A model of contextual interactions and contour detection in primary visual cortex

Mauro Ursino*, Giuseppe Emiliano La Cara

Department of Electronics, Computer Science, and Systems, University of Bologna, Cesena, Italy

Received 30 September 2003; revised 29 March 2004; accepted 29 March 2004

Abstract

A new model of contour extraction and perceptual grouping in the primary visual cortex is presented and discussed. It differs from previous models since it incorporates four main mechanisms, according to recent physiological data: a feed-forward input from the lateral geniculate nucleus, characterized by Gabor elongated receptive fields; an inhibitory feed-forward input, maximally oriented in the orthogonal direction of the target cell, which suppresses non-optimal stimuli and warrants contrast invariance; an excitatory cortical feedback, which respects co-axial and co-modularity criteria; and a long-range isotropic feedback inhibition. Model behavior has been tested on artificial images with contours of different curvatures, in the presence of considerable noise or in the presence of broken contours, and on a few real images. A sensitivity analysis has also been performed on the role of intracortical synapses.

Results show that the model can extract correct contours within acceptable time from image presentation (30–40 ms). The feed-forward input plays a major role to set an initial correct bias for the subsequent feedback and to ensure contrast-invariance. Long-range inhibition is essential to suppress noise, but it may suppress small contours due to excessive competition with greater contours. Cortical excitation sharpens the initial bias and improves saliency of the contours.

Model results support the idea that contour extraction is one the primary steps in the visual processing stream, and that local processing in V1 is able to solve this task even in difficult conditions, without the participation of higher visual centers.

Keywords: Primary visual cortex; Neural network; Contour; Orientation selectivity; Feed-forward and feedback mechanisms; Inhibition; Excitation; Saliency

1. Introduction

One of the fundamental tasks that the visual cortex can perform with apparent easiness is grouping different parts of the visual scene into organized objects. A pivotal role in this process is played by the Gestalt psychology criteria, which govern the organization of perception: they include proximity, continuity, similarity, common fate and closure (Koffka, 1935; Wertheimer, 1938). According to these criteria, perception of objects in any visual process depends on contextual influences.

Particularly, extraction of contours plays a pivotal role in any object recognition task. It is generally assumed that contours can ‘pop out’ from noisy environment if they satisfy the contiguity and continuity criteria, i.e. if successive elements in the contour are closed together and share similar orientations.

Traditionally, the process of global perception (including contour extraction) has been ascribed to high-order cortical areas. However, several recent psychophysical studies (Field, Hayes, & Hess, 1993; Kapadia, Ito, Gilbert, & Westheimer, 1995; Kapadia, Westheimer, & Gilbert, 2000; Kovacs & Julesz, 1993; Polat & Sagi, 1993, 1994), as well as anatomical and physiological investigation (Fitzpatrick, 1996; Gilbert, 1992; Gilbert, Das, Ito, Kapadia, & Westheimer, 1996; Kapadia et al., 1995, 2000; Polat, Mizobe, Pettet, Kasamatsu, & Norcia, 1998; Sillito, Griewe, Jones, Cudeiro, & Davis, 1995) reveal that contextual interactions may occur even at the early levels of visual processing, and that the primary visual cortex (V1) plays a role in the pre-attentive spatial integration.

Although the receptive fields of V1 cells, as measured by spot stimuli, are quite small, and signal just a tiny edge of specific orientation (Ferster & Miller, 2000;
Hubel & Wiesel, 1962), V1 cells have been observed to modify their response depending on the presence of surrounding stimuli, i.e. stimuli located outside their classic receptive field (Kapadia et al., 1995, 2000; Nelson & Frost, 1985). These contextual influences may allow V1 neurons to integrate spatial information from a large portion of the visual scene, in order to group perception into coherent features, such as contours or surfaces.

A common idea is that the main responsible for these contextual influences are long-range horizontal intracortical connections, emanating from excitatory cortical pyramidal cells (Gilbert, 1992; Gilbert & Wiesel, 1989; Rockland & Lund, 1983). These connections reach both excitatory and inhibitory post-synaptic cells, thus accomplishing both excitatory and inhibitory feedback connections inside the cortex.

Various recent physiological studies provide indications on the possible spatial arrangement of the facilitatory synapses. Field et al. (1993) observed that detection of a contour depends primarily on the relative orientation of its elements, in accordance with the continuation Gestalt rule: target could be detected if adjacent contour elements differed by less than 60°. Hence, the authors suggested the existence of an ‘association field’ linking neurons proximally located and with close orientation preference. Several other experimental studies suggest that excitatory horizontal connections link sets of neurons with similar orientation preference (Gilbert & Wiesel, 1989; Rockland & Lund, 1983; Ts’o, Gilbert, & Wiesel, 1986). This property has also been named ‘modular specificity’ (Shouval, Goldberg, Jones, Beckerman, & Cooper, 2000).

A second important property of excitatory synapses, named ‘axial specificity’, implicates that two neurons are connected by excitatory synapses only if their receptive fields are approximately displaced along an axis in visual space which correspond to the preferred orientation (Schmidt, Goebel, Lowel, & Singer, 1997; Shouval et al., 2000). This property implies that the response of a neuron to an optimally oriented bar is facilitated by a collinear bar placed outside its receptive field (Nelson & Frost, 1985; Polat et al., 1998). This effect decreases with the distance between the receptive fields (Kapadia et al., 1995, 2000; Polat & Sagi, 1993, 1994) and with the misalignments between the preferred orientations (Kapadia et al., 1995).

Besides the facilitatory influences mentioned above, inhibitory surrounding effects have also been demonstrated for V1 neurons (Bishop, Coombs, & Henry, 1973; Hubel & Wiesel, 1965; Li & Li, 1994). These inhibitory effects are especially located in regions orthogonal to the receptive field, although a large weaker field of diffuse isotropic inhibition is also well evident (Kapadia et al., 2000). While the excitatory links are thought to play a role in the contextual pop out of smooth contours, according to contiguity and continuation rules, inhibitory synapses may be important for noise suppression and to avoid the propagation of uncontrolled excitation in the network, i.e. instability.

A deeper insight into the mechanisms by which contextual influences in V1 may achieve contour integration and improve contour saliency can be obtained with the use of computational models and computer simulations. These models may assess whether local synapses are sufficient to extract contours without the participation of high-order visual processing stages, and may suggest the best disposition of intracortical synapses. Moreover, they may allow the mechanisms at the basis of psychophysical results to be explained in rigorous quantitative terms, and further experiments suggested. Finally, models based on V1 physiology may provide useful suggestions for the solution of artificial vision problems.

