A recurrent model of contour integration in primary visual cortex

Thorsten Hansen

Department of Psychology,
Justus-Liebig-University Giessen, Giessen, Germany

Heiko Neumann

Department of Neural Information Processing,
Ulm University, Ulm, Germany

Physiological and psychophysical studies have demonstrated the importance of colinearity in visual processing. Motivated by these empirical findings we present a novel computational model of recurrent long-range processing in the primary visual cortex. Unlike other models we restrict the long-range interaction to cells of parallel orientation with colinear aligned receptive fields. We also employ a recurrent interaction using modulatory feedback, in accordance with empirical findings. Self-normalizing shunting equations guarantee the saturation of activities after a few recurrent cycles. The primary computational goal of the model is to evaluate local, often noisy orientation measurements within a more global context and to selectively enhance coherent activity by excitatory, modulating feedback. All model simulations were done with the same set of parameters. We show that the model qualitatively reproduces empirical data of response facilitation and suppression for a single bar element depending on the local surround outside the classical receptive field (M. K. Kapadia, M. Ito, C. D. Gilbert, & G. Westheimer, 1995). Next we evaluate the model performance for the processing of artificial and natural images. We quantitatively evaluate the model using two measures of contour saliency and orientation significance. We show that both measures monotonically increase during the recurrent interaction and saturate after a small number of recurrent cycles. The model clarifies how basic tasks of early vision can be accomplished within a single, biologically plausible architecture.

Keywords: contour grouping, primary visual cortex, recurrent processing, long-range interaction, neural model, perceptual organization, shape and contour

Citation: Hansen, T., & Neumann, H. (2008). A recurrent model of contour integration in primary visual cortex. *Journal of Vision, 8*(8):8, 1–25, http://journalofvision.org/8/8/8/, doi:10.1167/8.8.8.

Introduction

The response properties of neurons in the early visual stages are characterized by their receptive field (RF) properties (Hubel & Wiesel, 1962). The classical RF is shaped by the feedforward connections of afferent fibers. More recent studies have shown that the response properties of cortical neurons are substantially influenced by stimuli outside their classical RFs (e.g., Gilbert, 1992; Gilbert & Wiesel, 1990; Kapadia, Ito, Gilbert, & Westheimer, 1995; Lamme, 1995). These contextual influences are mediated by lateral horizontal interactions and feedback connections (Gilbert, Das, Ito, Kapadia, & Westheimer, 1996). The precise functional role of both horizontal and feedback connections is unclear. It has been hypothesized that the more global integration of information from distant areas neurophysiologically "highlights" perceptually salient image features (Lamme & Spekreijse, 2000). The contextual influences seem to be related to low- and mid-level tasks such as contour and feature linking (Allman, Miezin, & McGuinness, 1985; Grossberg & Mingolla, 1985b; Li, 1998) or preattentive texture segmentation (Li, 2000b) as well as higher-level processes like perceptual organization, attention, and visual awareness (Lamme & Spekreijse, 2000).

In this article we focus on the role of recurrent long-range processing for the enhancement of contours. We present a model of contour integration using intralaminar long-range horizontal connections in V1. We employ a novel connection pattern where the horizontal connections are confined to parallel, near colinear orientations. Unlike cocircular filters that have been employed previously, connections between parallel, near colinear orientations have been found in vivo (Bosking, Zhang, Schofield, & Fitzpatrick, 1997; Schmidt, Goebel, Löwel, & Singer, 1997). One goal of the present paper is to investigate whether this connection pattern found in vivo can enhance contours in noisy artificial and natural images. The core of the model is a recurrent loop where top-down activity generated by the long-range pattern is fed back to a previous stage. Unlike previous models, our model uses modulatory feedback that cannot generate activity at the lower stage, but can only modulate the activity already present, in agreement with physiological findings (Hirsch & Gilbert, 1991; Hupé et al., 1998; Salin & Bullier, 1995). Again, our goal is to investigate whether such modulatory feedback is sufficient to enhance contours. The overall goal

doi: 10.1167/8.8.8 Received September 19, 2007; published June 18, 2008 ISSN 1534-7362

is to define and evaluate a complete model for contour enhancement that is based on known anatomical and physiological findings in the early visual cortex. We believe that such an algorithm is a valuable step from a purely conceptual model based on some intuition that some mechanisms should work as expected toward a rigorous, quantitative, and computational understanding of perceptual organization in primary visual cortex. The recurrent interaction that is the essential ingredient of our model is nontrivial and a full mathematical stability analysis for arbitrary two-dimensional input patterns has not been presented. When an analytical solution is missing a neural model is of particular value, since it allows us to investigate the recurrent interactions for a large variety of input stimuli.

The robust and reliable extraction of contours is an important task in early visual processing. However, initial contrast measurements which define the first processing stage in the computation of contour signals are often noisy and fragmented. Therefore, the salient or prominent contours have to be determined out of an array of noisy, cluttered contrast responses.

How is this accomplished? We suggest a computational framework involving *long-range connections, modulating feedback,* and *recurrent interactions*. The task of contour extraction cannot be solved solely on the basis on the incoming data (the feedforward excitatory input), but requires additional constraints and making assumptions on the shape of salient contours. The feedforward input is insufficient to define a contour, because the initial measurements are fragmented and noisy. Thus, additional knowledge about the shape a salient contour needs to be incorporated in the visual system. A main property of salient contours is that they are smooth or colinear, as reflected in the Gestalt law of good continuation (Wertheimer, 1923). It has been suggested that the neural implementation for the law of good continuation may be found in the horizontal *long-range connections* that exist in the superficial layers of early visual areas V1 and V2 (Schmidt et al., 1997). The assumptions or a priori information (such as expressed in the law of good continuation) have to be sufficiently supported by the incoming data to prevent the generation of arbitrary contour responses. We suggest that *modulatory feedback* between layers in V1 plays a central role in this matching process by selectively enhancing only those feedforward input signals that are consistent with the assumptions. The interaction between feedforward data and feedback assumptions requires certain time steps. In each step the result of the interactions is recursively fed into the same matching process. Such a process of *recurrent interaction* might be used by the brain to determine the most stable and consistent representation depending on both the assumptions and the given input data. Recent physiological findings have found a close correlation between the responses of neurons in V1 and the perceptual saliency of contours (Li, Piëch, & Gilbert, 2006). These results show that V1 plays a pivotal role in contour integration.

Motivated by empirical findings we present a model of recurrent long-range interaction in the primary visual cortex for contour processing. The model enhances initially weak oriented activity along a contour that fits into a global context, while suppressing spurious noisy activity. At the same time, activities for multiple orientations at corners and junctions are preserved. With the same parameter settings, the model is employed to simulate physiological data on contour grouping (Kapadia et al., 1995). Here, the model successfully accounts for the increase in activity by colinear flankers, and the surround suppression by noisy textures. We next applied the model for the processing of natural images, and found that contours were successfully enhanced. An extensive numerical evaluation of the temporal evaluation using two measures of contour saliency and orientation significance shows that the model robustly enhances contours for a wide range of input images. All model simulations were carried out with the same set of parameters.

Methods

In this section we describe the model and present the measures of contour saliency for the numerical evaluation of the model.

Model definition

We propose a biologically plausible model of contour integration using intralaminar long-range horizontal connections and interlaminar recurrent interactions in the primary visual cortex. The model connections create a recurrent network which transforms the feedforward input to a stable point where contours are more salient compared to noise.

Motivation of the model components and the model architecture

The model has several key components. The model incorporates (i) localized receptive fields for oriented contrast processing, (ii) cooperative horizontal long-range integration, (iii) inhibitory short-range connections, and (iv) feedforward and feedback processing. These components are motivated by empirical findings.

Oriented contrast processing

Simple and complex cells in the primary visual cortex respond best to a respond best to oriented edges and bars defined by a luminance contrast (Hubel & Wiesel, 1968).

Horizontal long-range connections

Contour integration of aligned edges requires a mechanism that links cells of proper orientation over larger

distances. Horizontal long-range connections found in the superficial layers of V1 may provide such a mechanism: They span large distances (Gilbert & Wiesel, 1983; Rockland & Lund, 1983) and selectively link cells with similar feature preference (Gilbert & Wiesel, 1989; Stettler, Das, Bennett, & Gilbert, 2002; Ts'o, Gilbert, & Wiesel, 1986) and colinearly aligned RFs (Bosking et al., 1997; Schmidt et al., 1997). Evidence for nonlocal integration also comes from psychophysical experiments for contrast detection (Kapadia et al., 1995; Kapadia, Westheimer, & Gilbert, 2000; Polat & Sagi, 1993, 1994) and contour integration (Field, Hayes, & Hess, 1993; Yen & Finkel, 1998). Facilitation is generally strongest for colinear flankers and rapidly declines with the angular difference between target and flankers. Here we focus on these strong facilitatory interactions between iso-oriented items.

Inhibitory short-range connections

Short-range connections are rather unspecific for a particular orientation (Amir, Harel, & Malach, 1993; Bosking et al., 1997; DeAngelis, Freeman, & Ohzawa, 1994) and most likely belong to an inhibitory system (Kisvárday, Kim, Eysel, & Bonhoeffer, 1994).

Modulating feedback

Several physiological studies indicate that feedback projections have a modulation or gating rather than generating effect on cell activities (Hirsch & Gilbert, 1991; Hupé et al., 1998; Salin & Bullier, 1995). Feedback alone is not sufficient to drive cell responses, i.e., initial bottom-up activity is necessary to generate activity (Sandell & Schiller, 1982).

The framework builds upon previous work by Grossberg and colleagues (Grossberg & Mingolla, 1985b; Grossberg & Raizada, 2000), and shares basic computational components such as divisive normalization by nonlinear so-called "shunting" inhibition, recurrent interactions and nonlocal long-range integration.

The core model architecture that we propose here consists of three main stages:

1. an initial preprocessing of the input and a recurrent processing within the two following stages,
2. a combination stage of modulatory feedback and feedforward input, and
3. a cooperative-competitive stage of center-surround long-range interaction (Figure 1).

The key component of this architecture is the recurrent processing at the stages 2) and 3) (Hansen, Sepp, & Neumann, 2001; Neumann & Sepp, 1999). In this recurrent scheme of two interacting regions, let them be cortical layers or areas, each region has a distinct purpose. The lower region serves as a stage of feature measurement and signal detection. The higher region represents expectations about visual structural entities and context information to be matched against the incoming data carried by the feedforward pathway. This architecture has been successfully applied to different domains, such as the disambiguation of local motion (Bayerl & Neumann, 2004), the detection of texture boundaries (Thielscher, Kölle, Neumann, Spitzer, & Grön, 2008; Thielscher & Neumann, 2003, 2005, 2007), the detection of junctions (Hansen & Neumann, 2004a) or to model feature attention (Bayerl & Neumann, 2007). Here we employ the same core architecture for a neural model of contour integration by inter-laminar recurrent interactions within V1. Horizontal long-range interactions are modeled by a spatial weighting function where the interaction is confined to parallel edge elements (i.e., of the same orientation as the target cell) at colinear or near-colinear spatial locations. Overall, the model implements a simplified architecture of V1 (Gilbert, 1993). In the following we shall present the components of the model in more detail.

