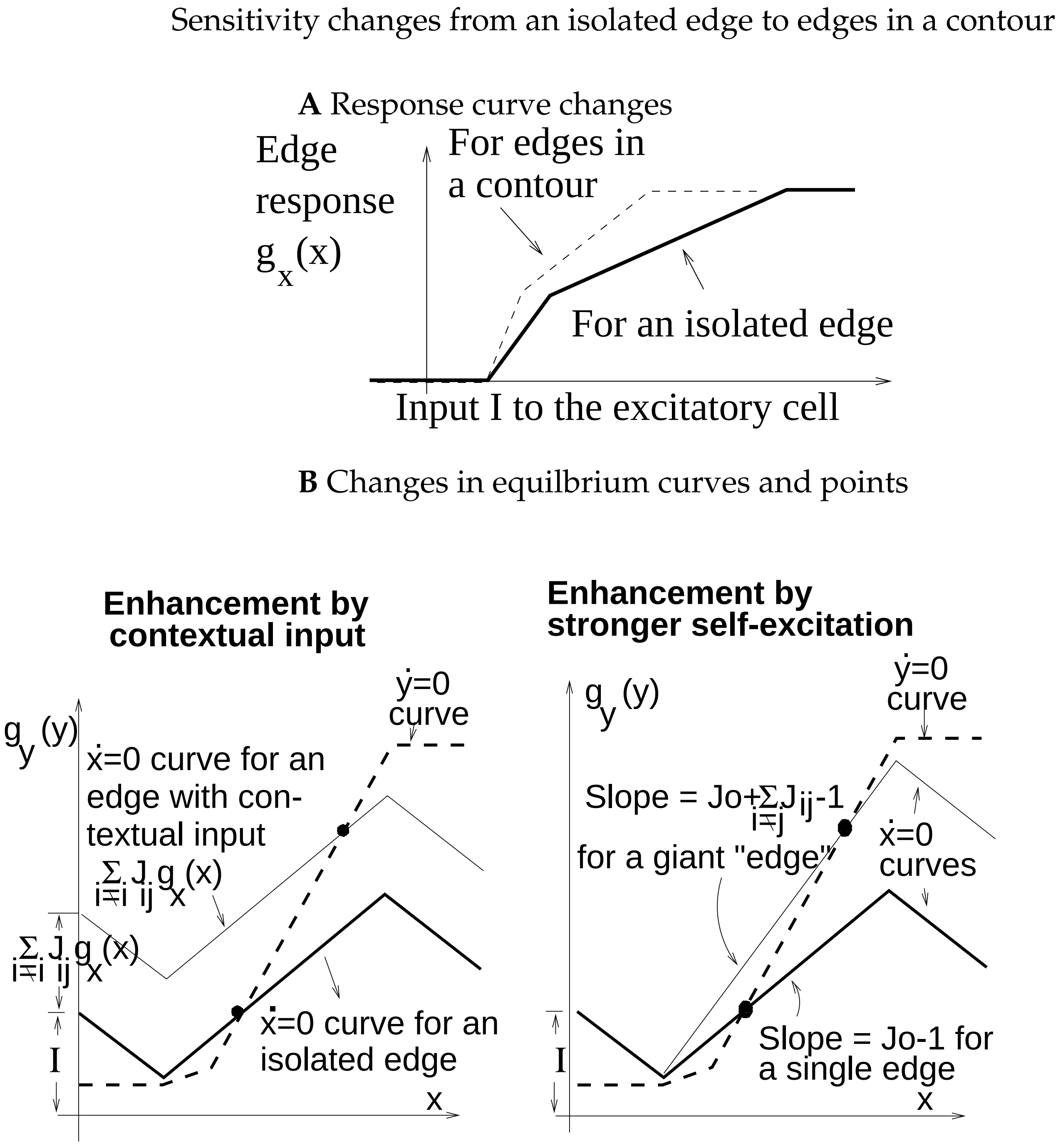

Figure 6: Response changes from an isolated edge to edges in contours. All edges in a contour are assumed to receive the same visual input strength. **A:** Changes in response curves. The solid curve is the response from an isolated edge, and dashed curve response from edges in contours. **B:** Changes in equilibrium curves and points. The two thick curves are the $\dot{x} = 0$ (solid) and $\dot{y} = 0$ (dashed) curves for an isolated edge. Only the $\dot{x} = 0$ curve is changed going from an isolated edge (thick solid curves) to edges in a contour (thin solid curve). Such changes can either be seen as caused by extra excitation $\sum_{j\neq i} J_{ij} g_x(x)$ from neighboring edge segments to a single edge (left figure), or as caused by an increase in the self-excitation or curve slope $J_o - 1 \rightarrow J_o + \sum_{j\neq i} J_{ij} - 1$ in a giant "edge". The equilibrium point changes from the lower black dot to the upper one.

It can be shown from Fig. 6B, after some geometrical calculation, that the response levels for a segment in a contour and an isolated segment differ by a factor $(s_y - s_1)/(s_y - s_2)$ where $s_1 = J_o - 1$, $s_2 = J_o + \sum_{i \neq j} J_{ij} - 1$, and s_y is the slope of the $\dot{y} = 0$ curve. The quantity $\sum_{i \neq j} J_{ij}$ is the sum of horizontal connection strengths between one contour segment to all others, hence it is larger for smoother contours by design, thus the response enhancement is also larger for smooth contours[4]. Furthermore, since s_y is usually smaller for smaller input strength, the contour enhancement is stronger for low input contrasts which are the case in some physiological and psychophysical experiments (e.g. Kapadia et al 1995, Kovacs and Julesz 1993). Roughly, this model enhances the saliencies in a smooth contour against background by up to 200-300% near threshold and around 100% at higher input strengths. In a contour of finite length, translation invariance (symmetry between segments) along the contour breaks down near its two ends. There the enhancement is less since the end segments get comparatively less excitation from other segments. In particular, this predicts that a closed or longer contour will have higher saliency than an open or shorter one, as seen in Fig. 7. These predictions are expected to hold also for contours of non-constant curvature. Note that the line in Fig. 7 is actually a closed line by the periodic boundary condition used in the model's visual space. Since the line has zero curvature, its salience is higher than that of the circle. Note that edge segments in an open and a closed contour (Fig. (7b) and Fig. (7c)) have roughly the same saliency except near the ends of the open contour where saliency decreases. This closure effect is weak compared to what is implied from psychophysical observation (Kovacs and Julesz 1993, see Discussion). Exactly how the saliency decays towards the ends depends on how $J_{i\theta,j\theta'}$ decays with inter-segment distances, and the longest connection length. Note that the model used a discrete hexagonal grid in the visual space, the apparent small gaps in the circle and curves are the artifact of the coarseness of the grid.

3.6 Neural oscillations and synchrony between contour segments

So far, we have only analyzed the equilibrium points $(\bar{x}, \bar{y})$, which are also roughly the average responses. Upon visual stimulation, the neurons may either approach these equilibrium after a transient phase, or they may sustain dynamic activities about these equilibrium, such as oscillations. Here we show how variations of neural activities about their averages reflect characteristics of the contour.

Each edge segment, a pair of connected excitatory and inhibitory cells, can be modelled as a neural oscillator (Li and Hopfield 1989) oscillating around the equilibrium point $(\bar{x}, \bar{y})$. With interactions between the segments, the oscillators are coupled and exhibit collective behavior reflecting contour characteristics embedded in the coupling. To analyze

[4] It can be shown that this conclusion still holds if some contour segments exert relatively weak, non-zero suppressive components (by connection W) on other segments, especially in curves of higher curvatures. In such cases, a modification can be derived in the expression above for the amount of overall contour enhancement.