Indeed, several neural models of contour enhancement have been proposed in recent years with different purposes. Earlier relevant models are those by Grossberg and Mingolla (1985) and Zucker, Dobbins, and Iverson (1989) which, however, are not based on recent physiological findings. More recent models exploit the presence of co-axial excitatory connections, including some kind of modular and axial specificity to emphasize smooth contours, while an inhibitory normalization rule obliges the network to converge (Li, 1998; Pettet, McKee, & Grzywacz, 1998; Yen & Finkel, 1998). These models are able to simulate various psychophysical and physiological effects, and succeed in identifying and enhancing salient contours. The models by Li (1998) and Yen and Finkel (1998) use oscillatory neurons, which synchronize over a fast time scale, to represent contours. Results of these models demonstrate that interactions, similar to those postulated in V1, are able to emphasize contours, without the need for the participation of additional higher centers. However, the two models use a different disposition for intracortical synapses (for instance, Li assumes that cells with orthogonal preferred orientation are connected by inhibitory synapses, whereas Yen and Finkel assume a facilitatory connection, that they call ‘trans-axial’, between orthogonal cells). A different approach was used by VanRullen, Delorme, and Thorpe (2001). These authors realized contour integration using spike generating neurons and feed-forward mechanisms, avoiding iterative feedback mechanisms.

A recent model was developed by Ross, Grossberg, and Mingolla (2000), based on a previous sketch of possible intracortical circuits provided by Grossberg, Mingolla, and Ross (1997). This is probably the most sophisticated and advanced model of perceptual grouping presently available. However, it differs from the others, and from the present one, since the focus is on the laminar organization of the visual cortex, including interlaminar cortical circuits, a corticogeniculate feedback and the cortical area V2. By contrast, the previous studies, and the present, are aimed at investigating the organization and functional role of synapses in V1 only.
Despite these recent important contributions, there are still problems and unsolved questions which may deserve further analysis. First, processing in V1 cells is the result of a sophisticate combination of both feed-forward and feedback mechanisms [see Ferster and Miller (2000) for a discussion and Ringach, Bredfeldt, Shapley, and Hawken (2002)]. Previous models, with the exception of the comprehensive model by Ross et al. (2000), emphasize just feedback (Li, 1998; Pettet et al., 1998; Yen & Finkel, 1998) or feed-forward (VanRullen et al., 2001), without caring the concurrent action of the two mechanisms. Moreover, in these models description of cortical cells is simplified, generally assuming that cells extract orientation without a spatial description of their receptive field (Li, 1998; Yen & Finkel, 1998) or that only one cell in each hypercolumn is simultaneously active (Pettet et al., 1998). Moreover, none of the previous models investigated the effect of changes in excitatory and inhibitory synapses on contour saliency. Are V1 networks robust versus changes in synapse disposition (provided the general rules delineated above are fulfilled) or a moderate change in synapses can induce large variations in the final perception of contours? How a different strength of excitation versus inhibition affects the nature of contour extraction? A further problem, raised by VanRullen et al. (2001), concerns the time required for a feedback mechanism to converge. Is this time compatible with psychological results on object recognition? How feedback and feed-forward mechanisms participate in this process?

The aim of this work is to present a model of contour enhancement in V1, based on recent physiological data, which overcomes some of the limitations mentioned above and aspires to analyze the previous problems. The model implements a feed-forward mechanism, which sets an initial orientation preference and enhances contours, and a feedback intracortical mechanism, which exploits a narrow coaxial excitation but a broader non-axial inhibition. Simulations are aimed at investigating how feed-forward and feedback synapses may affect the characteristics of extracted contours and influence the convergence time. The results may be useful to provide indications on the possible mechanisms (feed-forward + feedback) implemented in V1, and may provide useful suggestions for contour extraction tools in artificial vision.

2. Model description

As originally described in Hubel and Wiesel (1962), the primary visual cortex is composed of ‘hypercolumns’, which consist of cells responding to the same spatial position in the retina, but with different orientation preferences, and ‘orientation columns’ which consist of cells responding to the same orientation but with different position in the visual space (see Fig. 1). In the present model, we consider an array of 50 × 50 hypercolumns, with a number of distinct orientation preferences for the excitatory neurons in each hypercolumn equal to 16. Moreover, each hypercolumn includes four inhibitory interneurons, which also have a distinct orientation preference. The ratio between excitatory neurons and interneurons agrees with that reported in the physiological literature (Gabbott & Somogyi, 1986). According to neurophysiological data, model assumes that orientation selectivity varies gradually within a hypercolumn. Hence, two consecutive excitatory neurons in the hypercolumn have orientation preferences which differ by 11.25°, while orientation preferences of consecutive interneurons differ by 45°.

In the following, the notation \((x, y, \theta)\) will be used to represent a cell whose receptive field is positioned at the coordinate \(x, y\) of the visual space, with preferred orientation, \(\theta\). In particular, the subscript \(c\) (hence, the notation \((x_c, y_c, \theta_c)\)) will be used to denote an excitatory cortical cell, while the symbol \(i\) (hence \((x_i, y_i, \theta_i)\)) a feed-forward inhibitory interneuron. When two excitatory cells are connected with an intracortical synapse, we will denote with the subscripts \(c\) and \(h\) (hence, with coordinates \((x_c, y_c, \theta_c)\) and \((x_h, y_h, \theta_h)\)) the post-synaptic and presynaptic cells, respectively.

2.1. The geniculate input

As it is well known, cortical cells receive their afferent inputs from cells in the lateral geniculate nucleus (LGN). In the present model, we assume that the receptive field of each cortical cell consists of a central ON region (that is, a region excited by light) surrounded by two lateral OFF regions (excited by darkness). Each region is elongated along a preferred direction. According to Jones and Palmer (1987a, b), these spatial receptive fields can be reproduced fairly well using a Gabor function.

If we consider a cortical cell whose receptive field is centered at the position \(x_c, y_c\) of the visual image, with preferred orientation \(\theta_c\), the receptive field assumes the following expression

\[
R(x - x_c, y - y_c, \theta_c) = R_0 \exp \left( -\frac{v_1^2}{2\sigma_1^2} \right) \exp \left( -\frac{v_2^2}{2\sigma_2^2} \right) \cos(2\pi f v_2) \tag{1}
\]
where
\[ v_1(x - x_c, y - y_c, \theta_c) = (x - x_c)\cos \theta_c + (y - y_c)\sin \theta_c \]  
(2)
\[ v_2(x - x_c, y - y_c, \theta_c) = -(x - x_c)\sin \theta_c + (y - y_c)\cos \theta_c \]  
(3)

Eq. (1) represents the Gabor function; Eqs. (2) and (3) describe a rotation of the receptive field by an angle \( \theta_c \) around the central point of coordinate \((x_c, y_c)\). \( x, y \) represent a generic coordinate in the input image, \( \sigma_1^2 \) and \( \sigma_2^2 \) are spatial variances, which establish the dimension of the receptive field in the preferred and non-preferred orientations, and \( \theta_c \) is a spatial frequency, which determines the width of the ON and OFF subregions.