Feedforward preprocessing

In the feedforward path, the initial luminance distribution is processed by isotropic LGN cells, followed by

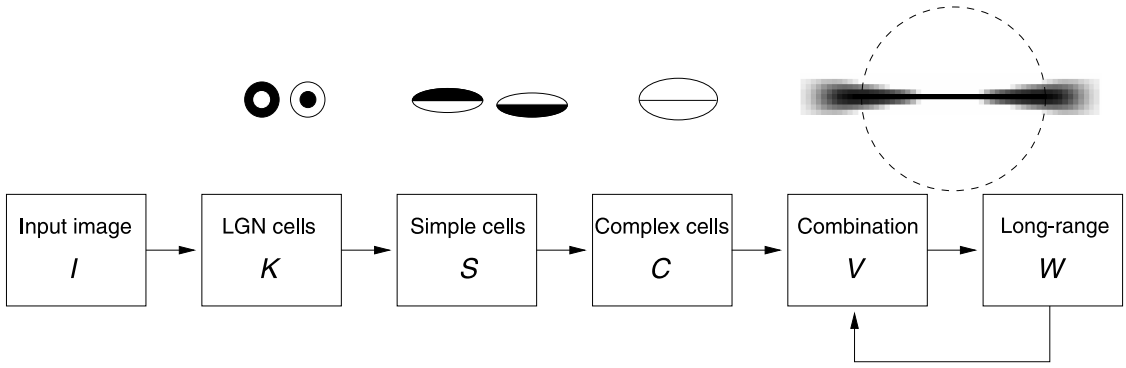

Figure 1. Overview of the model stages. Icons above each stage sketch the receptive field of cells with an orientation of 0 deg. For the long-range stage, the spatial weighting function of the bipole filter is shown together with the spatial extent of the inhibitory short-range interactions (dashed).

orientation-selective simple and complex cells. The interactions in the feedforward path are governed by basic linear equations to keep the processing in the feedforward path relatively simple and to focus on the contribution of the recurrent interaction. A more elaborated processing in the feedforward path would make use of, e.g., nonlinear processing at the level of LGN cells and simple cells (Hansen & Neumann, 2004b; Neumann, Pessoa, & Hansen, 1999). In our model, complex cell responses as output of the feedforward path (cf. Figure 1) provide an initial local estimate of contour strength, position and orientation which is used as bottom-up input for the recurrent loop.

LGN on- and off-cells

Retinal ganglion cells and cells in the LGN have receptive fields with a circular center-surround organization (Kuffler, 1953) that can be modeled by a difference-of-Gaussians (DoG) operator (Enroth-Cugell & Robson, 1966). Responses of isotropic LGN-cells are modeled by the correlation of the initial input stimulus I with values in the range [0, 1] by a DoG operator. Two types of LGN cells are modeled, namely on and off, which generate rectified output responses $K_{\text{on/off}}$

$$\begin{aligned} K &= \text{DoG}_{\sigma_c,\sigma_s} * I \\ K_{\text{on}} &= [K]^+ \\ K_{\text{off}} &= [-K]^+. \end{aligned} \tag{1}$$

Here and in the following * denotes the spatial correlation operator and $[x]^+ := \max\{x, 0\}$ denotes half-wave-rectification. The DoG operator is parameterized by the standard deviation of the center and surround Gaussian ($\sigma_c = 1$, $\sigma_s = 3$), respectively.

The chosen ratio of the size of the center and surround $\sigma_s/\sigma_c = 3$ is larger than the ratio 1.6 that would approximate a Laplacian (Marr & Hildreth, 1980) and smaller than the ratio of about 5 found in vivo (Enroth-Cugell & Robson, 1966) and the ratio of 6 that has been used in a recent investigation of the center-surround properties of retinal ganglion cell RFs (Graham, Chandler, & Field, 2006). We choose a ratio of 3 that allows a robust computation of contrast while at the same time keeping the size of the DoG operator reasonably small. This reduces the computation time of the initial stages and allows us to focus on the contributions of nonlocal cortical interactions.

Simple cells

Simple cells in V1 have elongated subfields (on and off) which sample the input of appropriately aligned LGN responses. Input sampling is modeled by correlation with rotated, anisotropic Gaussians. The Gaussians are shifted perpendicularly to their main axis by $\tau = \pm 3$ to model left and right subfields of an odd-symmetric simple cell. Thus, e.g., for the on-channel, the equations read

$$\begin{aligned} R_{\text{on,left},\theta} &= K_{\text{on}} * G_{\sigma_x,\sigma_y,0,-\tau,\theta} \\ R_{\text{on,right},\theta} &= K_{\text{on}} * G_{\sigma_x,\sigma_y,0,\tau,\theta}. \end{aligned} \tag{2}$$

The activations of the off-channel are computed analogously.

Simple cells are modeled for two polarities (dark–light and light–dark) in $O_{\max} = 4$ orientations ($\theta = 1, \pi/O_{\max}, \ldots, (O_{\max} - 1)\pi/O_{\max}$). The standard deviations of the anisotropic Gaussians are set to $\sigma_y = 1$, $\sigma_x = 3\sigma_y$. For each orientation, the simple cell activity is computed by pooling the two subfield responses. The equation for light–dark (ld) and dark–light (dl) simple cells read

$$\begin{aligned} S_{\text{ld},\theta} &= R_{\text{on,left},\theta} + R_{\text{off,right},\theta} \\ S_{\text{dl},\theta} &= R_{\text{off,left},\theta} + R_{\text{on,right},\theta}. \end{aligned} \tag{3}$$

We employed separate equations for the on and off subfields here because this makes the connection to a proposed a model variant of early feedback clearer (Model variant using early feedback section). In this variant of the model we replace Equation 2 by a feedback-controlled inhibition from cells of opposite contrast polarity.

Complex cells

Cortical complex cells are polarity insensitive. Their response is generated by pooling simple cells of opposite polarities. Before pooling, simple cells of opposite polarities compete and are spatially blurred. The corresponding equations read

$$\begin{aligned} \widetilde{S}_{\text{ld},\theta} &= \left[\left(S_{\text{ld},\theta} - S_{\text{dl},\theta}\right) * G_{\sigma_x,\sigma_y,0,0,\theta}\right]^+ \\ \widetilde{S}_{\text{dl},\theta} &= \left[\left(S_{\text{dl},\theta} - S_{\text{ld},\theta}\right) * G_{\sigma_x,\sigma_y,0,0,\theta}\right]^+ \\ C_\theta &= \widetilde{S}_{\text{ld},\theta} + \widetilde{S}_{\text{dl},\theta}. \end{aligned} \tag{4}$$

The final complex cell responses represent oriented contrast energy similar to the energy models proposed to measure motion energy (Adelson & Bergen, 1985). The present model is a simplified version of the simple cell model we have proposed for visual contrast measurement based on segregated on and off pathways (Neumann et al., 1999).

Recurrent long-range interaction

The output of the feedforward preprocessing defines the input to the recurrent loop which has two stages, namely a combination stage where bottom-up and top-down inputs are fused, and a stage of long-range interaction.

Combination stage

At the combination stage, feedforward complex cell responses C_θ and feedback long-range responses W_θ are added and subject to "shunting interaction," i.e., a nonlinear compression of high amplitude activity following the Weber–Fechner law (Fechner, 1889; Weber, 1846):

$$\partial_t V_\theta = -\alpha_V V_\theta + (\beta_V - V_\theta)\mathrm{net}_\theta, \quad \text{where} \tag{5}$$

$$\mathrm{net}_\theta = C_\theta + \delta_V W_\theta. \tag{6}$$

Solving the equation at equilibrium $\partial_t V_\theta = 0$ results in a normalization of activity

$$V_\theta = \beta_V \frac{\mathrm{net}_\theta}{\alpha_V + \mathrm{net}_\theta}. \tag{7}$$

The weighting parameter $\delta_V = 2$ is chosen so that dimensions of C_θ and W_θ are approximately equal, the decay parameter $\alpha_V = 0.2$ is chosen small compared to net_θ, and $\beta_V = 10$ scales the activity to be sufficiently large for the subsequent long-range interaction. For the first iteration step, feedback responses W_θ are set to C_θ.

Long-range interaction

At the long-range stage, the contextual influences on cell responses are modeled. Orientation-specific, anisotropic long-range connections provide the excitatory input. The inhibitory input is given by isotropic interactions in both the spatial and orientational domain. Long-range connections are modeled by a filter whose spatial layout is similar to the bipole filter as first proposed by (Grossberg & Mingolla, 1985a). The spatial weighting function of the long-range filter is narrowly tuned to the preferred orientation, reflecting the highly significant anisotropies of long-range fibers in visual cortex (Bosking et al., 1997; Schmidt et al., 1997). The size of the long-range filter is about four times the size of the RF of a complex cell.

Essentially, excitatory input is provided by correlation of the feedforward input with the long-range filter B_θ. A cross-orientation inhibition prevents the integration of cell responses at positions where responses for the orthogonal orientation also exist. The excitatory input is governed by

$$\mathrm{net}_\theta^+ = [V_\theta - V_{\theta_\perp}]^+ * B_\theta, \tag{8}$$

where * denotes spatial correlation and $[x]^+ := \max\{x, 0\}$ denotes half-wave-rectification.

The long-range filter is defined as a polar-separable function

$$B_\theta(\varphi, r) = B_{\mathrm{ang}}(\varphi) B_{\mathrm{rad}}(r). \tag{9}$$

The angular function B_{ang} is maximal for the preferred direction θ and smoothly rolls off in a cosine fashion, being zero for angles deviating more than $\alpha/2$ from the preferred orientation:

$$B_{\mathrm{ang}}(\varphi) = \cos(2\pi/2\alpha(\theta - \varphi)) \quad \text{if} \quad |\varphi - \theta| \le \alpha/2, \quad \text{else } 0. \tag{10}$$

The parameter α that defines the opening angle of the long-range filter is set to 20 deg. The radial function B_{rad} is constant for values smaller than $r_{\max} = 25$ and smoothly decays to zero in a Gaussian fashion for values larger than $r_{\max}$:

$$B_{\mathrm{rad}}(r) = \exp(-r^2/(2\sigma)) \quad \text{if } r > r_{\max}, \quad \text{else } 1. \tag{11}$$

The standard deviation of the Gaussian is set to $\sigma = 3$. The long-range filter is finally normalized such that the filter integrates to one. A plot of the long-range filter for a reference orientation of 0 deg is depicted in Figure 2.

Long-range filters of similar spatial but different orientational layout have been proposed previously (Grossberg & Mingolla, 1985b; Grossberg & Raizada, 2000; Grossberg & Williamson, 2001; Guy & Medioni, 1996; Heitger, von der Heydt, Peterhans, Rosenthaler, & Kübler, 1998; Li, 1998, 1999a, 1999b, 2000a, 2000b; Neumann & Sepp, 1999; Parent & Zucker, 1989; Peterhans & Heitger, 2001; Ross, Grossberg, & Mingolla, 2000). In particular, all these formulations allow interactions between elements of different orientations, such as cocircular orientations. The physiological evidence for these types of connections is sparse. On the contrary, a large number of studies have found maximal interaction between cells of the same orientations (Bosking et al., 1997; Buzás, Eysel, & Kisvárday, 1998; Gilbert & Wiesel, 1989; Malach, Amir, Harel, & Grinvald, 1993; Schmidt et al., 1997; Schmidt & Löwel, 2002; Sincich & Blasdel, 2001; Ts'o et al., 1986; Weliky, Kandler, Fitzpatrick, & Katz, 1995), to name a few. Here we investigate how the long-range facilitation when restricted to the biologically plausible pattern of co-aligned cells of parallel orientation can be used to enhance contours in artificial and natural images.