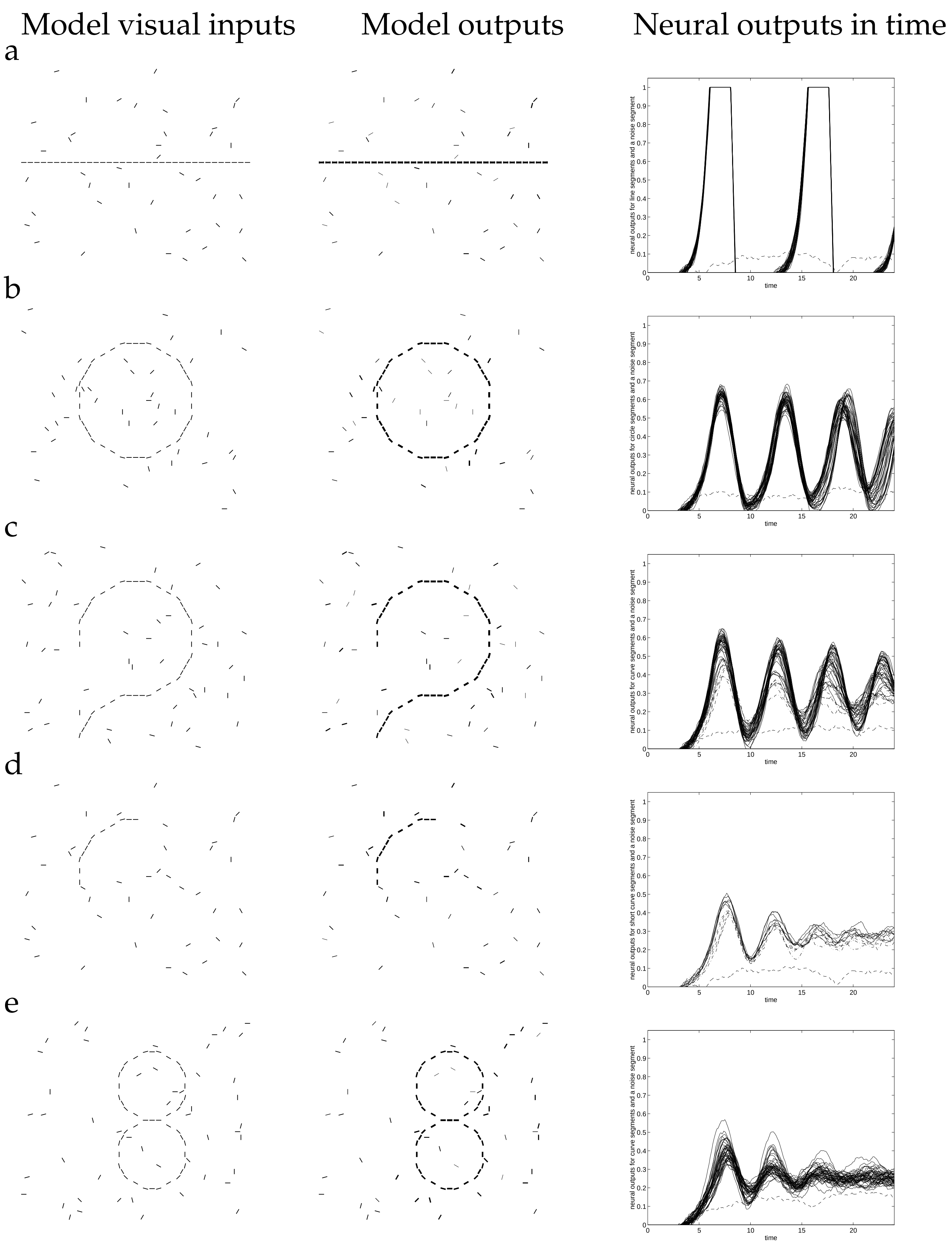
Model visual inputs
Model outputs
Neural outputs in time
a
b
c
d
e
neural outputs for line segments and a noise segment
neural outputs for circle segments and a noise segment
neural outputs for curve segments and a noise segment
neural outputs for short curve segments and a noise segment
neural outputs for curve segments and a noise segment
time

Figure 7: Model performance for input contours amidst noise. Each row is dedicated to one input condition. The left column shows the visual input; the middle column shows the average neural responses, and the right column shows the outputs of the segments as a function of time. The visual inputs are presented at time zero and stay on. The neural signals in right column are shown superposed. In the graph, the solid curves plot outputs of the segments along the contour, away from its ends; the dash-dotted curves the outputs of segments near the contour's ends; and the dashed curves, usually the lowest lying ones, are the output for a single noise segment. Note the decrease in average neural activities, in the amplitudes of the oscillations, and in the synchrony between the activities of the segments, as the contours become open, shorter, or more curled, or for segments near to the contour ends. Note that the line in the top row is actually a closed line by the periodic boundary conditions, and so by having zero curvature, its salience is higher than that of the circle. Also, the discrete hexagonal grid visual space gives apparent small gaps in the circle and curves because these "gaps" fall on no grid points — no filling-in is needed. Time is in units of the cell membrane time constant, and so a time constant of 5-10 msec would lead to oscillation frequencies in the figures around 10-40 Hz.

the dynamics, denote for simplicity $x_i - \bar{x}_i \rightarrow x_i$, $y_i - \bar{y}_i \rightarrow y_i$, and write as vectors $\mathbf{X} = (x_1, x_2, ...,)^T$ and $\mathbf{Y} = (y_1, y_2, ...)^T$. For small $\mathbf{X}$ and $\mathbf{Y}$, we approximate by a linear expansion of equations (10) and (11) about the equilibrium point $(\bar{x}_i, \bar{y}_i)$:

$$\dot{\mathbf{X}} = -\mathbf{X} - \mathbf{G}'_y\mathbf{Y} + \mathbf{J}\mathbf{X} \quad (15)$$

$$\dot{\mathbf{Y}} = -\mathbf{Y} + \mathbf{G}'_x\mathbf{X} \quad (16)$$

where[5] $\mathbf{J}$ is a matrix with elements $(\mathbf{J})_{ij} = J_o g'_x(\bar{x}_j)$, if $i = j$, and $(\mathbf{J})_{ij} = J_{ij} g'_x(\bar{x}_j)$ otherwise; $\mathbf{G}'_x$ and $\mathbf{G}'_y$ are diagonal matrices with elements $(\mathbf{G}'_x)_{ii} = g'_x(\bar{x}_i)$ and $(\mathbf{G}'_y)_{ii} = g'_y(\bar{y}_i)$. For a contour with a constant curvature (i.e., a circle), all its segments have the same equilibrium point if they receive the same input strength, $\bar{x}_i = \bar{x}_j \equiv \bar{x}$ and $\bar{y}_i = \bar{y}_j \equiv \bar{y}$. Then $\mathbf{G}'_x$ and $\mathbf{G}'_y$ are proportional to the identity matrix, and $\mathbf{J}$ is symmetric (since we imposed symmetry along contour directions, i.e., $J_{ij} = J_{ji}$). Then $\mathbf{J}$ has an orthogonal set of eigenvectors $\{\mathbf{X}^k\}$ and real eigenvalues λ^k for $k = 1, 2, ...$, which we order such that $\lambda^1 \geq \lambda^2 \geq ....$ Take $\{\mathbf{X}^k\}$ as the new basis to represent $\mathbf{X}$ and $\mathbf{Y}$, we have $\mathbf{X} = \sum_k x^k\mathbf{X}^k$, and $\mathbf{Y} = \sum_k y^k\mathbf{X}^k$, and

$$\dot{x}^k = -x^k - g'_y(\bar{y})y^k + \lambda^k x^k \quad (17)$$

$$\dot{y}^k = -y^k + g'_x(\bar{x})x^k \quad (18)$$

[5]Here we take for simplicity that contour segments do not link to each other with connection $W_{i\theta,j\theta'}$. The analysis here needs a little modification, but the general conclusion still holds, when W connections are included.

which has solution

$$x^k(t) = x^k(0)e^{-(1-\lambda^k/2)t}\cos(\omega^k t + \phi^k), \tag{19}$$

where the oscillation frequency is $\omega^k = \sqrt{g'_y(\bar{y})g'_x(\bar{x}) - (\lambda^k)^2/4}$, and initial conditions determine the amplitude $x^k(0)$ and the phase ϕ^k. The exponential in equation (19) suggests that the system will be dominated by the first oscillation mode[6] $\mathbf{X}^1$ since $\lambda^1 \geq \lambda^k$ for all k. The relative oscillation amplitudes and phases of the segments in a contour are determined by the components of the complex vector $\mathbf{X}^1$.

Let us suppose for simplicity, that the edge segments concerned are in the linear operating region where $g'_x(\bar{x}) = 1$. For a contour of constant curvature with uniform inputs to its segments, matrix $\mathbf{J}$ is Toplitz (i.e., $\mathbf{J}_{ij} = \mathbf{J}_{(i+a)\ \mathrm{mod}_N,(j+a)\ \mathrm{mod}_N}$ for all a, where N is the contour length or matrix dimension) under translation and rotation invariance of the model, and has non-negative elements. It can then be shown that the eigenvectors are the cosine and sine waves on the contour segments, i.e., $\mathbf{X}^k_j \propto e^{if_k j}$ with spatial frequency f_k and the eigenvalues are the corresponding Fourier coefficients of the row vector of the $\mathbf{J}$ matrix. In particular, the eigenvector $\mathbf{X}^1$ is the zero frequency Fourier wave, hence all components of $\mathbf{X}^1$ are equal, $x^1_i = x^1_j$, thus all segments in a contour oscillate with the same amplitude and phase. The eigenvalue for this mode is the zero frequency Fourier coefficient and thus $\lambda^1 = J_o + \sum_{\text{all } j \neq i \text{ on contour}} J_{ij}$. We can therefore relate λ^1 to the characteristics of the contour, as reflected by the connections J_{ij} along its length. It follows that the strength of the oscillation is largest in a long line, decays with increasing contour curvature (or decreasing contour length for circles), and is weakest for an isolated edge. The isolated edge is a special case with a single oscillator ($\mathbf{X}^1$ is a scalar) when $\lambda^1 = J_o$. When $\lambda^1 < 2$, the oscillation is damped and disappears after some transient phase. Otherwise, if segment couplings are sufficiently strong, the oscillations[7] will grow until the non-linearity invalidates the linear analysis and constrains the oscillation to have a finite amplitude (see the activities for the line segments in Fig. 7). These predictions are expected to hold approximately for general contours with non-constant curvatures or non-uniform inputs. However, because the translation invariance is compromised in these cases (i.e., $\mathbf{J}$ deviates from being Toplitz and symmetric), some differences in oscillation amplitudes and the relative phases are expected. Similarly, non-zero relative phases may emerge between edge segments near the end of a contour, or between such segments and

[6]From the analysis in the next paragraph, it will be apparent that it is unlikely to have degeneracy in the first two modes (the case when $\lambda^1 = \lambda^2$).