Starting from Eqs. (1)–(3), the geniculate input to the cortical cell (say \( g(x_c, y_c, \theta_c) \)) is obtained by performing the inner product of the visual image, \( I(x, y) \), and the receptive field, i.e.
\[ g(x_c, y_c, \theta_c) = \left[ \int \int I(x, y)R(x - x_c, y - y_c, \theta_c)dx \ dy \right]^+ \]
\[ \equiv \left[ \sum_{m=1}^{N} \sum_{n=1}^{M} I(m\Delta x, n\Delta y) \times R(m\Delta x - x_c, n\Delta y - y_c, \theta_c)\Delta x \ \Delta y \right]^+ \]  
(4)

The two sums in the right hand member of Eq. (4) signify that the two-dimension integral has been approximated with the histogram method, and \( \Delta x, \Delta y \) represent the dimensions of the single pixel in the input image. \( N \) and \( M \) are the number of pixels in the horizontal and vertical directions. In this work, input images with 101 \( \times \) 101 pixels were used. Finally, the symbol \([ \ ]^+\) denotes the positive part.

### 2.2. The intracortical synapses

In the model, we considered three kinds of intracortical synapses. They have a different role in the process of contour extraction:

(i) a short-range feed-forward inhibition (synapses \( W_i < 0 \));
(ii) a mid-range feedback intracortical excitation (synapses \( W_i > 0 \));
(iii) a long-range feedback intracortical inhibition (synapses \( W_i < 0 \)).

The three kinds of connections aspire to reproduce the spatial distribution of contextual interactions in V1, according to data reported in Kapadia et al. (1995, 2000). Relationships with anatomical data will be discussed in the last section. While the connections described at points (ii) and (iii) share some aspects with those used in previous neural models (Li, 1998; Pettet et al., 1998; Yen & Finkel, 1998), the feed-forward inhibition at point (i) is quite original, and warrants contrast-invariance of the orientation tuning.

In the following, the general structure of synapses is first presented. Subsequently, each kind of synapse is described and justified in detail.

#### 2.2.1. The general form of synapses

As a general rule, we assume that the synaptic strengths decrease with the distance between the centers of the receptive fields of the pre-synaptic and post-synaptic cells. Moreover, we assume that the strength of the synapses also depends on the angle between the preferred orientation of the post-synaptic cell (\( \theta_p \)) and the line connecting the centers of the receptive fields (see Fig. 2). In other words, the strength of synapses may be different along the direction of the receptive field compared with the orthogonal direction. This choice implements the co-axial (or transaxial) specificity principle. Looking at Fig. 2, we can write
\[ W_j = W_{0j} \exp \left( \frac{-d_0^2}{2\sigma_0^2} \right) \exp \left( \frac{-d_{\theta_0-\pi/2}^2}{2\sigma_{\theta_0-\pi/2}^2} \right) \]  
(5)

where \( W_j \) is a generic synapse, \( d \) represents the distance between the centers of the receptive fields, and \( d_0, d_{\theta_0-\pi/2} \) represent the projections along the preferred and non-preferred orientations of the post-synaptic cell. The subscript \( j \) indicates the kind of synapse (either \( j = i, e \) or \( l \) for the three type of synapses). According to Fig. 2, we have:
\[ d = \sqrt{(x_c - x)^2 + (y_c - y)^2}; \]
\[ \varphi = \arctan((y_c - y)/(x_c - x)) \]  
(6')

Fig. 2. Computation of the horizontal connectivity between a pre-synaptic and a post-synaptic (target) cell, located in two different hypercolumns. The center of the figure represents the center of receptive field of the target cell, while \( O \) represents the position of the receptive field of the pre-synaptic cell. \( \theta_c \), preferred orientation of the target cell; \( \varphi \), orientation of the line connecting the two receptive fields; \( d_0 \), distance between the two receptive fields; \( d_{\theta_0-\pi/2} \), projection of the distance along the preferred and non-preferred directions, respectively.
\[ d_0 = d \cos(\varphi - \theta_0); \quad d_{\theta + \pi/2} = d \sin(\varphi - \theta_0) \quad (6') \]

where \( \varphi \) is the orientation of the line connecting the pre-synaptic and post-synaptic cells and \( x, y \) are the coordinates of the pre-synaptic cell. However, it is worth noting that the present model does not include self-coupling terms, i.e. a neuron is not connected with itself.

Eq. (5) implies that the synaptic strengths decrease with distance according to a classical Gaussian function, with a different variance \( (\sigma^2_0 \text{ and } \sigma^2_{\theta + \pi/2}) \) in the preferred and non-preferred directions. Eq. (5) applies to all intracortical synapses, but with different choices of parameters \( \sigma^2_0 \) and \( \sigma^2_{\theta + \pi/2} \). Eqs. (6') and (6") consider only the distance between hypercolumns (i.e. the distance between the centers of the receptive fields) and implements the axial specificity.

In order to implement the modular specificity, the quantity \( W_0 \) in Eq. (5) may not be constant, but may depend on the difference between the orientation preferences of the pre-synaptic and post-synaptic cells (i.e. the distance inside the hypercolumn).

In the following, the maps of the different synapses are described in detail.

### 2.2.2. The map of feed-forward inhibition

First we assumed that cortical cells, besides feed-forward input from the LGN (i.e. quantity \( g(x_i, y_i, \theta_i) \) in Eq. (4)), also receive feed-forward inhibition from cortical interneurons located in the same hypercolumn or in close hypercolumns. As pointed out by Ferster and Miller (2000) and Ringach et al. (2002), this inhibition is essential to account for the property of contrast invariance of orientation selectivity. Without feed-forward inhibition, in fact, a stimulus with high-contrast but not-optimal orientation would evoke a suprathreshold response, i.e. the width of the orientation-tuning curve of cortical neurons would increase with contrast. On the contrary, experimental results on cats and monkeys [see Ferster and Miller (2000)] agree in showing that the orientation selectivity of V1 cells is largely independent of contrast.

Mathematical models of orientation selectivity in V1 demonstrate that contrast-invariance can be explained by adding an inhibitory input to the excitatory input from relay cells [see Ferster and Miller (2000), p. 454 for a discussion]. This inhibitory input may come from inhibitory interneurons with identical receptive field but opposite polarity (i.e. ON instead of OFF, OFF instead of ON), thus realizing a sort of push–pull mechanism (Troyer, Krukowski, Priebe, & Miller, 1998) or may come from interneurons with the same polarity as their target cells (Somers, Nelson, & Sur, 1995). A critical summary of these models can be found in a recent paper by our group (Ursino & La Cara, 2004).

In the present model, we decided to implement the feed-forward mechanism using inhibitory neurons with the same polarity as their target excitatory cells, but with a ratio 1:4 (Gabbott & Somogyi, 1986). In order to warrant contrast independence of orientation tuning, this inhibition must exhibit a broad dependence on the orientation difference of the pre-synaptic and the post-synaptic cells (Somers et al., 1995; Ursino & La Cara, 2004). This assumption is further supported by the observation that the response to a grating of optimal orientation is suppressed by a second non-optimally oriented grating, and this inhibition is quite similar in magnitude for all orientations of the superimposed grating (De Angelis, Robson, Ozhawa, & Freeman, 1992).

Accordingly, we assumed that feed-forward inhibitory synapses from interneurons are independent of the orientation difference (i.e. of the difference \( \theta_i - \theta_j \)), but depends only on the distance between the receptive fields (i.e. the distance between the hypercolumns).