Responses are not salient if neighboring cells of random orientation show strong responses. Such activity has an inhibitory effect on the target cell (Kapadia

Figure 2. Spatial weighting function for the long-range interaction for a reference orientation of 0 deg.

et al., 1995). This inhibitory effect is modeled by an sampling of activity with isotropic Gaussians from the orientational $\widetilde{g}_{\sigma_o,\theta}$, $\sigma_o = 0.5$ and spatial neighborhood $G_{\sigma_{sur}}$, $\sigma_{sur} = 8$:

$$\text{net}_\theta^- = \text{net}_\theta^+ * \widetilde{g}_{\sigma_o,\theta} * G_{\sigma_{surf}}. \quad (12)$$

The standard deviation of the Gaussian in the spatial domain is set to $\sigma_{sur} = 8$ to model the smaller extend of the inhibitory short-range connections. This parameterization results in an effective spatial extension of about half the size of the excitatory long-range interaction modeled by the long-range filter. The standard deviation in the orientational domain is set to $\sigma_o = 0.5$ to give near-zero input for the orthogonal orientation. The orientational weighting function $\widetilde{g}_{\sigma_o,\theta}$ is implemented by a 1D Gaussian g_σ, discretized on a zero-centered grid of size O_{max}, normalized, and circularly shifted so that the maximum value is at the position corresponding to θ. The spatial profile of the 2D Gaussians weighted by the orientational Gaussian is visualized in Figure 3.

The final activity of the long-range stage results from interactions between the excitatory long-range input net_θ^+ and the inhibitory input net_θ^-. The excitatory long-range input net_θ^+ is gated by the activity V_θ to implement a modulating rather than generating effect of lateral interaction on cell activities (Hirsch & Gilbert, 1991; Hupé et al., 1998). Similar to Equation 7, the inhibition from net_θ^- is divisive. The shunting for the long-range stage reads

$$\partial_t W_\theta = -\alpha_W W_\theta + \beta_W V_\theta (1 + \eta^+ \text{net}_\theta^+) - \eta^- W_\theta \text{net}_\theta^-. \quad (13)$$

The equation is solved at equilibrium, resulting in the following nonlinear, divisive interaction at the long-range stage:

$$W_\theta = \beta_W \frac{V_\theta (1 + \eta^+ \text{net}_\theta^+)}{\alpha_W + \eta^- \text{net}_\theta^-} \quad (14)$$

where $\alpha_W = 0.2$ is the decay parameter and $\eta^+ = 5$, $\eta^- = 2$, and $\beta_W = 0.001$ are scale factors.

The multiplicative contribution of V_θ ensures that long-range connections have a modulating rather than generating effect on cell activities (Hirsch & Gilbert, 1991; Hupé et al., 1998). If the bottom up-input V_θ is zero, no activity will be generated, because the multiplication by $V_\theta = 0$ would annihilate any long-range activity net_θ^+.

The result of the long-range stage is fed back and combined with the feedforward complex cell responses, thus closing the recurrent loop. The divisive shunting interactions governing both the long-range interactions and the combination of feedback and feedforward input ensure a saturation of activities after a few recurrent cycles and result in graded responses within a bounded range of activations (Grossberg, 1973, 1981; Grossberg, Mingolla, & Ross, 1997).

To understand the general behavior let us consider the second term of Equation 13, $\beta_W V_\theta (1 + \eta^+ \text{net}_\theta^+)$. This term denotes the soft-gating of the activity from the long-range integration net_θ^+ by the response of the combination stage V_θ. Stability concerning the boundedness of input and output activation is achieved by the combined effect of this soft-gating mechanism together with the divisive, or shunting, inhibition that is effective by incorporating the third term of Equation 13, $\eta^- W_\theta \text{net}_\theta^-$. In addition, the gating variable V_θ is also bounded by employing a shunting mechanism to achieve mass action for the excitatory term (compare Equation 7).

The model is robust against parameter changes mainly due to the compressive transformation equations. For the combination of responses (Equation 7), however, it is crucial to have activities in both streams of similar order of magnitude. Also the relative RF sizes must not be substantially altered. The current parameter setting results in relative RF sizes of complex cells: isotropic short-range filter: long-range interaction of about 1:2.5:4, assuming a cut-off of the Gaussians at 2σ (or 95% of the total energy).

The model is implemented in Matlab; the mfiles of the model implementation are provided as Supplementary Material.

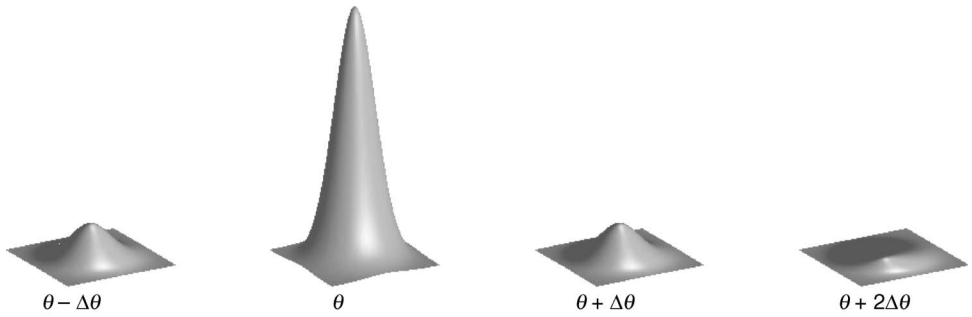

Figure 3. 3D Gaussians in (x, y, θ) space are visualized as $O_{max} = 4$ Gaussians in (x, y) space. Left to right: Gaussians for $\theta - \Delta\theta$, θ, $\theta + \Delta\theta$, and $\theta + 2\Delta\theta$. Note that the Gaussian for the orientation $\theta + 2\Delta\theta$ orthogonal to θ is near zero at all positions.

Model evaluation

Two measures of contour saliency and orientation significance are used to numerically evaluate the competencies of the model.

Contour saliency

To quantify the contour enhancement, we use a saliency measurement as suggested by Li (1999a, 1999b). Li defined the net saliency S at each position as the response of the maximally activated orientation

$$S = \max_{\theta} X_{\theta}. \tag{15}$$

The relative enhancement of contour activity was then defined as the ratio of the mean saliency along the contour and the mean saliency measured over all positions:

$$r = \frac{\overline{S}_{\text{contour}}}{\overline{S}_{\text{all}}}. \tag{16}$$

A second measurement compared the standard deviation of the saliencies at all positions σ_{all} with the difference of the mean saliencies:

$$z = \frac{\overline{S}_{\text{contour}} - \overline{S}_{\text{all}}}{\sigma_{\text{all}}}. \tag{17}$$

A salient contour is characterized by high values of r and z.

Modifications of the saliency definition

We also investigated modifications of the saliency definition as given above. In these modifications, we defined the net saliency as the sum across all orientations at a given position:

$$\tilde{S} = \sum_{\theta} X_{\theta}. \tag{18}$$

Further, we used an alternative formulation for the ratio where the mean saliency along the contour was compared with the mean saliency of the background:

$$\tilde{r} = \frac{\overline{S}_{\text{contour}}}{\overline{S}_{\text{background}}}. \tag{19}$$

Orientation significance

We used a measure of orientation significance to quantify this relative enhancement of contour response across orientations. Orientation significance is a measure of orientation bandwidth that is bounded between zero and one. Cells not tuned for orientations have zero orientation significance. Cells that are very sharply tuned have an orientation significance value close to one. The orientation significance is defined as the length of the vector resulting from summing all orientations normalized by the absolute sum of orientations (Batschelet, 1981; Ringach, Hawken, & Shapley, 1997). The formal definition reads

$$\text{osgnf}(W) = \frac{|\sum_{\theta} W_{\theta} \exp(2i\theta)|}{\sum_{\theta} W_{\theta}.} \tag{20}$$

The factor of 2 in the argument of the exponential function stretches the range of orientations $\theta \in [0, \pi]$ to the full turn $[0, 2\pi]$, such that circular statistics can be applied.

Results

In this section we show the competencies of the model in a number of simulations. We shall begin with an artificial image that allows us to highlight and evaluate various properties of the proposed model. We shall then demonstrate the performance of the model on a number of real world images. The resulting edge images for both the complex cell stage and the long-range stage show pooled responses which are obtained by summing over all orientations.

The values of the model parameters as specified in Model definition section are employed in all simulations. The model response saturates after a few recurrent cycles. Unless denoted otherwise, a number of 12 recurrent cycles is employed. The resulting edge images for both the complex cell stage and the long-range stage show pooled responses which are obtained by summing over all orientations.

Processing of noisy artificial images

In a first simulation a synthetic stimulus of a noisy square is employed. The image is heavily corrupted by high-amplitude additive Gaussian noise of standard deviation equal to 100% of the luminance difference at the edge (so-called 100% Gaussian noise). Figure 4 demonstrates the functionality of lateral long-range interaction for the enhancement of coherent structure. Outline contrasts are detected and subsequently enhanced such that the activities of salient contrast as well as orientation significance is optimized.

The capabilities of the model can be further assessed by a close-up of the top left corner of the square (Figure 5). The simulations demonstrate three important properties of the recurrent long-range interaction:

1. *Contour enhancement:* The weak initial orientation estimates of the square contour are enhanced.
2. *Noise suppression:* Spurious noisy activities in the background are suppressed.

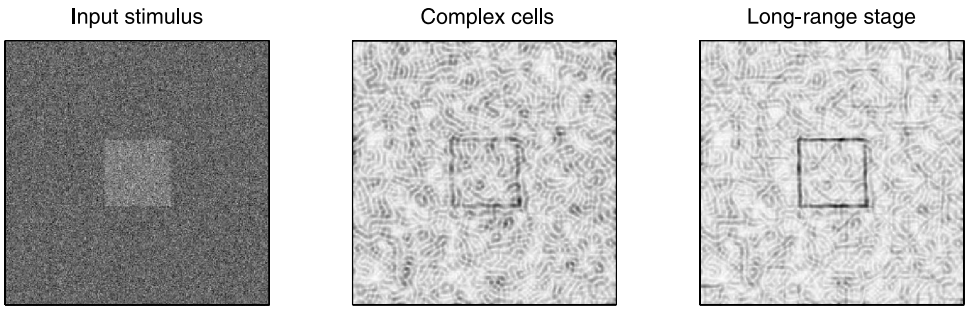

Figure 4. Processing of a square pattern with additive high amplitude noise. The size of the images is 256 × 256 pixels.

3. *Noise suppression:* Spurious noisy activities in the background are suppressed.

The preservation of multiple activities near corners and junctions provide an implicit signature or labeling of such points. These higher order features play a significant role in object recognition and depth segregation (e.g., Biederman, 1985). We have shown that corners and junctions can be extracted with high accuracy from the contour representation resulting from the recurrent interaction (Hansen & Neumann, 2004a).

Overall, the proposed model circuit is thus capable to perform main tasks of low and mid-level vision, namely contour enhancement, noise suppression and junction detection within a single architecture. The model thus defines a computational framework suggesting how these tasks can be accomplished by the neural machinery.

Simulation of empirical data

In order to address the empirical relevance of the model, the model is probed with fragmented contour patterns and texture stimuli such as used in the study of Kapadia et al. (1995). In particular, we investigated the relative contributions of surround inhibition by randomly oriented bars and long-range excitation from colinear flankers on the activity of a central bar element (Kapadia et al., 1995; Knierim & van Essen, 1992). The simulation results are depicted in Figure 6.