[7]For large enough λ^1, the oscillation frequency from the linear analysis $\omega^1 = \sqrt{g'_y(\bar{y})g'_x(\bar{x}) - (\lambda^1)^2/4}$ is imaginary. So the local dynamics about the equilibrium point is in fact not oscillatory. However, the global non-linear dynamics can be shown to be still oscillatory provided that system is not in the region of unstable self-amplification, i.e., $g'_y(\bar{y})g'_x(\bar{x}) > \lambda^1 - 1$.

those near the middle of the contour.

The role of λ^1 in the oscillation frequency $\omega^1 = \sqrt{g_y'(\bar{y})g_x'(\bar{x}) - (\lambda^1)^2/4}$ suggests a correlation between stronger contours and lower oscillation frequencies, as is the case when comparing the circle with the line in Fig. 7. However, in strong sustained oscillations, the small amplitude linear approximation no longer holds. The non-linearity greatly influences the frequency and makes this prediction imprecise. With strong visual input and contour enhancement, oscillation can be completely suppressed by the non-linearity near the saturation region. In realistic neural systems, however, saturation can be prevented by the gain control or adaptation for large inputs or activity levels.

Synchronization within a contour will happen even when the visual inputs are turned on at different times for different contour segments. On the other hand, synchrony is rare between different contours even when their visual inputs are turned on simultaneously. These are both demonstrated in Fig. 8 — segments within a line quickly reach synchrony after asynchronous stimulus onset, but circle segments and line segments become out of synchrony within 2 oscillation cycles after synchronous stimulus onset. In fact, when two contours are nearby, they tend to desynchronize largely due to the normalizing neural interactions (see Appendix) and the mutual suppressive couplings $W_{i\theta,j\theta'}$ between them. The dynamic coupling between two contours also causes frequency shifts. Furthermore, the nature and synchrony of the oscillations of the weaker contour tend to be distorted by the stronger contour such that some of the contour segments oscillate in a non-sinusoidal manner (Fig. 8B). We do not discuss the details of such dynamic coupling further as they are not used for contour integration in this model. However, synchronization within a contour and desynchronization between contours can be exploited for the purpose of contour segmentation (see discussion).

3.7 Control of the contour saliency by top-down feedback — selective contour enhancement/suppression, filling in, and contour segmentation

This section shows that our model of V1 provides a mechanism by which higher visual areas can selectively influence the response to given contours. We just assume that the higher centers already know which segments belong to a contour and what feedback signals to send back. In the model, the influence from higher centers is additional to, but not necessary for, the contour enhancing capabilities by the V1 neural circuit.

Higher visual areas are modelled as sending I_c as a feedback signal to the inhibitory cells, to influence the edge outputs according to equation (8):

$$\delta g_x(\bar{x})/\delta I_c = -\frac{g_y'(\bar{y})g_x'(\bar{x})}{1 + g_y'(\bar{y})g_x'(\bar{x}) - J_o g_x'(\bar{x})}. \qquad (20)$$

Let $I_c = I_{c,background} + I_{c,control}$. The background input $I_{c,background}$ is the same for all edge segments, and can be used to modulate the overall level of visual alertness. Higher areas can selectively enhance or suppress a given contour by providing a negative or positive $I_{c,control}$ (i.e., decreasing or increasing I_c) for the selected contour segments $(i\theta)$. A strong enough $I_{c,control} > 0$ on a given contour can completely suppress the outputs from that

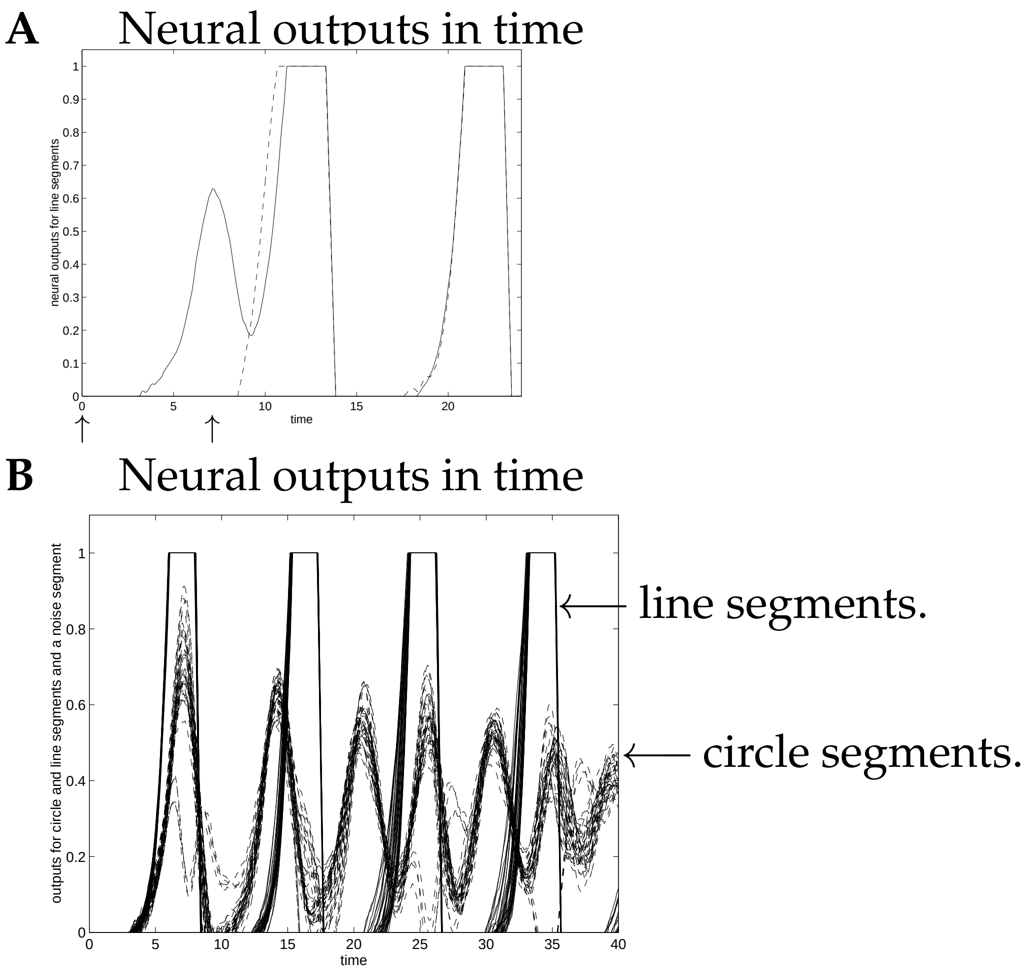
A
Neural outputs in time
neural outputs for line segments
time
B
Neural outputs in time
outputs for circle and line segments and a noise segment
time
line segments.
circle segments.

Figure 8: **A:** Synchrony within a contour even with asynchronous visual onset. Neural outputs for line segments as in the top row of Fig. 7, except that the visual input for the 17 leftmost line segments starts at t=0 (left arrow) whereas input for the 18 rightmost line segments starts at t=7 (right arrow). The solid curve and dashed curve correspond to one of the 17 leftmost and one of the 18 rightmost segments respectively, all segments in each group have roughly the same outputs. Note that synchrony is achieved within one oscillation cycle after the onset of the later segments. Compare with the top row of Fig. 7. **B:** Desynchronization between contours. Neural output based on visual input as in figure (9A), where a circle and line overlap and their inputs onset simultaneously at t=0. The edge outputs are plotted superposed, with solid curves for the line segments, and dash curves for the circle segments. Note that the line segments oscillate with smaller frequencies, and are more synchronized (the solid curves for different segments are almost identical). The circle segments differ somewhat in their oscillation amplitudes and phases, especially after two oscillation periods, much more so than they would if the line were absent (cf. Fig. 7). This is because the strong activities in the line segments interfere with the activities in the circle segments close to the line via local neural interactions. The coupling between the two contours also distorts the oscillation frequencies.

contour, leading to the effective removal or segmentation of this contour from other contours in the visual input.