A value to the parameters \( \sigma^2_0 \) and \( \sigma^2_{\theta + \pi/2} \) in Eq. (5) can be assigned by considering recent data published by Kapadia et al. (2000). According to these authors, V1 neurons receive a strong inhibitory contextual interaction preferentially along a direction in the visual space orthogonal to the preferred orientation of the post-synaptic cell (i.e. we have \( \sigma^2_0 < \sigma^2_{\theta + \pi/2} \) in Eq. (5)). The map is illustrated in Fig. 3, left panel. This contextual arrangement may have a supplementary functional role, beside warranting contrast invariance of the orientation tuning. With this choice of parameters, in fact, the feed-forward mechanism from interneurons may help contour extraction, by inhibiting the response of cortical cells which have a poorer probability of belonging to a smooth contour.

In order to compute the feed-forward inhibitory input to cortical cells, we need to calculate the output response of interneurons. In the model, these interneurons receive just the thalamic input (Fig. 3, bottom panel). Moreover, in this model we assume that the response of all cortical cells (both excitatory and interneurons) exhibits a lower threshold and upper saturation. This is realized by means of a sigmoidal function. With this choice of parameters, in fact, the feed-forward mechanism from interneurons may help contour extraction, by inhibiting the response of cortical cells which have a poorer probability of belonging to a smooth contour.

Starting from Eq. (7), and using the symbols \( W_{ij} \) to denote a synapse linking the interneuron to a cortical cell, we can compute the feed-forward inhibition to the target cell (say \( \theta_i \)) as follows:

\[ i(x_i, y_i, \theta_i) = \sum_j \sum_{\gamma_i} W_{ij}(x_i, y_i, \theta_i, x_j, y_j, \theta_j) \cdot \rho(x_i, y_i, \theta_i) \quad (8) \]

where the expression of \( W_{ij} \) is computed according to Eqs. (5), (6') and (6") with suitable parameter values (and the subscript \( j = i \)). Since these synapses are inhibitory, we have \( W_{ij} < 0 \) in Eq. (5).
2.2.3. The map of the excitatory intracortical synapses

These synapses realize a feedback mechanism within the visual cortex, which improves and sharpens the information coming from feed-forward inputs.

According to results reported in various physiological and psychophysical experiments (summarized in Section 1), we assume that excitation among cortical cells is maximal if the pre-synaptic cell lies along the preferred orientation of the post-synaptic cell (co-axial specificity). In other words, individual neurons in the cortex receive excitation from other neurons whose receptive field center is located along an axis in the visual space that coincides with the preferred orientation of the post-synaptic cell (co-axial specificity). In other words, individual neurons in the cortex receive excitation from other neurons whose receptive field center is located along an axis in the visual space that coincides with the preferred orientation of the post-synaptic cell (co-axial specificity). The situation is illustrated in Fig. 3 middle panel, and corresponds to the criterion named axial specificity in Shouval et al. (2000). Moreover, excitatory intracortical synapses also satisfy the second criterion, named modular specificity (Shouval et al., 2000). This means that excitatory synapses may link only intracortical neurons which share similar orientation preference; i.e. the synaptic strength decreases with the difference in the orientation preference between the pre-synaptic and the post-synaptic neurons.

The two previous rules for intracortical excitatory synapses can be summarized by the following equation, which is a particular case of Eq. (5).

\[ W_e(x_c, y_c, \theta_c, x_h, y_h, \theta_h) = W_{0,e} \exp\left(\frac{-\left(\theta_c - \theta_h\right)^2}{2\sigma^2_{\theta,e}}\right) \exp\left(\frac{-d^2_\theta}{2\sigma^2_{\theta,e}}\right) \exp\left(\frac{-d^2_{\theta+\pi/2}}{2\sigma^2_{\theta+\pi/2,e}}\right) \]

(9)

where \( \theta_c \) and \( \theta_h \) are the orientation preferences of the post-synaptic and pre-synaptic cells, respectively. The first exponential in the right hand member of Eq. (9) means that the strength of the synapses (i.e. parameter \( W_0 \) in Eq. (5)) decreases with the difference of the orientation preference. \( \sigma^2_{\theta,e} \) establishes the decrease in the synaptic strength with the difference in the orientation preference, and the meaning of the other symbols is analogous as in Eq. (5). Since these synapses are excitatory, we have \( W_{0,e} > 0 \).

It is worth noting that Eq. (9) describes both the excitatory connections within an orientation column (i.e. when \( \theta_c = \theta_h \)) and the connections within an hypercolumn (i.e. when \( d_\theta = d_{\theta+\pi/2} = d = 0 \)). In the first case, Eq. (9) reduces to Eq. (5). In the second case, the synaptic strength...
Eq. (9) has a functional role to improve extraction of smooth contours. According to this equation, in fact, excitation is high only among neurons which form a smooth contour, since only in this case both axial and modular specificity rules are simultaneously satisfied (Fig. 4). By contrast, neurons which do not concur to the formation of a smooth contour, since only in this case the function $r(x, y, \theta)$ represents the firing rate (i.e. the output) of a cortical cell, as specified at point (v).

2.2.4. The map of the inhibitory (long-range) synapses

According to Kapadia et al. (2000), a large weak field of diffuse inhibition is evident in the contextual map of V1 neurons. It is credible that this additional inhibition works to improve noise elimination, by realizing a competitive mechanism among cortical neurons: neurons which form a smooth contour, and which are mutually excited (according to Eq. (5)) with lower threshold and upper saturation. Finally, the response of a generic excitatory cortical cell via these lateral inhibition (say $l(x_c, y_c, \theta_c)$) is computed as:

$$l(x_c, y_c, \theta_c) = \sum_{x_h} \sum_{y_h} W_l(x_c, y_c, \theta_c, x_h, y_h, \theta_h) \cdot r(x_h, y_h, \theta_h)$$

where the meaning of symbols is as in Eq. (10).

2.2.5. The response of the cortical excitatory cells

Finally, the response of a generic excitatory cortical cell is obtained by computing the sum of its input quantities and passing it through a sigmoidal relationship (similar to Eq. (7)) with lower threshold and upper saturation. Finally, as a consequence of feedback interactions, the activity of a generic cortical cell exhibits time evolution. This is summarized by means of first-order differential equation with time constant $\tau$.

Hence, we can write

$$u(x_c, y_c, \theta_c) = g(x_c, y_c, \theta_c) + i(x_c, y_c, \theta_c) + e(x_c, y_c, \theta_c) + l(x_c, y_c, \theta_c)$$

$$\frac{dr(x_c, y_c, \theta_c)}{dr} = - r(x_c, y_c, \theta_c) + S((1 - e^{-2u(x_c, y_c, \theta_c)})^+ - (1 + e^{-2u(x_c, y_c, \theta_c)})^+)$$

where the input quantities $g, i, e$ and $l$ are provided by Eqs. (4), (8), (10), and (12), $u$ is the global input to the cortical cell, and $S$ is the upper saturation level of cortical cell activity.