We compared the results of the model simulations to results from single cell recordings (Figure 12C in Kapadia et al., 1995). To obtain a qualitative fit, the relative responses of the empirical data were scaled down by a factor of 10. The selectivity of real neurons typically varies over a large range, and the particular neuron we considered from the study by Kapadia et al. (1995) showed a high selectivity. Our model neurons all have the same selectivity which may be close to the average selectivity found in vivo. We cannot test this because data of only a few neurons were shown in Kapadia et al. (1995).

We first simulated the response to a single central bar element which serves as the reference activity. Adding two colinear flanking bars results in a increase of activity. The responses of feedforward complex cells to the flanking bars are integrated by the long-range filter, resulting in higher activation at the long-range stage. In the feedback loop, this higher activity selectively enhances the response to the central bar element. In another experiment, the central bar is embedded into a texture of randomly oriented bars. This results in a response decrease compared to the central bar alone. Here, the short-range inhibitory filter which samples activity summed across all orientations is activated by the randomly oriented bars.

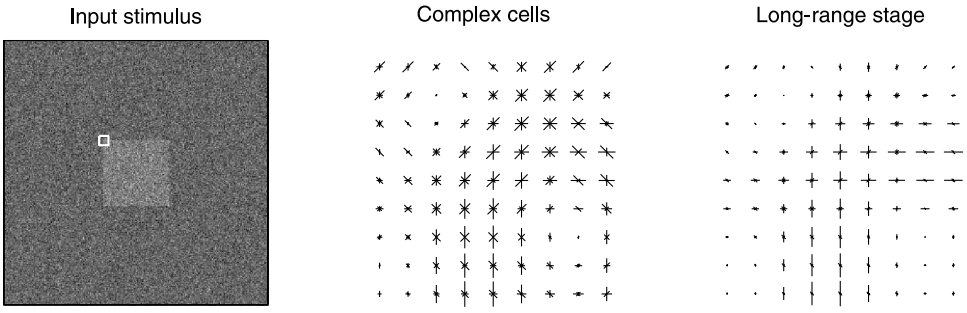

Figure 5. Orientation plot of a close-up of the top left corner of the noisy square. The size of the close-up images is 9 × 9 pixels. Left to right: Input image and close-up of the top left corner (white square inset in the input image) for complex cell responses and long-range responses. In the close-ups, three important properties of the long-range interaction can be seen: i) enhancement of the orientation coaligned to the contour, ii) suppression of noisy activity in the background, and iii) preservation of the significant orientations at corners.

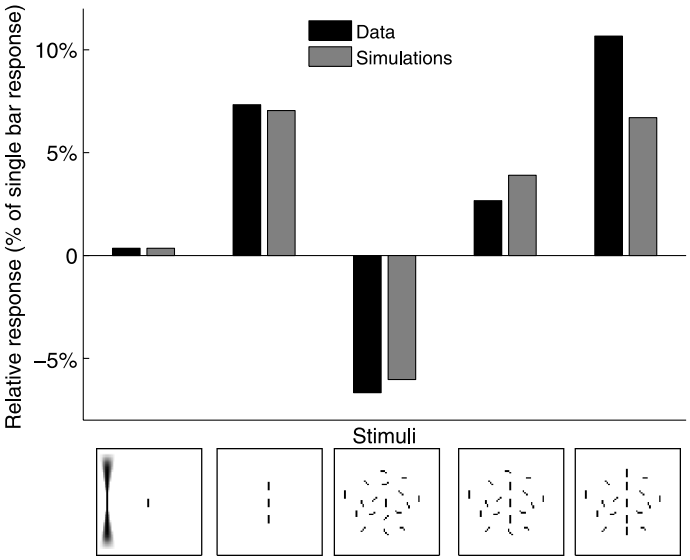

Figure 6. Model responses to generic contour patterns as used in an empirical study by Kapadia et al. (1995). The results depict the percent change of response for different input stimuli relative to the response to a single bar. The relative responses of the empirical data were scaled down by a factor of 10. The left inset in the leftmost stimulus indicates the relative size of the bipole filter. The model responses (gray bars) are in good qualitative agreement with the empirical data (black bars): Adding colinear flankers to a central bar element results in a response increase, adding a texture and randomly oriented bars results in a response decrease. The net effect of combined flankers and texture is a response increase, which strength depends on the strength of supporting activity from collinear flanking bars.

We finally simulated the response to the combined pattern of colinear flankers and a texture of randomly oriented bars. Here, the contributions of the surround inhibition and long-range excitation result in a net effect of excitatory feedback, which is weaker than the response without the textured surround. Adding two more colinear flankers further increases the excitatory feedback and causes a higher activity of the central bar.

Kapadia et al. (1995) found cells where the neuronal response to five colinear bars embedded in a texture was even larger than the response to three coaligned bars without the texture. The simulated data shows deviates from this pattern and the model response do not show such an increase. The empirically observed increase could be due to feedback from higher cortical stages such as V2, which acts on a larger scale and was not incorporated in the present model. A further increase in the model may be achieved by integrating feedback from boundary cells in V2 (Neumann & Sepp, 1999).

Overall, the model exhibits basic response characteristics of long-range excitation and surround suppression in good qualitative agreement with the empirical data.

Processing of natural images

To further examine the model, natural stimuli are used as input. We first employed a cell image depicted in Figure 7. The results show that the outlines of the cells are enhanced by the recurrent long-range interaction. In particular, week initial estimates and low contrast measurements are enhanced, e.g., the longer colinear structures at the top border of the rightmost cell or smaller salient structures like the nuclei. This demonstrates how the proposed circuit can accentuate meaningful structures like object boundaries and may thus serve as a prerequisite of figure-ground segregation.

We have also processed an image of a 3D laboratory scene (Figure 8). Those locations with high contrast complex cell responses and low orientation uncertainty are further stabilized by the recurrent loop. In addition, weak initial estimates such responses along the contour of the pedestal are enhanced, while weak spurious responses to the ground are suppressed. Note that no artifacts are introduced by the long-range interaction, and that, e.g., L-junctions are not turned into crosses. This is due to the modulatory feedback which ensures that no illusory activity can be generated at locations where no initial bottom-up activity exists.

In a final set of simulations we used images from a collection of fruit and vegetable images (Williams, 1998).

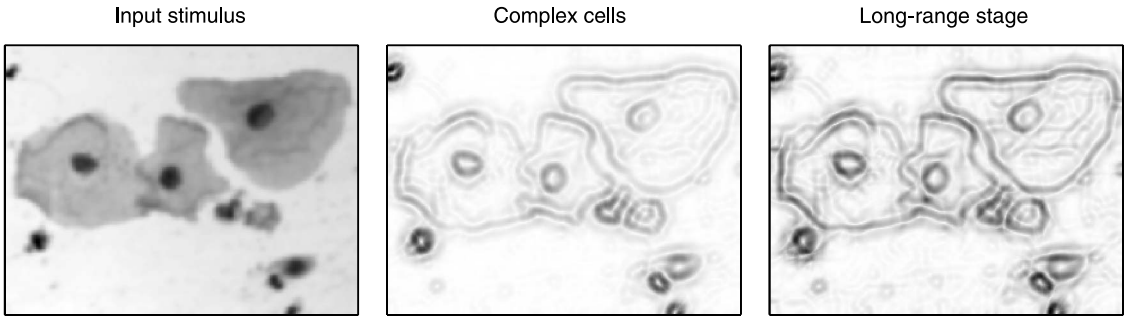

Figure 7. Model responses for a cell stimulus (256 × 195 pixels). Week responses to the cell border are enhanced during the recurrent long-range interaction.

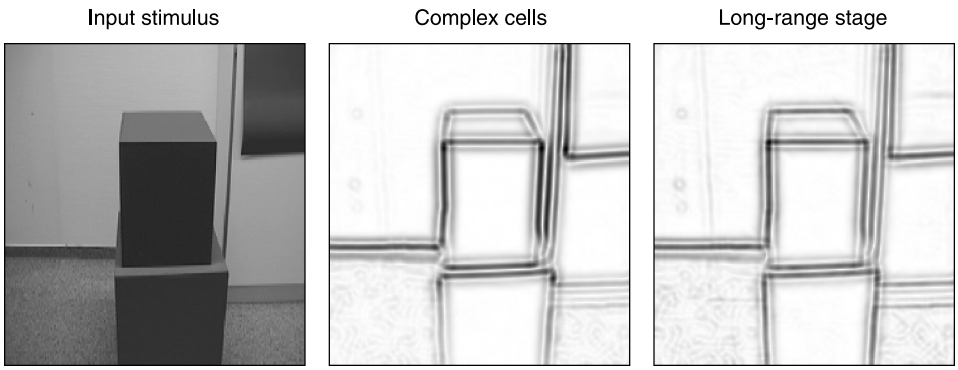

Figure 8. Model responses for a laboratory scene (230 × 246 pixels). The responses to the pedestal and the top of the cube are enhanced during the recurrent long-range interaction.

For the banana image, a close-up of the orientation distribution is shown for the complex cell responses and the long-range responses (Figure 9). Though the image is rather simple because the banana is shown against a uniform background, the initial estimates at the complex cell stage are still noisy and wrong orientation responses occur. These responses are suppressed by the long-range interaction, and only the salient responses along the contour remains. The result is a purified representation of only those contours that are coherent within a broader context.

In a last simulation on natural images we used an image of a sweet potato. Simulations results for the potato image are shown in Figure 10.

The initial complex cells responses to this stimulus already provide a rather good representation of the contours, but show considerable amplitude differences. These differences are compensated by the long-range processing, yielding a more equal contour activity. To demonstrate this property, each image is thresholded at 40% of its relative maximum value (Figure 10, bottom row). Thresholding of the long-range images results in longer continuous contour segments. Some gaps still remain in the thresholded contour, at those locations where the initial response amplitude is considerably lower compared to other parts of the contour. The model responses assume a range of values instead of providing only a binary "all-or-none" response, in accordance with the idea of "analog sensitivity" (Grossberg et al., 1997). Overall, the selective equalization of amplitude differences may ease the tasks of subsequent processing stages of, e.g., figure-ground determination and object recognition.

The simulations of natural images demonstrate core competencies of the model. Initial complex cell activations generated for localized high contrast contours are further stabilized. Initially weak activations in coherent spatial arrangements are enhanced. The results show that noisy low contrast arrangements can be significantly enhanced to form elementary items of smooth contour segments.

Quantitative evaluation and analysis

In the following, we shall define two measurements of *contour saliency and orientation significance* which allow for a quantitative, more rigorous assessment of the capabilities of the model.

One core property of the proposed circuit is to enhance initially weak and noisy responses along contours. This property is quantitatively evaluated using a measure of contour saliency. The results show that the model successfully enhances the saliency of contours in noisy images. The measure of contour saliency compares responses at different locations (i.e., at the contour and at other locations). A robust contour processing scheme however should also suppress conflicting responses at the same location, such that only the valid response in the contour directions remains. To

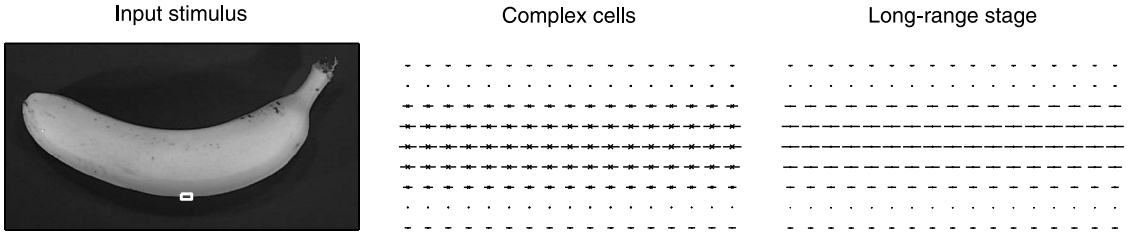

Figure 9. Close-up of the model responses for a banana image. The close-up has a size of 17 × 9 pixels and is shown as a small white square inset in the input image. Only those orientation responses which are coaligned along the contour are enhanced by the recurrent long-range processing relative to other spurious responses.