It is not yet clear from experimental data which are the cells in V1 that should be the target of feedback (Salin and Bullier 1995, and see discussion). This model chooses the inhibitory interneurons as the targets for the reason that it is computational desirable not to mix visual inputs from the external world with the internal feedback signals. Given that the excitatory cells are the visual input neurons in this model, directing the feedbacks $I_{c,control}$ to the excitatory cells would lead to a confusion between the internal and the external. Since it targets the interneurons in this model, the feedback cannot completely substitute for visual stimulation. As is evident from above, I_c is effective only when $g'_y(\bar{y})g'_x(\bar{x}) \neq 0$. Without visual input or excitation from other segments, an edge segment's membrane potential $\bar{x}$ is below threshold and $g'_x(\bar{x}) = 0$. Decreasing or removing the feedback I_c merely reduces or removes the inhibition onto this segment, and is not enough to activate the excitatory cell beyond threshold when the background (non-visual) input I_o to the excitatory cell is too weak. However, with some, even sub-threshold, visual input $I_{i\theta}$ or contextual excitation from neighboring edges, an edge segment can increase its activity or become active if I_c is reduced or removed. Therefore, this model can enhance and complete or fill-in a weak and incomplete input contour under the control of the feedback (Fig. 9C), but cannot enhance or "hallucinate" any contour that does not at least partially exist in the visual input I (Fig. 9B). See section 4 for more detailed discussion on the targets of the feedback, the related experimental findings, and computational considerations.

Fig. 9 demonstrates central control. Without central control, $I_{c,control} = 0$, a visual input consisting of two contours, a circle and a line, and some noise segments results in a salient line and circle, and less salient noise segments (Fig. 9A). By adding $I_{c,control} > 0$ for the line and $I_{c,control} < 0$ for the circle, the saliency of the line is suppressed and the circle becomes most salient. If feedback control $I_{c,control} > 0$ for the line is strong enough, then its neural activity can be completely eliminated, effectively segmenting it away from the circle (Fig. 9A). Without central control, the gaps in an input line are partially filled in by the excitation from other line segments (Fig. 9C); With reduced I_c on the line, the initially fragmented line becomes almost completely filled-in (Fig. 9C).

4. Summary and discussion

4.1 Summary of the motivation, components, and performance of the model

Although experimental data suggest that contour enhancement may be first attempted in V1 (Kapadia et al 1995), it has been difficult to model it by a realistic model of V1 largely because of the inherent dynamic stability problems in recurrent neural networks. The V1 network has to be sensitive enough to amplify all possible smooth contours it may receive, and at the same time should not have uncontrollable sensitivities that would make the contours and noises be amplified indiscriminately. Such difficulties have until now prevented a realistic V1 model of contour enhancement. Therefore, it is desirable to find out whether contour enhancement can indeed be accomplished by V1 and thus be modelled by a realistic V1 circuit, or whether the observed contour sensitivity enhancements in experiments (Kapadia 1995) have to be all attributed to higher visual areas and their feedbacks to V1. It is one of the goals of this study to demonstrate and understand how contour enhancement can indeed be reasonably carried out in a model using only the known elements of V1. Accordingly, our model uses orientation selective cells, a local neural circuit with recurrent interactions between excitatory and inhibitory cells, and the particular connection patterns suggested by experimental evidence (Gilbert 1992, Hirsh and Gilbert 1991, Weliky et al 1995, White 1989, Douglas and Martin 1990, Kapadia et al 1995). The neural interactions in the model enhance the cell activities for edge segments belonging to smooth contours against a background of random edge segments, and induce synchronized oscillatory neural activities between segments within a contour. We show analytically and empirically that the extent of contour enhancement and neural synchrony are stronger for longer, smoother, and closed contours. These behaviors of the model are consistent with experimental observations (Kapadia et al 1995, Field et al 1993, Kovacs and Julesz 1993, Gray and Singer 1989, Eckhorn et al 1988). In addition, this model introduces a mechanism which allows higher visual areas to feedback and selectively enhance or suppress activities for given contours, and even to achieve a crude form of contour segmentation. This model makes the following testable predictions which have not been systematically investigated experimentally: (1) the horizontal cortical connec-

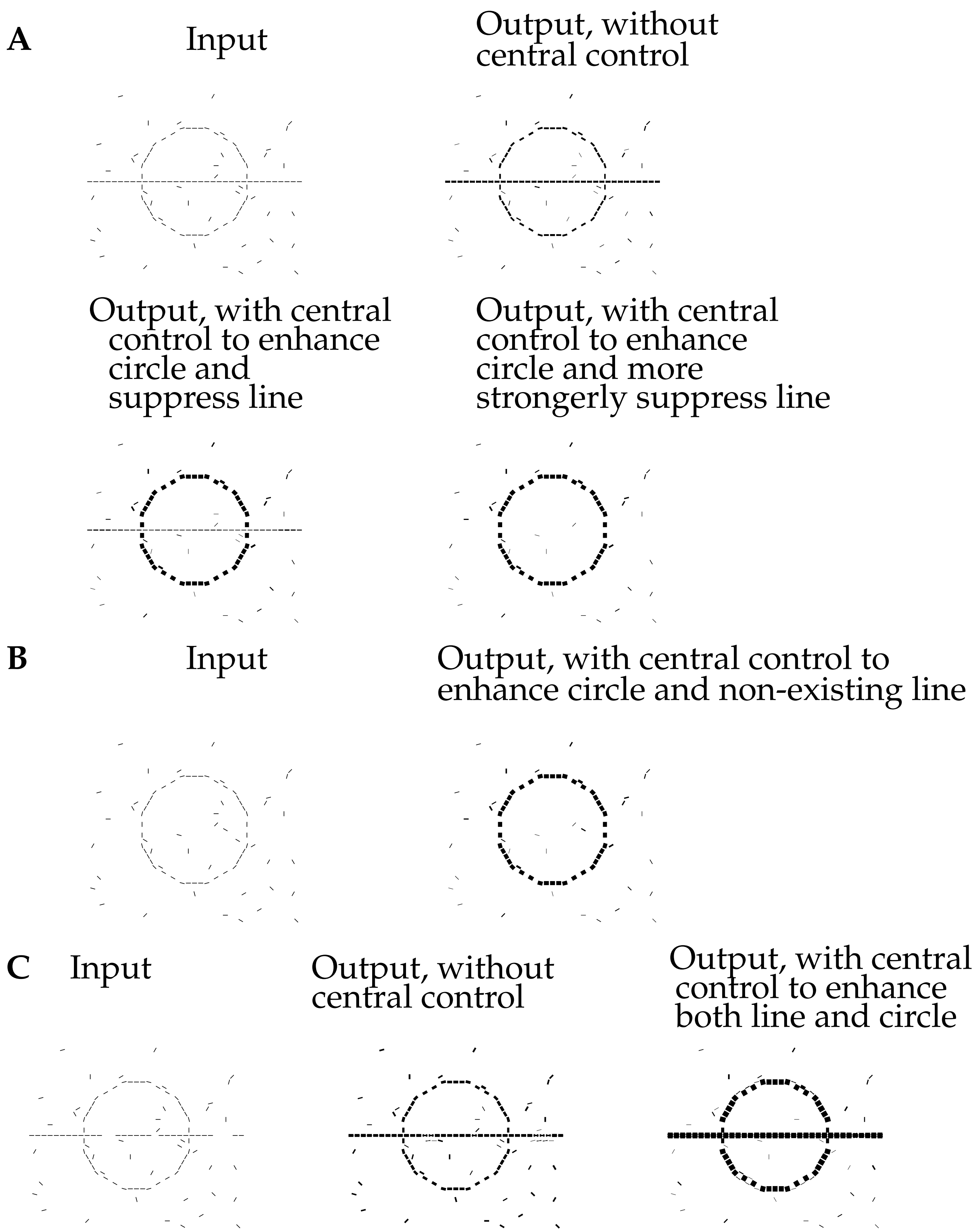
A
Input
Output, without central control
Output, with central control to enhance circle and suppress line
Output, with central control to enhance circle and more strongerly suppress line
B
Input
Output, with central control to enhance circle and non-existing line
C
Input
Output, without central control
Output, with central control to enhance both line and circle

Figure 9: Central feedback control. **A:** Without central control, both the line and the circle in the input (top left) are suitably enhanced against the background (top right). The bottom left graph demonstrates the selective enhancement of the circle by $I_{c,control} = -I_{c,background}/5$ and suppression of the line by $I_{c,control} = I_{c,background}/10$ on the respective edge segments. Stronger line suppression by $I_{c,control} = I_{c,background}/3$ can effectively remove all responses to the line (bottom right). **B:** Lack of hallucination: the non-existing line in the input can not be enhanced by central control $I_{c,control} = -I_{c,background}/5$ to enhance both the non-exiting line and the existing circle. **C:** Gaps in the line of the input (left) is partially filled in by the contour enhancement mechanism in V1 (center), and more completely filled in by the central control to enhance the contours with $I_{c,control} = -I_{c,background}/3$ (right).

tions from the excitatory cells should more likely contact excitatory or inhibitory post-synaptic cells if the two linked cells have their preferred orientations roughly parallel or orthogonal, respectively, to their relative RF displacement; (2), The strength of neural oscillation, as well as neural synchrony, should increase with contour length, smoothness, and closure.