Finally, the output from the network at position $(x_c, y_c)$, is computed as the sum of all activities in the same hypercolumn. By denoting with $z(x_c, y_c)$ this output, we
have

$$z(x_c, y_c) = \sum_{\theta_c} r(x_c, y_c, \theta_c)$$  \hspace{1cm} (15)

3. Parameter assignment

A value to the parameters in the model has been given on the basis of various complementary criteria:

1. Parameters which describe the classic receptive field of V1 cells (i.e. the parameters $R_0$, $\sigma_1$, $\sigma_2$ and $f$ in the Gabor function, Eq. (1)), have been given to reproduce experimental results obtained by fitting Gabor functions to spatial receptive fields measured with the reverse correlation technique (Jones & Palmer, 1987a,b). Since a large variability of results is obtained experimentally, the present parameters are just within the range of values reported in the literature.

2. Values to the standard deviations of the synaptic maps linking different hypercolumns (i.e. parameters $\sigma_{\theta j}$ and $\sigma_{\theta + \pi / 2 j}$ in Eq. (5), with $j = i$, $e$ or $l$) have been given to mimic results of physiological and psychophysical experiments reported in Kapadia et al. (2000). As described in previous paragraphs, these results suggest that excitation develops mainly along the preferred direction of the target cell, whereas short-range inhibition is mainly located along the orthogonal direction. Moreover, excitation extends along a distance which is approximately 2 or 3-fold longer than inhibition (see Figs. 5–8 in Kapadia et al. (2000)). The exact value of extension, however (which describes the range of contextual influences) depends on the eccentricity in the retina. In the case of foveal stimuli, the excitatory map extends for about 0.5° in the preferred direction, while the short range inhibitory map extends for about 0.1–0.2°. By contrast, in the near periphery, at an eccentricity of about 4°, the excitatory and inhibitory maps have greater extension (about 1 and 0.5°, respectively). In the present study, we used the standard deviations reported in Table 1. Since the Gaussian function is almost zero after three standard deviations, the excitatory and inhibitory maps in our model extend by about 0.6 respectively.

The standard deviation of the broad inhibition (parameter $\sigma_i$ in Eq. (11)) has been first assigned assuming that this inhibition extends over a long distance, in order to help noise elimination over a wide region of the visual space. As will be discussed later, however, excessive long-range inhibition may disrupt proximal contours and, moreover, may be questionable on the basis of anatomical data. These aspects will be analyzed in the sensitivity analysis section and in the final discussion.

3. The dependence of the excitatory map on the orientation distance within the hypercolumn (i.e. parameter $\sigma_{\Delta \theta e}$ in Eq. (9)) has been given on the basis of data reported in previous works (Field et al., 1993; Kapadia et al., 1995). According to these authors, in fact, the strength of facilitatory interaction is strongly reduced when the orientation distance between cells is about 60° (with a significant reduction already evident at about 40°). Hence, we used a standard deviation $\sigma_{\Delta \theta e} \equiv 20°$.

4. It is not easy to assign a numerical value to the strength of the synapses (i.e. to parameters $W_{\theta j}$ in Eq. (5)) on the basis of physiological and psychophysical experiments. Only the ratio among these synapses can be approximately deduced. A clear indication is that excitation is stronger than inhibition at short distances, while long-range inhibition is even weaker (Kapadia et al., 2000).

Hence, a value to these synapses has been given, in accordance with the previous constraint, using an a posteriori criterion: more precisely, we chose: (i) values for feed-forward inhibitory synapses which warrant an orientation-tuning curve in accordance with experimental results; (ii) values of feedback synapses which warrant a good contour extraction for some exemplary images, in the presence of wide random noise (see Section 4). The effect of changing these parameters is analyzed in the sensitivity analysis section.

5. The values of the upper saturation for firing rate activity (parameter $S$ in Eqs. (7) and (15)) agrees with the maximal firing rate reported for simple cells in some studies (Skottun, Bradley, Sclar, Ohzawa, & Freeman, 1987; Somers et al., 1995), although a large variability can be found to as this value. The time constant of neural dynamics (parameter $\tau$ in Eq. (14)) agrees with values used in deterministic mean-field equations [a few milliseconds, see Ben-Yishai, Bar, and Sompolinsky (1995)]. In particular,
this value can be chosen significantly lower than the membrane time constant (Treves, 1993).

A complete list of parameters is reported in Table 1.

4. Results

In order to verify model’s capacity to extract contours in the presence of noise, we used three exemplary images with different geometrical characteristics: (i) a square, characterized by only two orientations and sharp corners; (ii) a circle, characterized by all orientations without corners; (iii) a bottle, which exhibits all possible orientations but with different curvature and rounded corners. In all cases, we used images with 100% contrast and superimposed noise. Assuming that $I = 0$ represents the average luminance of the environment, we have $I(x, y) = 0.5 + v(x, y)$ in bright regions and $I(x, y) = -0.5 + v(x, y)$ in dark regions. $v$ is a Gaussian random noise with zero mean value and standard deviation 0.4.

The simulation results are shown in Figs. 5–7. Each figure displays the input image (upper left panel), the network output resulting from the feed-forward input only (right upper panel) and the network output at several steps of the numerical integration, when feedback mechanisms are operating.

Results show that the model can extract the contour quite exactly despite the very large amount of noise superimposed on the image. Moreover, the response of the feedback mechanisms converges to a steady-state solution in about 30–40 ms.

Examples of the orientation-tuning curve of single hypercolumns are shown in Fig. 8. The left panel shows the activity of all 16 excitatory neurons in the hypercolumn.
at position \( x_c = 35, y_c = 38 \) at the end of the simulation in Fig. 7 (‘bottle’). These neurons are responding to a vertical edge. The right panel shows the response of the 16 neurons in the hypercolumn at position \( x_c = 36, y_c = 36 \) at the end of the simulation in Fig. 5 (‘square’). Neurons are placed at the upper right corner, hence are responding to two simultaneous edges, with orientation 0 and 90°.

Fig. 9. Model response to a circle with broken contours. The meaning of panels is the same as in Fig. 5. It is worth noting that an intact contour is reconstructed by the intracortical feedback dynamic.

Fig. 10. Example of orientation-tuning curves within a hypercolumn at different levels of contrast. The figure refers to the simulation of a circle without noise, in a position where the hypercolumn responds to a vertical edge. It is worth noting that contrast changes do not affect the width of the orientation-tuning curve, but only the peak of the response is contrast dependent. Similar results can be obtained by using different figures.
and $\sigma_{\theta+\pi/2, e}$ in Eq. (9)). Indeed, these parameters were not assigned on the basis of physiological experiments, but using empirical a posteriori criteria. The sensitivity analysis is shown in Figs. 11 and 12 with reference to the bottle, but similar results were obtained on the other images too. The bottle was chosen since it exhibits more difficult contour extraction, especially at the convergence of the neck with the body.