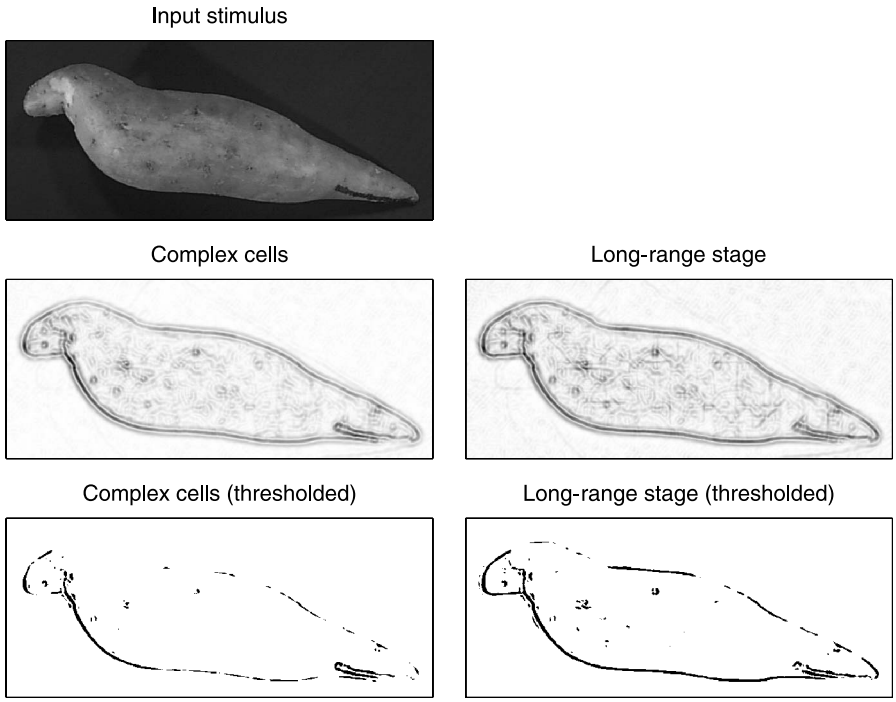

Figure 10. Model responses for a sweet potato image (610 × 256). The binary images were thresholded at 40% of the maximum value.

quantitatively address this property, a complementary measure of orientation significance is introduced.

We evaluated the temporal evolution of orientation significance for a variety of different contrast and noise values for a number of different realizations of the noise process. In all cases the orientation significance at the contour is considerable increased. At the background, on the contrary, the orientation significance remains almost the same. Next we analyzed the orientation significance for a synthetic orientation distribution to develop some insight into the range of significance values assumed. In particular we show that for a prototypic synthetic orientation distribution the significance values at the background cannot decrease below a certain limit.

Contour saliency

One core property of the proposed circuit is contour enhancement: The activity at contour locations should be increased compared to background locations. This property has already been demonstrated in the initial simulation of the noise square (Figure 4) and becomes even more prominent in a close-up of the top border of the square (Figure 11). Besides the strengthening of the contour activity, the results also demonstrate that the model circuit can close gaps (as can be seen in the right part of the contour), as long as some amount of initial bottom-up activity is present.

To quantify the contour enhancement, we used a saliency measurement as suggested by Li (1999a, 1999b). Li employed a measure r to quantify the ratio of the mean saliency along the contour and the mean saliency across the whole stimulus, and second measure z that compares the standard deviation of the saliencies at all positions with the difference of the mean saliencies, as detailed in Contour saliency section. A salient contour is characterized by high values of r and z.

Li used contour saliency to determine the saliency of boundaries in synthetic textures consisting of bars of different orientations. Here we use the saliency to quantify the enhancement in the processing of noisy artificial stimuli. The saliency values we obtained are of the same order of magnitude as those computed by Li. Because of differences in the input stimuli, the results are not directly comparable.

Temporal evolution of contour saliency

In this section we evaluate the contour enhancement property of the proposed model in terms of the saliency measure. In particular, we are interested in the time course of both measures during the recurrent interaction: Does the saliency of the contour increases, as expected? Is the increase monotonically, and does the network converge to stable point, or does it oscillate? Does the network dynamics saturate, and if so, after how many recurrent interactions?

A plot of the temporal evolution of both saliency values (r, z) for the noisy square stimulus is given in Figure 12. For the contour values, we considered positions defined by stripes of a width of two pixels along each of the four sides of the square. Mean saliency values at these positions are then compared against mean saliency values measured over all pixels of the whole input stimulus. The curves show an increase in saliency for

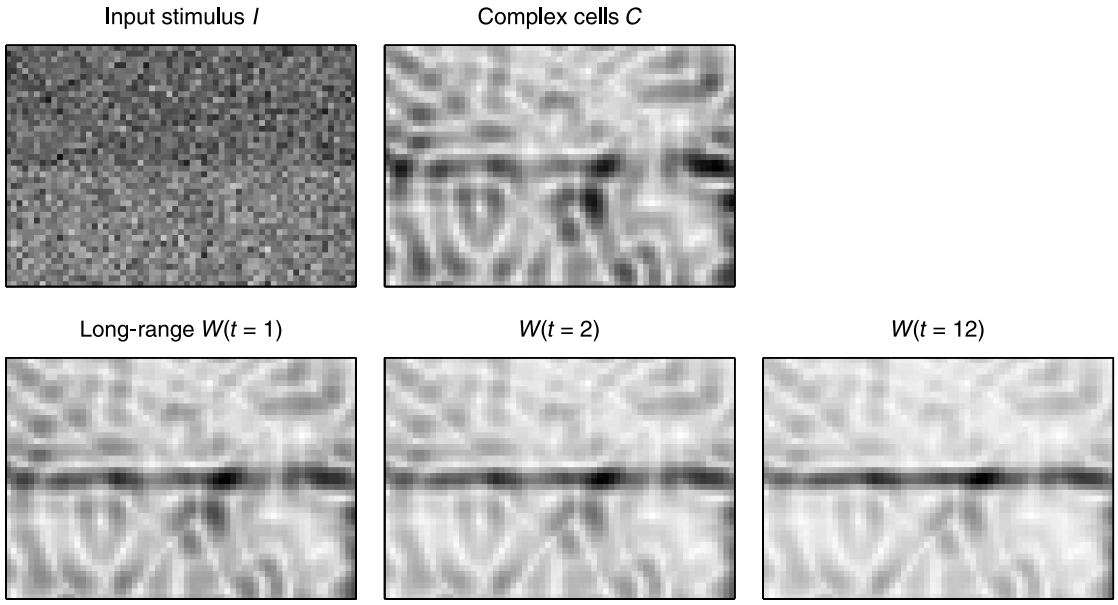

Figure 11. Close-up of the processing results obtained for the top contour of the noisy square. The size of the close-up images is 61 × 42 pixels. Top row, left to right: Input image and the initial complex cell responses. Bottom row, left to right: Result of the recurrent long-range processing at three discrete time steps (*t* = 1, 2,…, 12).

both values during the recurrent processing: initial saliency values of (*r, z*) = (2.3, 2.9) as obtained at the complex cell level are increased to (*r, z*) = (5.7, 7.0) after 12 recurrent cycles.

Regarding our initial questions, the simulation results show i) that the model successfully increases the contour saliency of an initially fragmented contour of low saliency, ii) that this increase is monotonic and saturates after about 12 recurrent cycles, and iii) that the largest increase occurs within the first recurrent interaction.

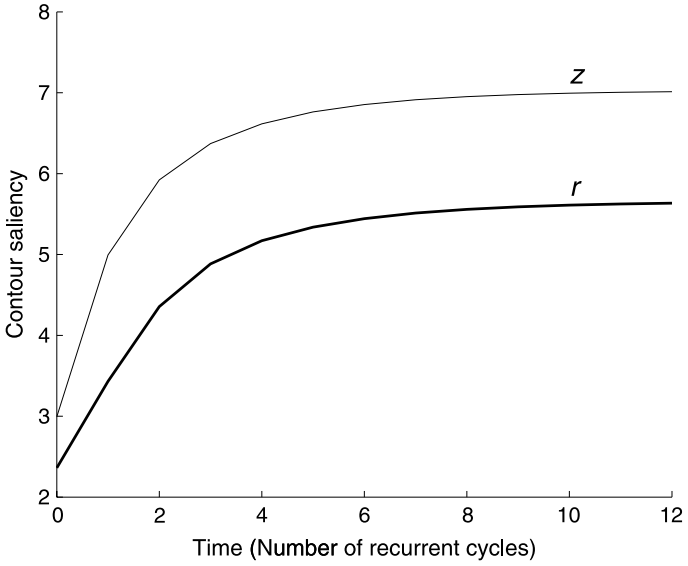

Figure 12. Temporal evolution of contour saliency for the noisy square. Top solid curve denotes the temporal evolution of the *z* value, the bottom bold curve denotes the *r* value. The initial values (*r, z*) = (2.3, 2.9) monotonically increase during the recurrent interaction and saturate at about 12 recurrent cycles at values (*r, z*) = (5.7, 7.0).

The saliency measure is based on the comparison of values along the contour against all the values in the image, including the contour. We investigated a modification of the saliency measure where we compared disjunctive data sets, i.e., the values along the contour with the values of the background (Equation 19). We have also investigated a modification of the net saliency measure based on the sum across all orientation instead of the maximum orientation (Equation 18). The results of these modifications of the saliency measures were virtually indistinguishable from the original measure such that we kept the original definition.

Having established contour saliency as a suitable measurement of the model competencies we now employ this measure to investigate the effect of parameter variations on the behavior of the model.

Effects of the extent of the long-range interactions on contour saliency

One central parameter of the model is the extent of the long-range interactions. In this section we investigate how of a smaller size of the long-range filter effects contour saliency. For a fixed scale of the preprocessing stages we modified the parameters of the long-range interaction. In particular, we varied the effective radius $r_{\max}$ of the long-range filter, and the standard deviation σ_{sur} of the Gaussian used for the short-range inhibition.