For analytical tractability and simplicity, the model adopts the following idealizations: a 1:1 ratio between the excitatory and inhibitory cell numbers, the lack of connections between the inhibitory cells, and the lack of direct visual input to the interneurons. Without essential changes to the model performance, these idealizations can be relaxed to give additional complexities in model behavior. For instance, each model cell really should be seen as modeling a local group of cells of similar types. Hence the 1:1 ratio between the excitatory and inhibitory model cell numbers is really a ratio between local cell groups, and the recurrent local connections between them model the recurrent connections in the local cell groups. Also, introducing direct visual inputs to the inhibitory cells can give additional input gain control to allow larger dynamic range for the system.

The recurrent local interactions between excitatory and inhibitory cells used in this model have long been part of a "basic circuit" for the cerebral cortical organization (Shepherd 1990). They have been used, for instance, in a model of the olfactory bulb for odor recognition and segmentation (Li and Hopfield 1989, Li 1990). A closely related version of this circuit is also used in a model of visual cortical RFs and surround influences (Somers, Todorov, Siapas, and Sur 1995).

Our model requires a neural connection structure with both a colinear excitatory component $J_{i\theta,j\theta'}$ and flanking or orthogonal and disynaptically inhibitory component $W_{i\theta,j\theta'}$. A similar connection structure is proposed by Zucker et al (1989), where the inhibitory component is modelled monosynaptically. A connection structure like our colinear excitatory connection component $J_{i\theta,j\theta'}$ is termed "association field" by Field et al (1993), and is generic for many contour enhancement models (e.g., Zucker et al 1989, Braun et al 1994, Yen and Finkel 1996). Experimentally however, cortical cells have horizontal axonal fields that extend orthogonally as well as parallel (colinearly) to the preferred orientation of the cells (Gilbert and Wiesel 1983). The "association field" like connections can only account for those axons that extend in a roughly parallel direction. Our model suggests that the orthogonally extending axons (Fig. 4) should be found to contact preferentially inhibitory post-synaptic cells (the connection $W_{i\theta,j\theta'}$), for the computational purpose of mediating competition between alternative routes of a contour. The different functions served by the different branch directions of the horizontal axons lead naturally to the anisotropic horizontal axonal fields as observed in V1 (Gilbert and Wiesel 1983, Fitzpatrick 1996b). The quantitative degree of anisotropy in the axonal field is not crucial in this model. The predicted correlation between axon directions and post-synaptic cell types has yet to be experimentally tested.

4.2 Higher center feedback

The feedback control mechanism in the model may also relate to the attentional effects observed in V1 cells (Motters 1993). It has the desirable property that while higher areas can enhance input contours, and even fill-in the gaps in an incomplete contour, they can not create a contour that does not exist in visual input. This property could be exploited by the higher visual centers to test hypotheses about the visual input and cooperate with V1 to reconstruct a coherent percept [8]. A complete model of top-down bottom up cooperation should include a mechanism by which higher areas can respond to V1 outputs and construct the requisite top-down control signals. This mechanism is left for future work.

Experimental data have suggested both excitatory (Mignard and Malpeli 1991, Nault et al. 1990) and inhibitory (Alonso, Cudeiro, Rerez, Gonzalez, and Acuna 1993, Fitzpatrick 1996a) effects of feedback signals. Most data on feedback fibers show that feedback terminals synapse onto dendritic spines (Rockland 1996, Johnson and Burkhalter 1992) which are usually associated with excitatory cells, though earlier observations from Johnson and Burkhalter (1991) suggest that feedback fibers terminate near inhibitory interneurons. Although these evidences are not consistent or clear cut (Salin and Bullier 1995), some of them point in the opposite direction from our model construction that higher center feedbacks are directed to the inhibitory interneurons. Computationally, I believe that it is important not to mix bottom-up visual input signals with top-down feedback ones. Hence the feedback fibers should avoid the input neurons, which in this model are the excitatory cells. Indeed, in the brain, feedback fibers generally avoid cortical layer 4 which is the input layer (Salin and Bullier 1995). Experimental evidence also suggest that top-down feedback to V1 can modulate V1 activities, but cannot substitute for visual input to activate V1 (Salin and Bullier 1995). If hallucinations should be avoided in the visual system, then, as is shown in section 3, it helps to send the feedback via

[8] In this model, the "hallucination" is prevented by setting the background input I_o sufficiently low such that excitation by I_o alone without visual input is impossible even with feedback enhancement. Hallucination could occur if I_o is large enough.

the inhibitory interneurons, as is done in this model. More realistically, V1 has different layers and groups of excitatory cells. Different excitatory cell groups are likely to serve different functions, and "hallucinations" may be avoided when the feedback fibers target other excitatory cells and avoid the excitatory cells in the input layer. It is also likely that top-down feedback first contact the non-input excitatory cells, which then transform the signals to the inhibitory interneurons. Such signal transformations may be needed in order to render the visual representations in the higher areas in terms of the representation in V1. It is also likely that the feedback effects are dynamically modulated and can be excitatory or inhibitory depending on the levels of neural activity and contextual conditions that are beyond the current model. More consistent and informative experimental data are desired to guide our further understanding. It is more clearly established in the the lateral geniculate neuclus (Salin and Bullier 1995) and the olfactory bulb (Shepherd 1990) that higher area feedback terminates mostly onto the inhibitory interneurons. It will be interesting to explore the extent of the universality of computational mechanisms across stages and sensory modalities. An analogous model of higher center control and olfactory segmentation has been proposed for the olfactory bulb (Li 1990).

4.3 Limitations and extensions of the model

As a first attempt to model contour integration using V1 components, this model has many weaknesses. The degree of closure effect in particular is much weaker than what is implied by psychophysical observations (Kovacs and Julesz 1993). It is only (Fig. (7)) in the edge segments near the ends of the open contour that saliencies decrease by up to $20-30\%$. Saliencies of the segments far from the ends are comparable to those in a closed contour. It is not clear whether a stronger closure effect can be achieved by more sophisticated neurons and their interactions, or by introducing mechanisms beyond V1 and the scope of this model. It is conceivable that closure effect can be stronger by introducing the notion of object surfaces and adding surface saliency to a region bounded by a closed curve. Such means are likely to involve higher visual centers. It is desirable to find out experimentally whether the closure effect, so far investigated only psychophysically, is significant in V1 without the involvement of the higher center feedback. Another noticeable fact is that, if saliencies are proportional to the temporally averaged neural activities, the quantitative changes in saliency with contour length, curvature, and closure for reasonably smooth contours are not very significant. This could be a weakness or strength of this model, depending on whether one desires sensitivity or robustness of the model performance. However, if one uses the maximum neural activities over a time window as a measure of saliency, then saliencies change more significantly with contour characteristics. This model addresses only the local saliencies of individual edge segments. One may conceivably use synchrony between segments to obtain the global saliency of a whole contour. That should make the contour saliency more sensitive to the contour characteristics. However this is beyond the scope of this paper.

This model uses an idealized image sampling grid in a single scale, making it difficult to handle cases when contours and their locations are defined in multiple scales, as is the case in many natural images. This difficulty is noticeable and explained in Fig (5) for the photo input. Substantial work, at least in the scale of simulations, will be required to make the model multiscale. Fig (5)B also makes it clear that the problems of image sampling and signal interpolation/interpretation, though beyond the contour enhancement model and probably outside V1, should be solved to give better inputs to the model and to better interpret the outputs.

In addition to orientation and spatial location, RFs in V1 are tuned for motion direction, motion speed, disparity, ocularity, scale, and color (Hubel and Wiesel 1962, Livingstone and Hubel 1984). Object contours can thus exist in all these dimensions. The current model can be extended to stereo, time, and color dimensions. The extended model will link edge segments with compatible selectivities in these dimensions as well as aligned orientations. Indeed, experiments reveal that horizontal connections tend to link cells with similar RF properties in dimensions other than orientation preference (e.g., Gilbert 1992), and that activities of cells with similar ocular dominance, color selectivity, and other complex RF properties tend to be correlated (Ts'o et al 1986, Ts'o and Gilbert 1988, Singer and Gray 1995). Such an extension has not been carried out yet.

The interactions designed to instantiate contour integration can be explored to see whether they also account for other visual contextual phenomena such as figure-ground and texture segmentation, the tilt illusion, and non-contour perceptual filling-ins (e.g., Gilbert 1992, Field et al. 1993). Given its significant simplification, I expect that the model will not explain many visual phenomena beyond contour integration. For example, it does not capture the highly flexible gain control and receptive field adaptation observed in the primary visual cortex (Gilbert 1992). However, some model limitations can be interpreted in other ways. For example, experiments on contextual influence on the orientation selectivity of cat V1 cells (Gilbert and Wiesel 1990) have found some data consistent but other data inconsistent with the perceptual tilt illusion phenomena. This model also displays similar ambiguous results, depending on how the contextual stimuli are placed in the visual fields.

Recent experimental evidence suggests that V1 cells contribute to figure ground distinctions (Zipser Lamme, and Schiller 1996, Lamme 1995), a fundamental problem that has been addressed by many models (e.g., Kienker, Sejnowski, Hinton, and Schumacher 1986). Although our model is originally aimed at grouping edges to contours, my more recent studies find that the model can also signal figure-ground differences and contribute to region grouping and the phenomena of pop-out. Since boundaries and regions are complementary to each other, it is reasonable to expect or require a model of boundary enhancement to signal region differences. This potential of the model is yet to be fully explored, and some early results can be found in (Li, 1997).