Fig. 11 shows the results of the sensitivity analysis performed on parameters in the feedback inhibition, i.e. the strength $W_{0,l}$ (upper panels) and standard deviation $\sigma_l$ (bottom panels) of the synapses. Results show that the inhibitory feedback plays a major role in noise elimination. In fact, a reduction of the inhibition strength causes the appearance of large noise superimposed on the correct contour. However, inhibition strength cannot be increased too much. In the latter case, in fact, competition among active neurons would become too strong, thus resulting into a broken contour: only those neurons which belong to portions of the contour with low-curvature would remain active, whereas high-curvature portions of the contour would be totally inhibited.

Similarly, an increase in the standard deviation of feedback inhibition, which establishes the extension of the inhibitory contextual influence, results in a more polished image. However, we do not think that increasing this parameter is suitable: in fact, if inhibition is too wide, the contour of an image would interfere with another contour placed in a not-too-distant position (see Fig. 13 and Section 5 for more details). A decrease in this parameter causes a poor noise elimination, further emphasizing the importance of long-range inhibition in the management of isolated noise.

Fig. 12 shows results of the sensitivity analysis performed on excitatory synapses. As it is clear from this
figure, increasing the strength of excitatory synapses results in a broken contour, while some portions of the contour become too long and overcome the corners of the figure. The reason is that competitive mechanisms between neurons are too strong, thus causing suppression of weak segments (generally, located close to corners or where a contour abruptly changes its direction) and an excessive extension of strong segments. Moreover, rounded portions are often approximated by straight lines, resulting in a squared contour.

By contrast, if excitatory synapses are reduced down to zero, a correct contour still pops out from the image, although with low excitation. This result, quite unexpected, demonstrates that the combination of feed-forward mechanism + long-range inhibition is already able to extract contours. The main worsening compared with the basal case consists in the presence of some noise, and in the low-level of network activity.

An increase in the standard deviation of excitation in the preferred orientation, \( \sigma^2_{\theta=\omega} \), has a very negative effect. In fact, excitation spreads along straight lines, well beyond the termination of a line at a corner. Increasing the standard deviation in the orthogonal direction, \( \sigma^2_{\theta=\pi/2,\omega} \), results in a thicker contour (which might also be a positive effect). However, some ‘weak’ portions of a contour may be broken, since the higher level of excitation in the network increases the competitive mechanism. In fact, in this condition much noise is suppressed. Finally, increasing the standard deviations in both the optimal and the orthogonal directions results in a very thick contour, with almost complete suppression of any external and internal noise.

Finally, we applied our model to two real black and white photographs, which exhibit multiple close contours. The results are summarized in Fig. 13. As it is clear from this figure, the model, with basal parameter values, can recognize the strongest contours of these real images quite well, but it partially eliminates small contours or portions with great curvature. We think that this effect is a consequence of the wide extension used in our model for long-range inhibition (several degrees), which helps noise elimination (see Fig. 11) but may also cause competition between close contours. Hence, the simulation of real objects have been repeated with a significant reduction in the extension of long-range inhibition. In this condition, the model is able to detect also the secondary contours of the images quite well. Finally, if the extension of inhibition is further reduced, we can observe a significant noise superimposed on the image.

In conclusion, our simulations demonstrate a conflict between noise elimination (which requires a wide inhibition) and detection of proximal contours (which may benefit from a shorter inhibition). The problem of long-range inhibition will be further discussed in the final section of this paper.

5. Discussion

The objective of the present work was to develop a neural network model of object contour extraction in V1, based on physiological knowledge. Model behavior has then been tested using simple artificial images and a couple of black and white photos, and a sensitivity analysis has been performed in order to emphasize the role of the different synapses in the final results.

Results show that the particular arrangement of input from the LGN and intracortical contextual synapses yield satisfactory contour extraction, even if images are corrupted by high levels of random noise. This result supports the idea, already exploited in previous models (Li, 1998; Pettet et al., 1998; Yen & Finkel, 1998), that local processing in V1 is already able to emphasize contours, even without participation of high-level neural structures, and that contour extraction represents one of the first steps in the visual processing stream. In a recent paper, we suggested that contour extraction may help the solution of the binding and segmentation problem, favoring desynchronization of oscillators belonging to different portions of the visual scene (Ursino, La Cara, & Sarti, 2003).

The present model incorporates four main elements, the superimposition of which determines the final behavior: these are the classic elongated receptive fields of simple cells, arranged in hypercolumns and orientation columns according to a regular lattice; a feed-forward inhibition, which suppresses non-optimally oriented stimuli; a mid-range feedback excitation, which implements the co-axial and co-modularity criteria to emphasize smooth contours, and a weak but diffuse isotropic feedback inhibition, which aids noise suppression. We are not aware of previous models which incorporate all these elements into a single unitary theoretical structure.

Perhaps the main original aspect of this model, compared with previous ones, lies in the synergic cooperation of feed-forward and feedback mechanisms to achieve contour extraction form visual images. This cooperation ensures several benefits. First, an important function of feed-forward + feedback processing is to guarantee contrast invariance of the response, avoiding the so-called ‘iceberg effect’, i.e. a progressive loose of orientation selectivity for V1 cells with increasing contrast. This point has been thoroughly analyzed in previous studies (Ferster & Miller, 2000; Somers et al., 1995; Troyer et al., 1998; Ursino & La Cara, 2004). A second important aspect, particularly underlined by the present simulation results, is that cooperation of feed-forward and feedback is crucial to extract contours from noise avoiding too long a processing time. Recently, some authors (VanRullen et al., 2001) speculated that the inherent dynamic of a feedback intracortical mechanism would be too slow to converge to an equilibrium solution within the short time required by human or animal object recognition tasks (usually as short as 150–250 ms) (Thorpe, & Gauthrais, 1996; Vogels, 1999).
For this reason, the use of pure feed-forward has recently been proposed to extract contours from images in V1 models, thus avoiding time consuming iterative computation (VanRullen et al., 2001). The choice adopted in the present model assures an acceptable network convergence time, exploiting advantages of both feed-forward and feedback. (Of course, the term ‘convergence time’ is always used here to indicate time in the neural processing system, nor the time required for integration of numerical equations by the serial computer. In the present model, the latter is approximately as high as 1 h using a modern Pentium processor). The feed-forward mechanism realizes a preliminary contour extraction in a single step. This represents a strong initial bias for the subsequent feedback mechanisms, which implement a weak competitive dynamic. In this way, the feedback mechanisms already start reasonably close to the final equilibrium, and converge quite rapidly to the desired steady-state solution within an acceptable time. By using a physiological time constant for the feedback dynamic (Eq. (14)), and prudent values for excitatory and inhibitory synapses, we demonstrated that equilibrium is reached within 30–40 ms (Figs. 5–7). This time is small enough to allow subsequent visual processing layers to complete an object recognition task, within the time observed in physiological and/or psychophysical studies (Thorpe & Gautrais, 1996; Vogels, 1999).

The mathematical form adopted for the various parts of the model, and the numerical parameter values require a critical analysis, on the basis of the obtained results and indications from the sensitivity analysis.