Recall that the parameters are set to $r_{\max} = 25$ and $\sigma_{\text{sur}} = 8$ in the original model as specified in Model definition

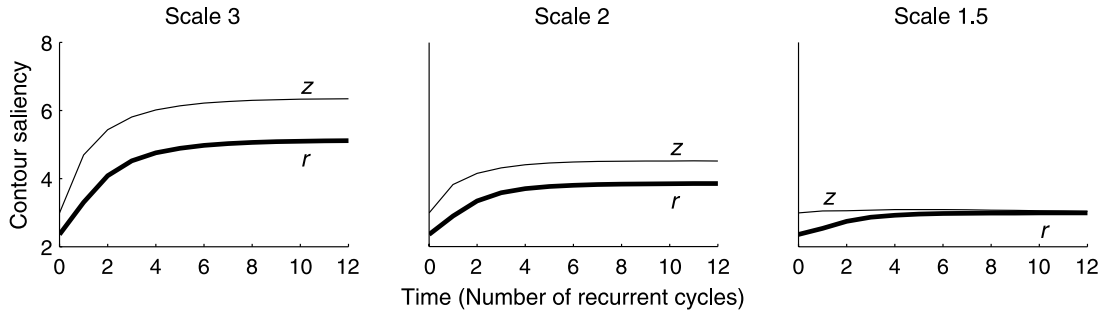

Figure 13. Temporal evolution of contour saliency for the noisy square for different scales of the long-range interactions. The top solid curve denotes the temporal evolution of the *z* value, the bottom bold curve denotes the *r* value. The increase of the initial values (*r, z*) = (2.3, 2.9) is considerably smaller if the range of the long-range interaction decreases: the final significance values are (*r, z*) = (5.1, 6.4) for scale 3, (*r, z*) = (3.9, 4.5) for scale 2, and (*r, z*) = (3.0, 3.0) for scale 1.5, as compared to (*r, z*) = (5.7, 7.0) for the original parameter setting (see Figure 12).

section. This setting results in a relative RF size of complex cells: isotropic short-range filter: long-range interaction of about 1:2.5:4. For the scale variations, we simulated three different scenarios, with decreasing relative RF size of the long-range filter, namely 3, 2, and 1.5 times the filter size used at the complex cell level. The relative size of the short range filter is kept fixed at approximately 2.5/4 = 0.625 the size of the long-range filter. The parameter settings for the three scenarios are as follows:

1. scale 3, relative RF sizes 1:1.875:3, (r_{max} = 19, σ_{sur} = 6)
2. scale 2, relative RF sizes 1:1.25:2, (r_{max} = 13, σ_{sur} = 4)
3. scale 1, relative RF sizes 1:1:1.5, (r_{max} = 9, σ_{sur} = 3)

The resulting temporal evolution of the saliency values are depicted in Figure 13. The plots show that the saliency drops to considerably smaller values if the range of interactions at the long-range stage is decreased, validating the relative scale of the long-range interaction and the inhibitory short-range interaction that was chosen in the definition of the model.

Orientation significance

The saliency measure as introduced above allowed us to compare the response at different positions, but does not allowed us to quantify the strength of a particular response relative to other responses of different orientations at the same position. For this purpose we used a measure of orientation significance that allows us to quantify the relative enhancement of contour response across orientations. Orientation significance, as defined in Orientation significance section is a measure of orientation bandwidth which is bounded between zero and one. Cells not tuned for orientations have a zero orientation significance. Cells that are sharply tuned have an orientation significance close to one.

Temporal evolution of orientation significance

In a pilot study, we examined the changes of orientation significance during recurrent long-range interaction for the image of the noisy square. Similar to the investigations for the measure of contour saliency, we were interested how the orientation significance changes during the time course of the recurrent long-range interaction. Does the orientation significance also increases monotonically and then saturates, similar to behavior found for the measure of contour saliency?

We therefore investigated the temporal evolution of the mean orientation significance for two patches of size of 2 × 40 pixels. Results are depicted in Figure 14. One patch is placed at the contrast boundary (upper solid line) and the other at the background (lower solid line). The curves for the two patches show that the recurrent interaction increases orientation significance only at the borders while leaving the significance at the background almost unchanged.

Orientation significance of the border increases monotonically and saturates at about 12 recurrent cycles, similar to the behavior found for the contour saliency. This shows that the model not only increases the activity at the contour (as shown by the measure of contour saliency), but that in addition this increase is highly selective and only increases the orientation that is tangent to the object contour.

Analysis of the domain of the orientation significance function

The numerical evaluation in the previous section has shown that the orientation significance differs considerable between the contour and the background. However, one may wonder about the values taken by the orientation significance function: Why do the values at the border level off at about 0.8 instead of approaching 1, and why, on the other hand, does the background significance assumes values about 0.33 instead of lower values near

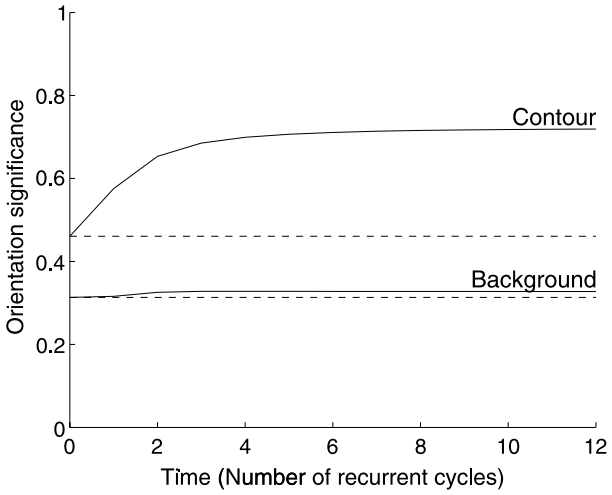

Figure 14. Temporal evolution of mean orientation significance for a patch placed at the contrast boundary (upper solid line), and a second patch placed at the background (lower solid line). The dashed horizontal lines indicate the corresponding mean initial significance obtained for the complex cell responses. The abscissa denotes the discrete time steps (t = 0, 1, …, 12), the ordinate denotes mean orientation significance. The curves show a pronounced increase in orientation significance at the borders (increasing from 0.46 to 0.72), while the significance at the background remains almost unchanged (increasing only slightly from 0.31 to 0.33). This corresponds to an increase in the ratio of the orientation significance at the border and at the background from 1.47 to 2.20 during the recurrent processing.

zero? Further, one may intuitively presume a decrease of orientation significance at background locations. The purpose of the following analysis is to address these points and to develop some intuition about the values assumed by the orientation significance function.

We consider an idealized distribution of orientation responses that has a response $w_{opt} > 1$ at the preferred orientation, zero response at the orthogonal orientation and unit-valued residual responses at all other orientations. The function that describes the values of orientation significance depending on the value of w_{opt} is derived in Orientation significance section for a synthetic distribution in Appendix A. A sample plot of the orientation significance values for $O_{max} = 4$ orientations and for varying strengths of the optimal orientation $w_{opt} = 1, 2, \ldots, 10$ is depicted in Figure 15. The minimal and maximal values of the optimal orientation ($w_{opt} = 1$ and $w_{opt} = 10$) result in orientation significances of 0.33 and 0.83, respectively. These values are in good agreement with the orientation significance at the background and the final orientation significance at the contour as measured in the simulation of the noisy square stimulus (compare Figure 14). Increasing the number of orientations results in a shift of the orientation significance curve toward zero and in a slightly steeper slope of the curve. For $O_{max} = 8$, e.g., the significance values for $w_{opt} = 1$ and $w_{opt} = 10$ are given by 0.14 and 0.63, respectively.

The above analysis has clarified that the orientation significance cannot decrease below 0.33 (given a zero response at the orthogonal orientation), and that an orientation significance of about 0.8 results from a tenfold higher response at the preferred orientation compared to the residual responses. Consequently, values significantly below 0.33 can only arise for equal responses along all orientations. However, the local image structure in noisy artificial and natural images almost always has a small orientation bias that results in higher responses for this orientation and vanishing or near zero responses at the orthogonal orientation. A pronounced decrease of orientation significance during long-range processing would require an increase or even generation of activity orthogonal to the local image structure, thus violating the idea of modulating feedback.

Orientation significance for different contrast and noise levels

So far, we have shown the competency of the model to increase orientation significance only for a single noisy image, i.e., for a single contrast and noise level and a particular realization of the noise process. For a more complete assessment of the competencies of the model, the temporal evolution of mean orientation significance is evaluated for a square image using a larger variation of the input parameters (contrast and noise levels) and for a number of different realizations of each noise level. The term "contrast" here denotes the contrast amplitude, i.e., the difference between maximal and minimal luminance values in the input image. We have simulated a square of 0.1 contrast, corrupted with 10%, 20%, 50%, and 100%

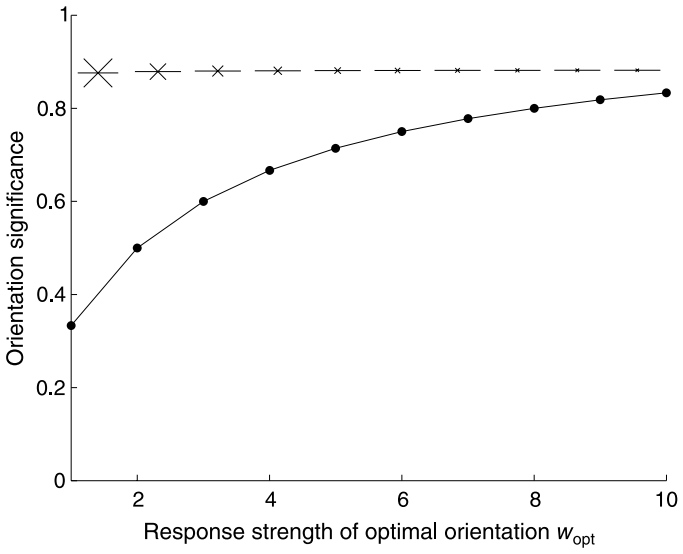

Figure 15. Evaluation of orientation significance for a synthetic orientation response distribution. The abscissa denotes the response strength of the preferred orientation, $w_{opt} = 1, 2, \ldots, 10$, the ordinate denotes the corresponding orientation significance. The 10 orientation plots inset at the top depict the corresponding orientation distribution.

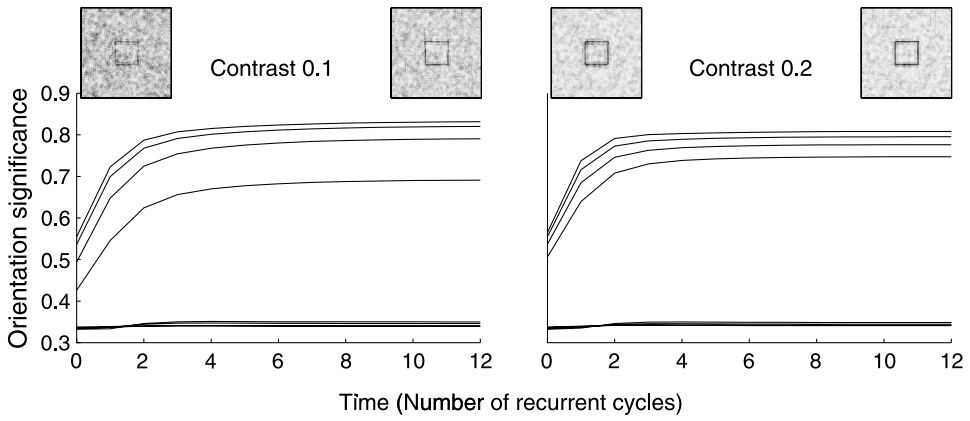

Figure 16. Temporal evolution of mean orientation significance under variation of the input contrast and noise level. Results are averaged over 100 different realizations of the noise process. Left: Contrast of 0.1, corrupted with 10%, 20%, 50%, and 100% Gaussian noise. Right: Contrast of 0.2, corrupted with 5%, 10%, 20%, and 40% Gaussian noise. The top four curves in each plot show the mean orientation significance along the contours of the square for increasing noise levels (top to bottom). The bottom curves show the corresponding orientation significance at the background. The insets show the initial complex cell response and the final long-range response for the highest noise level. The results show an enhancement of orientation significance along the contour, while at the background the orientation significance remains almost the same.

Gaussian noise (Figure 16, left) and a square of higher contrast of 0.2, corrupted with 5%, 10%, 20%, and 40% Gaussian noise (Figure 16, right). The different and smaller noise values for the high-contrast stimulus are chosen to guarantee input luminance values in the range [0, 1]. The simulation results are averaged over 100 different realizations of the noise process to exclude effects resulting from a particular realization. The results of this extensive numerical evaluation are shown in Figure 16.