4.4 Relating to previous models

Many other models are related to at least parts of our contour integration model. We relate our model to the most representative and relevant ones of them, and acknowledge that many more models exist and can be found, for instance, in the references of the papers cited here. One class of models addresses the underlying computation without using biologically plausible model elements, interactions, or algorithms. Shashua and Ullman (1988) modeled the perceptual saliencies of contour elements using a simple iterative network with local interactions between image elements. The model performs well, and Fig. 5A of this paper is patterned after an example of theirs. Guy and Medioni's model (1993) lets local image features vote on underlying global contours. The contours are then extracted by combining the votes with several methods of combination. Motivated by the existence in V2 of cells responsive to illusory contours (von der Heydt, Peterhans, Baumgartner 1984), Heitger and von der Heydt (1993) suggested a model that *infers* illusory occluding contours from T-junctions, corners, and line ends extracted from the image. This inference requires a highly non-linear operation to check the consistency of the occluding contour interpretations based on "end-stopped" signals. A recent model by Williams and Jacobs (1996) generates in the image plane a probability distribution over all the contours that can join two separated edge fragments, modeling the linking process as a random walk in position and orientation in the image plane. This model can be implemented in a network with local, albeit non-neural, interactions. However, the network has to know which edges to join before it can complete a contour between them.

Another class of models is closer to neurobiology. For example, the model by Grossberg and Mingolla (1985) aims for boundary or contour completion. The algorithm requires dipole fields in V1 which prevent two perpendicular contours crossing each other, and bipolar cells in V2 which provide cooperative feedback. The neural bases for the dipole fields is not clear, and contour enhancement necessarily requires feedback from V2. The model by Zucker et al (1989) used an relaxation labeling algorithm (Hummel and Zucker 1983) to infer contour segments from an initial image measurement resembling that in V1. The edge segments are labeled by their local orientations and curvatures. Two segments excite or inhibit each other depending on whether they are compatible with each other, in a way similar to our connection structure J and W. The orientation and curvature labels are updated iteratively by the compatibility interaction. The algorithm performs well after 2 iterations and should converge to some final configuration (Hummel and Zucker 1983). Braun, Niebur, Schuster and Koch (1994) suggested a model for contour integration that maps well to the human contour perception in the same visual displays. This model requires fast-adapting neural synapses and dendritic gating which have yet to be confirmed by experiments. The model by Yen and Finkel (1996) uses the "association field" like horizontal connections in their network and it works well. However, the model algorithm includes non-neural operations such as a global normalization of unit responses after each network iteration and a rule based algorithm for inter-unit synchronization.

Compared with these models, we stress the restriction to the V1 elements and operations. Only by doing so, can we ascertain whether contour enhancement can really be first attempted in V1 or has to be attributed to top-down feedback. Such consideration also accounts for our decision not to employ in this model image signals from T-junctions and corners, for there is no evidence for such units in V1. Consequently, this model *enhances* existing weak contours rather than *inferring* (invisible) foreground occluding (illusory) contours from visible image signals in the background. Furthermore, we study the model analytically in addition to the empirical simulation study, so as to grasp the relationship between neural interactions, contour characteristics, and the corresponding model behavior, and to overcome the formidable dynamic stability problem for reliable contour enhancement in a recurrent neural network model of V1. This model also avoids some of the undesirable features associated with some other models. For instance, this model does not give unreasonably high saliencies to short segments of contours that are attached smoothly to long and smooth contours. This is a problem for, e.g., Shashua and Ullman's model (1988) partly because an edge is defined in their model as directed, i.e., an edge of orientation θ is different from another edge of orientation $\theta + \pi$ at the same location. Such directed edges also appear in other non-biological models such as Williams and Jacobs (1996) chiefly to suit the particular contour enhancement algorithms. Our model is also unique in addressing the V1 response to top-down feedback, not as a requirement for contour integration, but as an additional feature that is simple but computationally powerful.

4.5 Neural oscillations

Oscillations do not seem to be essential to solve the computational problem of contour enhancement and could be just an epiphenomenon. They are intrinsic properties of recurrently interacting excitatory and inhibitory cell populations used in our model. On the other hand, since the strength of the oscillations and degree of synchrony do correlate with characteristics of the contours, the extra information carried in the oscillation could be exploited for other computations such as feature linking (von der Malsburg 1981). This model predicts weaker and more transient oscillations for shorter or weaker contours or for isolated edges. This may explain the failure to observe oscillatory neural behaviors in some experiments (Singer and Gray 1995). There are many other models of cortical neural oscillations and their possible roles in feature linking and segmentation (e.g., Baldi and Meir 1990, Sporns, Tononi, and Edelman 1991, König and Schillen 1991, von der Malsburg and Buhmann 1992 Sompolinsky, Golumb, and Kleinfeld 1991, Murata and Shimuzi 1993, Wang 1995 and references therein). Some model neural oscillators by oscillation phase variables only (e.g. Sompolinsky et al 1991, Baldi and Meir 1990), making the actual neural activity levels unavailable to study feature enhancement. Our oscillator, with its interconnected excitatory and inhibitory cells, is suggested by models of the olfactory bulb (Freeman 1987, Li and Hopfield 1989). Such neural oscillator models are also employed by many visual cortical models where both the phases and amplitudes of the oscillations can be studied (e.g. Sporns et al 91, König and Schillen 1991, von der Malsburg and Buhmann 1992, Murata and Shimuzi 1993, Wang 1995). However, these visual cortical models do not address the contour integration problem.

In summary, we have introduced a biological plausible model of contour integration in V1. The model exhibits experimentally observed behaviors and makes testable predictions about V1 anatomy and physiology. Extensions of the model to other visual input dimensions can be explored. This model also provides an analytical framework to study the neural dynamics and other visual computations in V1 (e.g., Li 1997). Some of the properties and mechanisms of the model can be exploited to study other computational problems such as hypothesis testing through top-down control, feature linking, and figure-ground segmentation.

Appendix

This appendix gives the detailed model parameters and a derivation of the model connection structure.

Model parameters

The number of orientations or edge elements at each spatial grid point is $K = 12$. Equations (1) and (2) determine the model dynamics.

The gain functions for the neurons are:

$$g_x(x) = \begin{cases} 0 & \text{if } x < T_x \\ (x - T_x) & \text{if } T_x \leq x \leq T_x + 1 \\ 1 & \text{if } x > T_x + 1 \end{cases} \tag{21}$$

$$g_y(y) = \begin{cases} 0 & \text{if } y < 0 \\ g_1 y & \text{if } 0 \leq y \leq L_y \\ g_1 L_y + g_2(y - L_y) & \text{if } 0 < L_y \leq y \end{cases} \tag{22}$$

where $T_x = 1$, $L_y = 1.2$, $g_1 = 0.21$, and $g_2 = 2.5$.

Except for the case of Fig (5B) where variable input strengths are used for different edges, the edge input strength for all visible edge segments in all other cases is the same $\hat{I}_{i\theta} = 1.02$, and $\hat{I}_{i\theta} = 0$ otherwise.

The weighting function $\psi(\theta)$ for the inhibitory cell to the local excitatory cells at the same grid points is:

$$\psi(\theta) = \begin{cases} 1 & \text{when } \theta = 0 \\ 0.8 & \text{when } |\theta| = \pi/K = 15^o \\ 0.7 & \text{when } |\theta| = 2\pi/K = 30^o \\ 0 & \text{otherwise} \end{cases} \tag{23}$$

The background inputs to the inhibitory cells is $I_{c,background} = 1.0$.

The central feedback control $I_{c,control}$ is applied to the inhibitory cells of the edge segments in an analogous way as the visual inputs are applied to the excitatory cells. If the higher centers intend a control to grid point i with orientation β, then $I_{c,control}$ on the edge segment $(i\theta)$ is

$$I_{c,control}(i\theta) = \hat{I}_{c,control}(i\beta)\psi(\theta - \beta) \tag{24}$$

where $\hat{I}_{c,control}(i\beta)$ serves a function analogous to $\hat{I}_{i\beta}$ in the visual input.

The background inputs I_o to the excitatory cells includes the following:

$$I_o = I_{e,background} + I_{normalization} \tag{25}$$

where $I_{e,background} = 0.85$, and $I_{normalization}$ is a normalization current that depends on the local edge activities, so its value for the edge element $(i\theta)$ is

$$I_{normalization}(i\theta) = -2.0 \left[\frac{\sum_{j \in S} \sum_{\theta'} g_x(x_{j\theta'})}{\sum_{j \in S_i} 1} \right]^2 \tag{26}$$

where S_i is a neighborhood of all grid point j that are no more than 2 grid distance away from i. This normalization $I_{normalization}$ is after the model by Heeger (1992) for cortical cells to account for non-orientation specific local cortical activity normalization and non-linearity. It can be implemented by other inhibitory interneurons of relatively short time constants which receive inputs from local excitatory cell pools and feed back to them. In addition, each neuron receives an input I_{noise} which is a random noise with an average temporal width of 0.1 and an average height of 0.1. Noise input to different neurons are independent.