The description of intracortical synapses in our model agrees with recent results by Kapadia et al. (1995, 2000). These authors, via physiological and psychophysical experiments, demonstrated the existence of different contextual influences among cells in V1: a map of intracortical excitation in the preferred direction of the target cell, a map of intracortical inhibition in the orthogonal direction, and a diffuse isotropic but weaker long-range inhibition. Our model demonstrates that this particular arrangement provides good contour extraction.

Short-range inhibition in the orthogonal direction is realized by means of a pure feed-forward, via inhibitory interneurons. As shown in previous studies, and in Fig. 10 of this work, this kind of inhibition is important to suppress stimuli with non-optimal orientation (i.e. stimuli in the orthogonal direction of the target neuron) and to achieve contrast invariance of the orientation tuning for each cell. However, looking at the literature, it is not easy to decide the exact arrangement for local inhibitory circuitry, and various alternative choices have been proposed, generally producing only minor differences in the final outcome [see Ferster and Miller (2000), Ringach et al. (2002), and Ursino and La Cara (2004) for a comparison among existing models]. In the present study, we adopted a circuitry for the feed-forward inhibition similar to that proposed by Somers et al. (1995), in which inhibitory interneurons share the same spatial phase as the target excitatory cell (i.e. ON versus ON and OFF versus OFF). A different choice, but which produces comparable final results, has been adopted in the push–pull model by Troyer et al. (1998), where inhibition is in spatial opposition with excitation (i.e. ON versus OFF and OFF versus ON). We think that a push–pull arrangement for short-range inhibition may be adopted in our model too, without appreciable changes in the final results, provided correct parameter values are selected for the corresponding synapses. In this study, we used the same spatial phase for the excitatory input from relay cells and inhibitory interneurons to avoid the inclusion of OFF cells in the model. This reduces the number of cells and so the numerical complexity.

A possibility (exploited, for instance, in the models by Ben-Yishai et al. (1995)) is that local inhibition does not involve a feed-forward scheme only, but also exploits short-range intracortical inhibitory feedback. This aspect is discussed in Ringach et al. (2002), where different schemes for suppression of non-optimal stimuli are critically examined, including pure feed-forward and cortical feedback. The feeling of these authors is that ‘purely feed-forward circuitry, trough inhibitory interneurons is doubtful given our knowledge of the cortical anatomy. It appears likely that inhibitory interneurons also receive substantial cortical input...’. In our model, we decided to maintain short-range inhibition as purely feed-forward for the sake of simplicity. It is possible that model behavior may be improved assuming that short-range inhibitory interneurons also receive a feedback input back from excitatory cortical cells. This might be attempted in future analysis of the model.

Intracortical excitations implement the co-axial and co-modularity criteria (Shouval et al., 2000). This aspect of the model, which agrees with experimental data by Kapadia et al. (1995, 2000), resembles that already used in previous recent models (Li, 1998; Pettet et al., 1998; Yen & Finkel, 1998). These contextual excitatory connections are important to sharpen the orientation-tuning curve (Fig. 8), to increase the level of neural activity, and to emphasize smooth contours starting from the initial bias provided by feed-forward activity. Another role of these excitatory synapses is to help the closure of broken contours (for instance, contours lacking of small portions or accidentally broken by noise, Fig. 9). However, we wish to remark that, in our model, the role of excitatory cortical synapses is not so absolute as assumed in other studies. In fact, a part of the contour extraction and noise elimination is accomplished by the feed-forward processing and by long-range intracortical inhibition. Looking at the sensitivity analysis in Fig. 12, for instance, we can observe that feed-forward mechanisms + intracortical long-range inhibition are already adequate to accomplish a discrete contour extraction, even in the absence of intracortical excitation (i.e. when \( W_r = 0 \)). Moreover, the sensitivity analysis shows that an excessive strength of intracortical excitation (simulated by increasing
excitatory strength or the standard deviations of synapses) may result either in a thicker contour or, sometimes, in a broken contour due to excessive competition among different segments. In other words, the competitive mechanism realized by excitation must not be excessively strong, to allow easy attainment of an adequate equilibrium condition. Moreover, increasing the strength of excitation causes a longer time for the network to converge to a steady-state level, with the risk of instability (for instance, the simulations in Fig. 12 performed by increasing synapses are still not in equilibrium, although equilibrium is approaching slowly).

Another aspect, which deserves a critical comment, concerns the extension of lateral excitation. In the present work excitation extends for about 0.6°, which agrees with results of psychophysical studies in Kapadia et al. (2000). However, other authors report that long-range connections from pyramidal excitatory neurons can expand to 6–8 mm in the cortex (Gilbert & Wiesel, 1989). This value signifies that excitatory connections can extend up to 2–3° close to the fovea. As shown in the sensitivity analysis of Fig. 12, the use of longer excitatory connections may favor the pop up of regular lines or smooth contours, but with the negative effect of a continuation of the lines well beyond their termination. We think that the use of longer excitatory connections can be implemented in the model in future works, but this requires the inclusion of end-stopped cells, which recognize corners and avoids spreading excitation beyond the end of a contour. In this regard, the use of longer excitatory connections may be at basis of the formation of illusory contours. Although illusory contours have been especially demonstrated with reference to cells in V2, more recent studies suggest that V1 cells can also respond to illusory contours although at a smaller proportion (Sheth, Sharma, Rao, & Sur, 1996). Analysis of illusory contours is beyond the aim of the present work, but it may represent a further application of the model, joined with a sensitivity analysis of long-range connections. We expect that excitatory connections in our model may help completion of illusory contours when two thick borders are aligned and are not affected by intermediate borders. However, in order to decide whether an illusory contour is actually formed or not, it will be important to have end-stopped cells in the model. Without the presence of end-stopped cells, illusory contours may appear repeatedly and inconsistently if excitatory connections have a too long-range extension, whereas the use of excitation with shorter extension probably excludes the formation of these contours. We will analyze illusory contours in a subsequent more complex study, after inclusion of end-stopped cells in the model and using both ON and OFF cortical cells together. However, it is worth noting that top-down strategies (not included in our model) may play a significant role in the formation of most illusory contours, which may require the use of more sophisticate models, including V2.

The long-range isotropic inhibition plays a major role in reduction of isolated noise, i.e. in eliminating contrast portions of the image which do not belong to any smooth contour. The use of these synapses, together with excitatory connections, implements a classic competitive mechanism: only neurons, which receive enough excitation from the feed-forward input and/or from coaxial and co-modular neurons can remain active, whereas all other neurons are silenced. Of course, the level of competition, hence the suppression of weakly excited neurons depends, beside the level of excitation, also on the strength and extension of these inhibitory synapses. A weak or narrow inhibition results in insufficient noise elimination. However, an excessive inhibitory strength is also inadequate, since, in this case, just a few segments can remain active, resulting in broken contours with low activity levels.