The results show that the model interactions increase the orientation significance along contours for a large variety of contrast and noise levels. This is important for the robust processing of natural images, where the contrast along a contour is not constant and noise of different sources and strength may occur.

Model variant using early feedback

In the model proposed above, the feedback signal is combined with the feedforward signal at a relatively high level, namely with the feedforward signal as generated by the complex cells (see Figure 1). This circuit models the intralaminar recurrent interaction of layer 2/3 pyramidal cells. Physiological and anatomical studies indicate the existence of another interlaminar recurrent cycle in V1. This recurrent loop runs from layer 4 to layer 2/3 and feeds back via layer 5 and 6 to layer 4 (Bolz, Gilbert, & Wiesel, 1989), where the feedback signal is combined with layer 4 simple cells. In this section we introduce a modification of the model using early feedback, where the feedback signal also influences simple cell responses. The

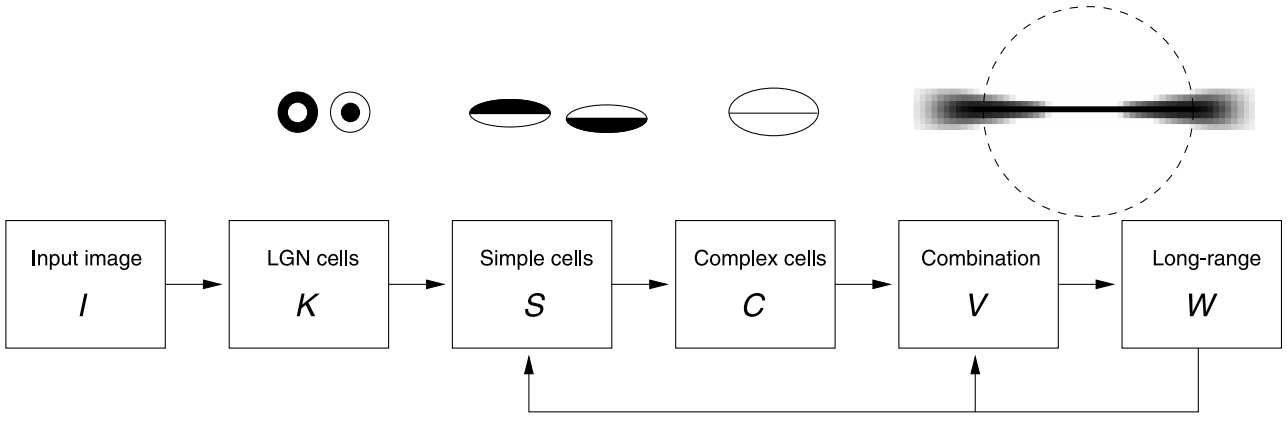

Figure 17. Overview of model stages of the new model using early feedback. Instead of terminating only at the combination stage (cf. Figure 1), the feedback signal now feeds also into the simple cell stage. Icons above each stage sketch the receptive field of cells with an orientation of 0 deg. For the long-range stage, the spatial weighting function of the bipole filter is shown together with the spatial extent of the inhibitory short-range interactions (dashed).

core model architecture using early feedback is depicted in Figure 17.

So far we have only motivated the new termination of the feedback connection within the simple cell layer. How should this combination be specified? We have proposed a mechanism of dominating opponent inhibition (DOI) that reduces the noise sensitivity of simple cells (Hansen & Neumann, 2004b). DOI introduces an adaptive threshold controlled by the activity in the opponent path. This adaptive threshold suppresses activity at the background, but not at locations of high contrast such as edges. In the simple cell model with DOI, the contrast locations are only determined locally in a single feedforward sweep. The DOI mechanism could be made even more robust by integrating more global contour information such as provided by the long-range interactions. The idea is to control the amount of DOI by long-range feedback signals, such that DOI occurs only at the background but not at contour locations signaled by high orientation significance. These ideas are made rigorous in the following section, were the precise equations of the model variant with early feedback are provided.

Modification of the model equations

We modify the model equations as introduced in Model definition section at the simple cell stage to introduce early feedback.

Simple cells

The simple cells combine input from different subfields. In the new model, the subfield responses R are generated by a feedback controlled DOI signal Ξ. Equation 2 are replaced by

$$\begin{aligned} R_{\text{on,left},\theta} &= \left[(K_{\text{on}} - \Xi \cdot K_{\text{off}}) * G_{\sigma_x,\sigma_y,0,-\tau,\theta}\right]^+ \\ R_{\text{on,right},\theta} &= \left[(K_{\text{on}} - \Xi \cdot K_{\text{off}}) * G_{\sigma_x,\sigma_y,0,\tau,\theta}\right]^+ \end{aligned} \tag{21}$$

where the feedback control is defined as

$$\Xi = \xi - \text{osgnf}(W). \tag{22}$$

The DOI parameter ξ is set to 2 (Hansen & Neumann, 2004b). In Equation 22, high orientation significance along contours leads to a decrease of inhibition of the opponent channel. Low significance values of osgnf(W) close to zero results in high dominating inhibition with $\Xi \approx \xi$. In the first iteration, W is set to zero. The activations of the off-channel are computed analogously. The old model results from setting $\Xi = 0$.

The use of the orientation significance to control the DOI signal Ξ is motivated mainly computationally. The orientation significance has three useful features: it robustly distinguishes between oriented and non-oriented structures, it is bound between [0, 1], and it grows during long-range processing. From a biological point of view, the computation of orientation significance requires only local computations within a hypercolumn. While the use of orientation significance for feedback control has not been shown empirically and hence is speculative, such a signal could principally be computed by the neural machinery.

Simulation results using early feedback

In this section we show simulation results for the model with early feedback. We start with a simulation of the noisy square image used in the simulations with the standard model. The simulation results are depicted in Figure 18. The recurrent processing with the new model generated a slightly more pronounced contour activity (Figure 18, top row as compared to simulation results for the standard model in Figure 4), while the orientation plot of the top right corner appears to be nearly the same (Figure 18, bottom row as compared to simulation results for the standard model in Figure 5).

For a rigorous comparison of the two models, more qualitative measurements are needed. Such measurements have been defined above as contour saliency and orientation significance.

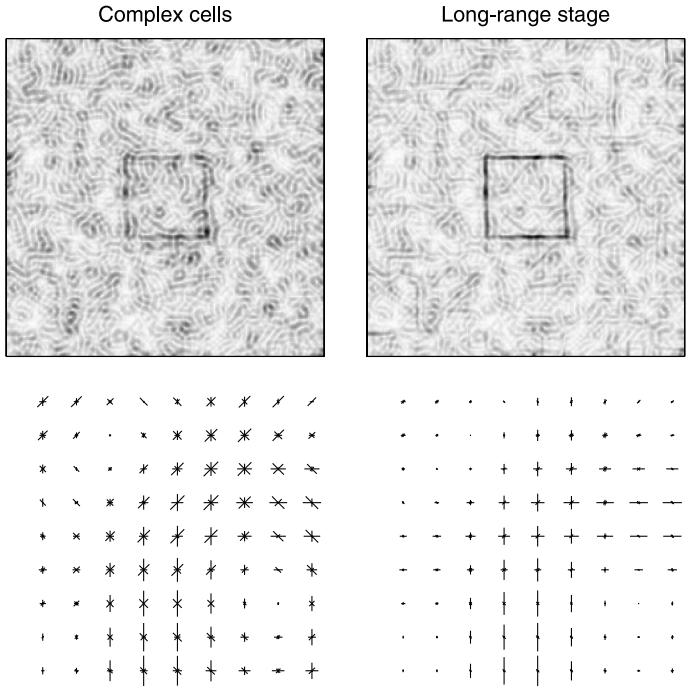

Figure 18. Processing of a square pattern with additive high amplitude noise by the model with early feedback. The size of the images is 256 × 256 pixels. Top row, left to right: Initial complex cell responses and the result of the recurrent long-range processing. Bottom row, left to right: Orientation plot of a close-up of the top right corner of the noisy square for the complex cells and the long-range stage. The size of the close-up images is 9 × 9 pixels.

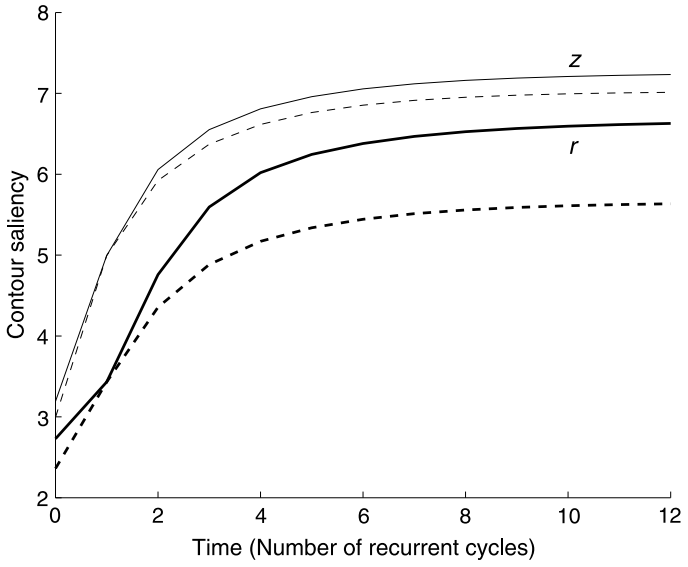

Figure 19. Temporal evolution of contour saliency for the noisy square generated by the model with early feedback. Top solid curve denotes the temporal evolution of the *z* value, the bottom bold curve denotes the *r* value. The dashed lines indicate the respective curves of the standard model. The saliency values monotonically increase during the recurrent interaction and level off at about 12 recurrent cycles at values (*r, z*) = (6.7, 7.3) as compared to (*r, z*) = (5.7, 7.0) for the standard model. The increase of the *z* value is nearly the same for both models (4.0 for the standard model and 4.1 for the new model), but the increase of the *r* value is considerably stronger for the model variant with early feedback (3.3 compared to 4.0).

In a first experiment, we compare the contour saliency of the two models Figure 19. The results show a stronger increase in saliency for the new model, especially in the relative enhancement as expressed by the *r* value.

We have also redone the extensive simulations measuring the temporal evolution of orientation significance for a variety of input parameters (noise and contrast) and different realizations of each respective noise level (Figure 20). The results show a slightly better performance of the old model. In particular, the orientation significance of background locations is increased less for the old model. Due to the opponent inhibition at the simple cell level, the new model has a sharper orientation tuning, resulting in a slightly higher initial orientation significance at the background (about 3.6 as compared to 3.3 in Figure 16, left) as well as a slightly higher final significance value (about 0.37 as compared to 3.4 in Figure 16, left).

The results can be interpreted in terms of complementary roles of the two kinds of feedback as employed in the model variants: while early feedback results in a more salient contour response, late feedback results in higher orientation significance. The results may indicate a functional role for the different kinds of feedback loops as observed in vivo.