The self-excitatory connection is $J_o = 0.8$. The long range synaptic connections $J_{i\theta,j\theta'}$ and $W_{i\theta,j\theta'}$ are determined as follows. Let the two edge elments $(i\theta)$ and $j\theta'$ be separated by a grid distance d, and denote the angles between the edge elements and the line connecting two edge elements by θ_1 and θ_2, where $|\theta_1| \leq |\theta_2| \leq \pi/2$, and $\theta_{1,2}$ are positive or negative depending on whether the edges rotates clockwise or counter clockwise towards the connecting line in no more than a $\pi/2$ angle. Denote $\beta = 2|\theta_1| + 2\sin(|\theta_1 + \theta_2|)$,

$\Delta\theta = \theta - \theta'$ with $|\theta - \theta'| \le \pi/2$, then [9]

$$J_{i\theta,j\theta'} = \begin{cases} 0.126 e^{-(\beta/d)^2 - 2(\beta/d)^7 - d^2/90} & \text{if } 0 < d <= 10.0 \text{ and } \beta < \pi/2.69 \\ & \text{or } 0 < d <= 10.0 \text{ and } \beta < \pi/1.1 \\ & \text{and } |\theta_1| < \pi/5.9 \text{ and } |\theta_2| < \pi/5.9 \\ 0 & \text{otherwise} \end{cases}$$

$$W_{i\theta,j\theta'} = \begin{cases} 0 & \text{if } d = 0 \text{ or } d/\cos(\beta/4) \ge 10 \text{ or } \beta < \frac{\pi}{1.1} \\ & \text{or } |\Delta\theta| \ge \frac{\pi}{3} \text{ or } |\theta_1| < \pi/11.999 \\ 0.14(1 - e^{-0.4(\beta/d)^{1.5}}) e^{-(\Delta\theta/(\pi/4))^{1.5}} & \text{otherwise} \end{cases}$$

Derivation of the connection structure

We derive the qualitative structure of the neural connections, using the following computational requirements. (1) Two nearby edge segments should enhance each other's activities if one could draw a smooth contour passing both of them; (2) A small enough gap in a smooth contour of sufficient input length and strength should be filled by the network; (3) A smooth contour of finite length and width in input should not grow in length or width by the network enhancement.

According to equations (7) and (8), an edge segment $(j\theta')$ can enhance or suppress edge segment $(i\theta)$ by sending monosynaptic excitatory input via connection $J_{i\theta,j\theta'}$ or di-synaptic inhibitory input via connection $W_{i\theta,j\theta'}$.

Hence, condition (1) above requires $J_{i\theta,j\theta'} \neq 0$ if $(i\theta)$ and $(j\theta')$ are nearby and roughly co-aligned, as illustrated in figure (4).

The bounds on the overall scale of the J connection can be obtained by considering a contour of a straight line on the x-axis like in equation (9). Omit the θ variables, let i^{th} segment be at $x = 0$, origin of the x-axis, and the j^{th} the other contour segments. If only segment i is missing in input, condition (2) requires $\sum_j J_{ij} g_x(x_j) > T_x$, or $\sum_j J_{ij} > T_x$ (for $g_x(x_j) \le 1$). If segment i and all segments j on the positive x-axis are missing in input, condition (3) requires $\sum_{j<0} J_{ij} < T_x$ if the strong contour should be prevented to grow into the positive x-axis. These requirements, together with the contour reflection symmetry $J_{ij} = J_{ji}$, constrain the overall scale of J within a factor of 2.

However, a segment near and roughly parallel to the x-axis is excited by the co-aligned contour segments on the x-axis, and is likely to be filled by condition (2) since it provides an alternative route for the contour. To prevent such a straight contour from thickening, condition (3) requires di-synaptic inhibition connections $W_{i\theta,j\theta'}$ connecting this segment from the non-aligned horizontal contour segments on the x-axis. A W connection struc-

[9] In Oct. 2000, an error was discovered in the published version of this paper for the expression of $W_{i\theta,j\theta'}$. The condition "$d/\cos(\beta/4) \ge 10$" to set $W_{i\theta,j\theta'} = 0$ in the expression below, was mistakenly typed in as "$d \ge 10$" instead in the published version. This error should lead to quantitative changes in the model behavior from those presented in this paper. Also, $\Delta\theta$ means $|\Delta\theta|$ in the expression for $W_{i\theta,j\theta'}$.

ture as in figure (4) results, a lower bound on the overall scale of the W connections can then be obtained.

Condition (1) discourages W connections between co-aligned segments, and condition (3) discourages J connections between roughly parallel segments displaced roughly orthogonally to their orientations. Our model connection is then derived from all these requirements, in addition to the space invariance (symmetry) requirements for the computation and the local and smoothness requirements expected as reasonable for cortical connections (see section 2).

References

[1] Allman J., Miezin F., and McGuinness E. "Stimulus specific responses from beyond the classical receptive field: neurophysiological mechanisms for local-global comparisons in visual neurons" *Ann. Rev. Neurosci.* 8:407-30, 1985.

[2] Alonso J. M., Cudeiro J., Perez. R., Gonzalez F., and Acuna C. "Influence of layer 5 of layer 6 of area 19 of the cat visual cortex on responses of cells in layer 5 of area 17 to stimuli of high velocity" *Exp. Brarin Res.* 93: 363-366, 1993.

[3] Baldi P. and Meir R. "Computing with arrays of coupled oscillators: An application to preattentive texture discrimination" *Neural Comput.* vol. 2, 458-471 (1990).

[4] Braun J., Niebur E., Schuster H. G., and Koch C. "Perceptual contour completion: A model based on local, anisotropic, fast-adapting interactions between oriented filters". (24th Annual Meeting of the Society for Neuroscience, Miami Beach, Florida, USA, November... Society for Neuroscience Abstracts, v.20, n.1-2, 1665 (1994).

[5] Douglas R.J. and Martin K. A. "Neocortex" in *Synaptic Organization of the Brain* 3rd Edition, Ed. G. M. Shepherd, Oxford University Press 1990.

[6] Eckhorn R. "Oscillatory and non-oscillatory synchronizations in the visual cortex and their possible roles in associations of visual features" *Progress in Brain Research* Vol. 102. 405-426, 1994

[7] Eckhorn R., Bauer R., Jordan W., Brosch M., Kruse W., Munk M., and Reitboeck H. J. " Coherent oscillations: a mechanism of feature linking in the visual cortex? Multiple electrode and correlation analysis in the cat." *Biol. Cybern.* 60:121-130, 1988.

[8] van Essen D. " Functional organization of primate visual cortex" in *Cerebral Cortex* Eds. A Peters and E G Jones, Plenum Press, New York, 259-329, 1985

[9] Field D.J., Hayes A., and Hess R.F. 1993. "Contour integration by the human visual system: evidence for a local 'association field'" *Vision Res.* 33(2): 173-93, 1993

[10] Fitzpatrick D. Private communication, 1996a.

[11] Fitzpatrick D. "The functional organization of local circuits in visual cortex: Insights from the study of tree shrew striate cortex" *Cerebral Cortex*, V6, N3, 329-341. 1996b

[12] Freeman W.J. "Simulation of chaotic EEG patterns with a dynamic model of the olfactory system" *Biol. Cybern.* 56(2-3):139-50. 1987

[13] Gilbert C.D. "Horizontal integration and cortical dynamics" *Neuron*. 9(1): 1-13. 1992

[14] Gilbert C.D. and Wiesel T. N. "Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex." *J. Neurosci.* 9(7): 2432-42. 1989

[15] Gilbert C.D. and Wiesel T.N. "Clustered intrinsic connections in cat visual cortex." *J Neurosci.* 3(5): 1116-33. 1983

[16] Gilbert C. D. and Wiesel T. N. " The influence of contextual stimuli on the orientation selectivity of cells in primary visual cortex of the cat." *Vision Res* 30(11): 1689-701. 1990

[17] Gray C.M. and Singer W. "Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex" *Proc. Natl. Acad. Sci. USA* 86: 1698-1702. 1989

[18] Grosof D.H., Shapley R.M. and Hawken M.J. "Macaque V1 neurons can signal 'illusory' contours" *Nature,* 365(6446):550-2. 1993

[19] Grossberg S. and Mingolla E. "Neural dynamics of perceptual grouping: textures, boundaries, and emergent segmentations" *Percept Psychophys.* 38(2): 141-71. 1985

[20] Guy G. and Medioni G. "Inferring global perceptual contours from local features." In *Proc. IEEE CVPR* 786-7. 1993

[21] Heeger D.J. "Normalization of cell responses in cat striate cortex" *Visual Neurosci.* 9, 181-197. 1992

[22] Heitger F. and von der Heydt R. " A computational model of neural contour processing: Figure-ground segregation and illusory contours", in *Proc. of 4th Intl. Conf. on Computer Vision* p. 32-40. 1993

[23] Hirsch J. A. and Gilbert C. D. " Synaptic physiology of horizontal connections in the cat's visual cortex." *J. Neurosci.* 11(6): 1800-9. 1991

[24] Hubel D. H. and Wiesel T. N. "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex." *J. Physiol.* 160. 106-154. 1962

[25] Hummel R. and Zucker S. W. "On the foundation of relaxation labeling processes " *IEEE Transactions on Pattern Analysis and Machine Intelligence* 5. 267-287, 1983

[26] Johnson R.R. and Burkhalter A. "Feedback connections in visual cortex contact inhibitory neurons" *Soc. Neurosci. Abst* 332.8 1991.