A fundamental question, which deserves attention, concerns the values adopted in the present model to describe the extension of long-range inhibition. In the present basal simulations (Figs. 5–7), we assumed that inhibition extends for several degrees in all directions. This choice agrees with Kapadia et al. (2000), who observed that ‘the four lobes of facilitation and inhibition were superimposed on a larger, weak field of diffuse inhibition’. More important, this choice was adopted to eliminate noise in our model even at large distance from the contours. It is possible, however, that in future studies this long-range inhibition should be critically re-considered for two major reasons:

(i) physiological data in the literature suggest that basket cells, which are the main responsible for horizontal long-range inhibition, have axonal fields covering up to 2.5–3 mm in the primary visual cortex (Buzas, Eysel, Adorjan, & Kisvarday, 2001; Kisvarday, Crook, Buzas, & Eysel, 2000; Kisvarday & Eysel, 1993). Of course, transforming this distance to degrees implies knowledge of the magnification factor of the cortex. At moderate eccentricity, close to the fovea (central vision) magnification is in the range 1–3 mm/deg (but may become even as high as 15 mm/deg in the very central point) (Levi, Klein, & Aitsebaomo, 1985; Virsu & Hari, 1996). At greater eccentricity (10°) magnification may be as low as 0.5 mm/deg or even less. Hence, the values of long-range inhibition adopted in our model may be adequate to describe peripheral vision, but are probably too high for central vision close to the fovea. However, it is important to remember that it is extremely difficult to evince functional connections from anatomical data only, since some connections may involve multiple synapses. In particular, pyramidal neurons may provide both direct excitatory input, and indirect inhibitory effects through a disynaptic circuit involving inhibitory interneurons (i.e. basket cells (McGuire, Gilbert, Rivlin, & Wiesel, 1991)). Although only excitatory populations can form long-range connections (6–8 mm) to other networks, both excitatory and inhibitory neurons can be the targets of
For the sake of computational simplicity, in our model a possible reason can be found in the particular disposition of lateral inhibition (see bottom panels in Fig. 13), in agreement with previous physiological considerations. In this case, however, the network becomes less robust to noise.

(ii) a second important argument, which may lead to reconsider the extension of lateral inhibition, arises from simulation of more complex figures (such as the real images in Fig. 13), consisting of several close objects or multiple close contours. The use of an excessively wide inhibition may cause the suppression of tiny contours located in the proximity of a strong contour (as in the case of the mid panels in Fig. 13). A better reproduction of images containing multiple contours may require a reduction in the extension of lateral inhibition (see bottom panels in Fig. 13), in agreement with previous physiological considerations. In this case, however, the network becomes less robust to noise.

Finally, it may be interesting to comment on some limitations and possible future improvements of the present model.

A puzzling aspect of our model concerns the large extension of inhibition and the trade off between noise suppression and co-existence of multiple contours. This apparent puzzle might derive from the use of a Gaussian pattern to describe the decrease of synaptic strength with distance. Of course, it is possible that synapses follow a different rule. For instance, we might have a strong constant inhibition close to the cell (in a range as short as 1°, according to anatomical data) to eliminate noise, and inhibition might suddenly falls to zero at greater distance (to avoid interference with other contours). This may allow noise elimination close to a contour, still avoiding interference between multiple contours. Furthermore, Kang, Shelley, and Sompolinsky (2003), in a recent paper, demonstrated that the visual cortex may operate in a ‘mexican-hat’ manner even using a spatially narrow inhibition, provided inhibition is faster than excitation. Hence, a possibility may be the use of different time constants for inhibition and excitation, with the use of short but faster inhibitory synapses, in greater accordance with anatomical data.

The condition occurring in the presence of multiple contours may also benefit, in future studies, from the use of oscillatory neurons, which synchronize within a contour and de-synchronize in different contours, as proposed in recent models (Li, 1998; Yen & Finkel, 1998).

A further limitation consists in the regular geometrical disposition of receptive fields. Looking at the patterns of feed-forward inputs (see the middle upper panels in Figs. 5–7 and 9) we can observe that the feed-forward mechanism emphasizes especially segments in vertical and horizontal directions as well in directions oriented by ± 45°. A possible reason can be found in the particular disposition of the Gabor receptive fields in contiguous hypercolumns. For the sake of computational simplicity, in our model receptive fields were centered on a 50 × 50 regular lattice. As a consequence of this particular geometrical arrangement, two contiguous receptive fields with orientation 0, 90 or 45° exhibit the greater superimposition, hence more easily concur to the formation of a smooth segment. We expect that the feed-forward input can be improved by using a random disposition of the receptive fields, which breaks any geometrical regularity. However, random disposition would require a more complex numerical procedure, and might be attempted in future studies.

Another limitation is that, in the model, we included only V1 simple cells which respond to elongated local stimuli of a particular orientation. It is well known that, besides these ‘simple cells’, the primary visual cortex also incorporates ‘complex cells’ which respond, for instance, to square corners or which indicate the termination of a segment (end-stopped cells) [see Hubel (1995)]. It is probable that inclusion of these cells may improve contour detection in the model. In fact, as it is evident in Figs. 5, 7, 11 and 12, an inherent difficulty for the model is represented by sharp corners, where a smooth contour changes abruptly. This is especially evident by increasing the strength or width of excitatory synapses. Inclusion of complex cells might lead to an improvement of model behavior, allowing the use of stronger excitatory synapses still avoiding breaking at the corners. As discussed above, this aspect may also be involved with the formation of illusory contours. It is worth noting that the difficulty to close contours at corners has also been underlined by Pettet et al. (1998). Using their model, these authors observed that just two sudden changes in orientation in a closed-path contour reduce detection performance.

Similarly, in the model we included only ON cells stimulated by elongated light bars, and which exhibit an even receptive field. The use of OFF cells, which should respond to elongated dark bars, and/or the use of odd receptive fields [as those described in Jones and Palmer (1987a,b)] may further improve contour extraction. For instance, the simultaneous presence of ON and OFF cells should detect both the internal and external portions of an object boundary, thus improving contour insensitivity to noise.

Another aspect, which may deserve attention in future studies, is the role of texture and color. In the present work, only information on luminance was used to distinguish the internal and external portions of an object, although corrupted by much noise. Recently, Mareschal and Baker (1998), via experimental measurements on cats, observed that the visual cortex is able to detect object boundaries not only by using information on luminance (first-order stimuli) but also information on texture (second-order stimuli).

A further possible extension of the model concerns the interaction with higher visual centers. An example of such interaction is reported in Li (1998). It is possible that some parameters of the model may be controlled by high level activity (for instance, spatial attention mechanisms...
(Motter, 1993) or previous knowledge about the contour shape); such a selection from higher centers may help to strengthen only those portions of a contour which require accurate detection, thus contributing to suppression of unfocused zones. All these aspects are certainly of great importance in the analysis of real vision problems, but are beyond the aim of the present study, which was only devoted to an analysis of local contextual influences in V1.

References


Schmidt, K. E., Goebel, R., Lowsel, S., & Singer, W. (1997). The perceptual grouping criterion of colinearity is reflected by anisotropies of
connections in the primary visual cortex. *European Journal of Neuroscience*, 9, 1083–1089.