[27] Johnson R.R. and Burkhalter A. "Corticocortical feedback connections in visual cortex synapse selectively with dendritic spines" *Soc. Neurosci. Abst* 134.15 1992.

[28] Kapadia M.K., Ito M., Gilbert C.D. and Westheimer G. "Improvement in visual sensitivity by changes in local context: parallel studies in human ob servers and in V1 of alert monkeys." *Neuron*. Oct; 15(4): 843-56. 1995

[29] Kass M., Witkin A. and Terzopoulos D. SNAKE: active contour models *Int. J. Computer Vision* Vol. 1. 321-332. 1988

[30] Kienker P. K., Sejnowski T. J., Hinton G. E. and Schumacher L. E. "Separating figure from ground with a parallel network" *Perception* vol. 15, 197-216, 1986.

[31] Knierim J. J. and van Essen D. C. " Neuronal responses to static texture patterns ion area V1 of the alert macaque monkeys." *J. Neurophysiol.* 67, 961-980. 1992

[32] König P. and Schillen T. B. "Stimulus-dependent assembly formation of oscillatory responses: I Synchronization." *Neural Comput.* vol. 3, pp. 155-166, 1991

[33] Kovacs I. and Julesz B. " A closed curve is much more than an incomplete one: effect of closure in figure-ground segmentation." *Proc Natl Acad Sci USA*. 15; 90(16): 7495-7 1993

[34] Lamme V.A. "The neurophysiology of figure-ground segregation in primary visual cortex." *Journal of Neuroscience* 15(2):1605-15. 1995

[35] Li Zhaoping and Hopfield J.J. "Modeling the Olfactory Bulb and Its Neural Oscillatory Processings" Biol. Cybern. 61(5) 379-392, 1989.

[36] Li Zhaoping, "A model of olfactory adaptation and sensitivity enhancement in the olfactory bulb". *Biological Cybernetics*, 62/4 349-361 1990

[37] Li Zhaoping "Primary cortical dynamics for visual computation" in *Theoretical Aspects of Neural Computation* workshop in Hong Kong May 1997. See also Li Z. "Visual texture segmentation via intra-cortical neural dynamics in V1" submitted to NIPS 1997.

[38] Livingstone M. S. and Hubel, D. H. 1984. "Anatomy and physiology of a color system in the primate visual cortex." *J. Neurosci.* Vol. 4, No.1. 309-356. 1984

[39] Maffei L. and Fiorentini A. 1976 "The unresponsive regions of visual cortical receptive fields" *Vision Res.*; 16(10): 1131-9 1976

[40] Malach R., Amir Y., Harel M. and Grinvald A. "Relationship between intrinsic connections and functional architecture revealed by optical imaging and *in vivo* targeted biocyting injections in primate striate cortex." *Proc. Natl. Acad. Sci. USA* 90, 10469-10473. 1993

[41] von der Malsburg C. "The correlation theory of brain function." Internal report, Max-Planck-Institute for Biophysical Chemistry, Gottingen, West Germany. 1981

[42] von der Malsburg, C. and Buhmann, J. "Sensory segmentation with coupled neural oscillators." *Biological Cybernetics* vol.67, no.3 p233-42. 1992

[43] Marr D. *Vision, A computational investigation into the human representation and processing of visual information* Freeman, San Francisco 1982

[44] McGuire B.A., Gilbert C.D., Rivlin P.K. and Wiesel T.N. " Targets of horizontal connections in macaque primary visual cortex." *J-Comp-Neurol.* 305(3): 370-92 1991

[45] Mignard M. and Malpeli J.G. "Paths of information flow through visual cortex." *Science* 251: 1249-1251, 1991

[46] Moran J. and Desimone R. "Selective attention gates visual processing in the extrastriate cortex" *Science* 229 (4715): 782-4 1985

[47] Motter B. C. " Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli." *J. Neurophysiol.* 70(3):909-19. 1993

[48] Murata T. and Shimizu, H. " Oscillatory binocular system and temporal segmentation of stereoscopic depth surfaces." *Biological Cybernetics* vol.68, no.5 381-91. 1993

[49] Nault B., Michaud Y., Morin C., Casanova C. and Molotchnikoff S. "Responsiveness of cells in area 17 after local interception of the descending path from area 18" *Soc. Neurosci. Abst.* 502.7 1990

[50] Polat U. and Sagi D. " The architecture of perceptual spatial interactions." *Vision Res.* 34(1): 73-8. 1994

[51] Rockland K.S., and Lund J.S. "Intrinsic laminar lattice connections in primate visual cortex." *J. Comp. Neurol.* 216(3):303-18, 1983.

[52] Rockland K.S., 1996 Private communication, also see Rockland and Douglas, Society for Neuroscience Abstract. 179.6, 1993.

[53] Salin P. and Bullier J. “Corticocortical connections in the visual system: structure and function” *Physiological Reviews* Vol. 75. no.1, 107-154, 1995

[54] Sengpiel R., Baddeley R. Freeman T., Harrad R., and Blakemore C. *Soc. Neurosci. Abstr.* 21, 1629 (1995).

[55] Shashua A. and Ullman S. “Structural Saliency” *Proceedings of the International Conference on Computer Vision*. Tempa, Florida, 482-488. 1988

[56] Shepherd G. M. *Synaptic Organization of the Brain* 3rd Edition, Oxford University Press 1990.

[57] Sillito A.M., Grieve K.L., Jones H.E., Cudeiro J. and Davis J. “ Visual cortical mechanisms detecting focal orientation discontinuities” *Nature* 378(6556):492-6. 1995

[58] Singer W. and Gray C. M. “Visual feature intergration and the temporal correlation hypothesis” *Ann. Rev. Neurosci.* 18:555-86. 1995

[59] omers D. C., Nelson S.B. and Sur M. “An emergent model of orientation selectivity in cat visual cortical simple cells.” *Journal of Neuroscience* 15(8):5448-65. 1995

[60] Somers D. C., Todorov E. V., Siapas A. G., and Sur M. “Vector-based integration of local and long-range information in visual cortex” A.I. Memo. NO. 1556, MIT. 1995

[61] Sompolinsky H., Golumb D., and Kleinfeld D. “Cooperative dynamics in visual processing” *Phys. Rev. A.* vol. 43, pp. 6990-7011, 1991

[62] Sporns O., Tononi G., and Edelman G. M. “Modeling perceptual grouping and figure-ground segregation by means of reentrant connections” *Proc. Natl. Acad. Sci. USA* Vol. 88, pp.129-133. 1991

[63] Ts’o D. and Gilbert C. “The organization of chromatic and spatial interactions in the primate striate cortex” *J. Neurosci.* 8: 1712-27. 1988

[64] Ts’o D.Y., Gilbert C.D., and Wiesel T.N. “ Relationships between horizontal interactions and functional architecture in cat striate cortex as revealed by cross-correlation analysis.” *J-Neurosci.* 6(4): 1160-70 1986

[65] Valverde F. “ The organization principles of the primary visual cortex in the Monkey” in *Cerebral Cortex* Eds. A Peters and E G Jones, Plenum Press, New York, p. 207-258. 1985

[66] von der Heydt R., Peterhans E. and Baumgartner G. “Illusory contours and cortical neuron responses” *Science* vol. 224, pp. 1260-1262, 1984.

[67] Wang D. "Synchronous oscillations based on lateral connections" in *Lateral interactions in the cortex: Structure and function* Electronic book, ISBN 0-9647060-0-8 http://www.cs.utexas.edu/users/nn/web-pubs/htmlbook96/ Eds: Sirosh J. Miikkulainen R. and Choe Y. 1995

[68] Weliky M., Kandler K., Fitzpatrick D. and Katz L. C. " Patterns of excitation and inhibition evoked by horizontal connections in visual cortex share a common relationship to orientation columns" *Neurons* Vol. 15, 541-552, 1995.

[69] White E. L. *Cortical circuits* 46-82, Birkhauser, Boston, 1989

[70] Williams, L. R. and Jacobs D. W. "Local parallel computation of stochastic completion fields" In *Proc. IEEE CVPR.* p.161-8 1996

[71] Yen S-C. and Finkel L.H. "Salient contour extraction by temporal binding in a cortically-based network" *Natural Information Processing Systems* Denver. 1996

[72] Zipser K. Lamme V.A., and Schiller P.H. " Contextual modulation in primary visual cortex." *J. Neurosci.* 15, 16(22):7376-89. 1996

[73] Zucker S. W., Dobbins A. and Iverson L. "Two stages of curve detection suggest two styles of visual computation" *Neural Computation* 1, 68-81, 1989