Discussion

Summary of findings

We proposed a model circuit of V1 contour processing utilizing long-range interactions and recurrent processing. The model realizes a biologically plausible algorithm of

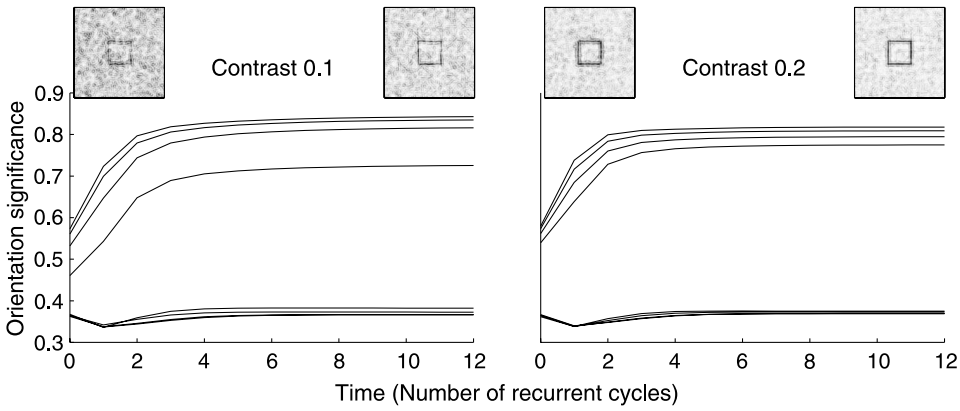

Figure 20. Temporal evolution of mean orientation significance under variation of the input contrast and noise level for the new model with early feedback. Results are averaged over 100 different realizations of the noise process. The corresponding results for the standard model are depicted in Figure 16. Left: Contrast of 0.1, corrupted with 10%, 20%, 50%, and 100% Gaussian noise. Right: Contrast of 0.2, corrupted with 5%, 10%, 20%, and 40% Gaussian noise. The top four curves in each plot show the mean orientation significance along the contours of the square for increasing noise levels (top to bottom). The bottom curves show the corresponding orientation significance at the background. The results show an enhancement of orientation significance along the contour, while at the background the orientation significance increases only slightly.

contour integration that allows us to retrieve the contours in a scene. We evaluated the model using a variety of artificial and natural images. We showed that the model enhances initially weak oriented activity along a contour that fits into a global context, while suppressing spurious noisy activity. The increase in activation is paralleled by an increase of orientation selectivity at contour locations. At the same time, responses to multiple orientations at corners and junctions are preserved.

We also evaluated the model response to contour fragments as used in an physiological study by Kapadia et al. (1995). The model results are in good agreement with the single-cell recordings: embedding a single line in a texture of randomly oriented lines results in a decrease of response due to surround suppression. Adding colinear flankers results in a response increase. Finally, we developed a variant of the model that incorporates early feedback; this model shows an even greater performance on contour salience.

Numerical evaluation

To numerically evaluate the model we used a measure of contour saliency (Li, 1999a, 1999b) and a second measure of orientation significance (Batschelet, 1981; Ringach et al., 1997). We show that both measures increase monotonically during the recurrent interaction and saturate after about 12 recurrent cycles.

Another approach to evaluate the model would be to first build a generative model of contours (for instance a model where contours are chains of Gabor-like edges and more likely to be cocircular) and then optimize the model by maximizing the efficiency of detecting the contours. This would allow us to write more general selectivity measures and would translate all choices of the model components and architecture to a framework of maximizing the efficiency of an algorithm. One problem with this approach is that no conclusive and generally agreed upon generative model of contours exists. Thus a number of arbitrary a priori choices would be necessary to define the generic properties of a contour (such that contours are chains of Gabor-like edges and more likely to be cocircular).

Comparison to other models

A number of different models have been proposed for contour integration. For a review, see Neumann and Mingolla (2001). Among the first approaches that utilize recurrent processing for contour extraction is the Boundary Contour System (Grossberg & Mingolla, 1985a, 1985b). A slightly revised version of the original Boundary Contour System serves as the basic building block for a model of recurrent *intracortical* contour processing between areas V1 and V2 (Ross et al., 2000). Our model focuses on the *intralaminar* processing within V1. Grossberg and coworkers suggest that V1 and V2 circuits are homologous and differ only in the size of their constituent neurons' receptive fields: they propose that V2 is basically V1 at larger scale. In contrast, we propose that V1 and V2 have different functional roles, such that, e.g., cells responding to illusory contours occur in V2 and corner selective cells occur in V1. Our model uses *modulating* feedback: initial bottom-up activity is necessary to generate activity. Consequently, in contrast to the model of Grossberg and coworkers, our model of V1 does not allow for the creation of illusory contours. Illusory contours evoke cell responses in V2 (von der Heydt, Peterhans, & Baumgartner, 1984) and have been investigated in a model of V1–V2 interactions (Neumann & Sepp, 1999). For natural stimuli, small responses due to sensor noise also occur between fragmented contours. We have shown that those small responses can be enhanced by our model of V1, making it suitable for contour formation.

Other models have explicitly used end-stop operators that provide the model input (Finkel & Edelman, 1989; Finkel & Sajda, 1992; Heitger et al., 1998; von der Heydt & Peterhans, 1989), while we used activity from initial contrast measurements that is sharpened by feedback modulation.

An alternative approach to model V1 recurrent interaction has been suggested by Li (1998, 1999a, 1999b, 2000b). The model of Li differs from our model in a number of features. Li uses a two-layer recurrent model of excitatory and inhibitory neurons, which interact by linear dynamics. In our model, nonlinear shunting equations are used to combine inhibitory and excitatory signals. Li also uses two types of highly tuned anisotropic filters for excitatory and inhibitory interactions that gather input from various orientations. Instead, we employed an excitatory long-range filter that collects input from the same orientation only. Further, we used a shorter range inhibitory isotropic Gaussian filter that is not tuned for orientation. These more basic interaction structures are in agreement with empirical findings in vivo. Long-range integration from different orientations based on a circularity constraint such as proposed by a large variety of models (Li, 1998; Parent & Zucker, 1989; Yen & Finkel, 1998) are speculative and have not been found in vivo. More recently, Ben-Shahar and Zucker (2004) have formulated a model of long-range interaction based on differential geometry, showing that texture and shading continuation could serve as complementary functional explanations to contour integration.

Multiscale processing

A number of studies have stressed the importance of multiple spatial scales for the proper extraction of images features (Lindeberg, 1998; Mokhtarian & Suomela, 1998; Würtz & Lourens, 2000). In particular, multiscale models inspired by the functional architecture of V1 have been

proposed for the processing of synthetic aperture radar images (Grossberg, Mingolla, & Williamson, 1995), to infer 3D shape from texture (Grossberg, Kuhlmann, & Mingolla, 2007), to model brightness perception (Grossberg & Hong, 2006; Hong & Grossberg, 2004; Sepp & Neumann, 1999), and to implement a sparse approximation of images (Fischer, Cristóbal, & Redondo, 2006, 2007). The present model operates only on a single scale. However, the model neurons integrate information over successively increasing regions of the visual space in the time course of the recurrent interactions. At the first iteration a target neuron receives information from neighboring neurons with a maximum distance as defined by the extent of the long-range filter. At the second recurrent iteration, the neighboring neuron itself has received information from its neighboring neurons and can convey this information to the target neuron. Thus, information of different scales is available during the temporal evolution of the orientation responses.

Convergence and stability of the recurrent model

Evaluation of the model with various input stimuli showed that the model dynamics saturate after a small number of recurrent cycles (about 12). This is due to the divisive shunting interactions that govern the computations at the long-range stage and the combination stage (compare the discussion at the end of Recurrent long-range interaction section). It has been shown that suitably designed shunting networks can adapt their sensitivity and give a sensitive response within finite intervals even if the input fluctuates in a much broader dynamic range (Grossberg, 1973, 1981; Grossberg et al., 1997).

Formally, absolute stability of global pattern formation has been proven in a classical paper by Cohen and Grossberg (1983) for a broad class of nonlinear systems. Our model equations are largely consistent with this scheme but differ in two aspects. First, we use an excitatory long-range input, i.e., an input that depends on the activity of neighboring neurons (analogous to the inhibitory input). Networks with an excitatory long-range input have been extensively evaluated numerically, suggesting that absolute stability should also exist for these networks (Cohen & Grossberg, 1983). Second, the gating activity of the excitatory input is from a previous stage and described by another nonlinear differential equation. Both extensions make the investigation of stability more complicated.

The present model defines a system of two coupled nonlinear shunting equations. The well-posedness for such a system has been proven under simplifying assumptions, such as negligible lateral interactions (Cardanobile, Cohen, Corchs, Mugnolo, & Neumann, 2008). Several extensions of the Cohen–Grossberg theorem have been investigated (e.g., Chen, Lu, & Chen, 2005; Guo & Huang, 2006; Lu & Chen, 2003), but a formal proof of stability for the network architecture used in our model is still a matter of future mathematical investigations.

Subsymbolic enhancement and symbolic representation of contours

The model we presented enhances contours in a large variety of contexts. The idea of enhancement is that the tasks of subsequent processing stages will be easier if the contours are less noisy. The enhanced contour in our model is still represented as a set of discrete orientation activities at different points in space, like the initial input activity. The model does not attach a common symbolic label that marks points as belonging to the same contour. In this sense, the present model does not group anything. One may argue that surfaces must still be discovered and represented and objects must be identified. The processing within our model is one step toward this ultimate goal. Our model aims at understanding how exactly this step can be realized by the neural circuitry in V1.

The present model could be extended in various ways to derive a more symbolic and explicit representation of a contour. For example, one could propagate labels along the contours. Another approach would be to have several independent and competitive layers each defining a contour (Wersing, Steil, & Ritter, 2001). Yet another approach would be to have the purified contour activity of the long-range stage providing the input to a subsequent stage where a symbolic representation is derived. In previous work we have pursued the latter approach. We have shown that the model output of the recurrent long-range interaction can be used to precisely detect corners and junction points in the image (Hansen & Neumann, 2004a).

One may question the need for an explicit, symbolic representation of contours. While the symbolic extraction of contours is the prevailing assumption and processing strategy pursued in computer vision, it is by no means clear that the visual system has an explicit, symbolic representation of contours and surfaces. There is some consensus about the existence of a level or stage where objects are represented, but the existence of symbolic representations at intermediate stages is more controversial (Peterson & Rhodes, 2003). For example, it has been shown that most cells in V2 and V4 respond well to complex stimuli, indicating that the population of these cells represents complex shape information (Hegdé & Van Essen, 2000, 2003, 2007). However, such a distributed representation is different from an explicit symbolic representation of single contours in the brain. The purpose of vision is to find out what is in the world by seeing, which may not necessarily involve an intermediate representation at which, e.g., a tree is a represented as a set of 12786 contour segments. As Wertheimer (1923) writes at the beginning of his seminal paper (p. 301) “I stand at the window and see a house, trees, sky. […] I see it in this definite common form, in this definite separation.

[…] What a surprise, if I discover after prolonged viewing, after many attempts, in a very unnatural attitude, that over there at the window parts of the dark frame together with a bare branch form the letter N."

Understanding an information processing system

In his influential monograph "Vision" Marr (1982) points out three levels at which an information processing system needs to be understood before one can claim complete understanding:

1. *Computational theory:* What is the computational problem that needs to be solved, i.e., what is the goal of the computation, and what is the underlying logic of the computational procedures?
2. *Representation and algorithm:* How can this computational theory be implemented? In particular, how is the input and output represented, and what is the algorithm of the transformation?
3. *Hardware implementation:* How can the representation and algorithm be realized physically?

Here we presented a neural model for contour integration by interlaminar recurrent interactions in V1. The model addresses the level of representation and algorithm with a consideration of the neuronal implementation. In particular, we used a set of discrete orientation activities (a hypercolumn) at each point in visual space to define the representation of the input and output. The algorithm is based on the integration of matching activity from neighboring neurons, modulatory feedback of this integrated activity to local measurements, and a recurrent interaction of this feedback process. At each stage, divisive shunting normalization ensures that the activity is bounded in a certain range (Grossberg, 1973, 1981; Grossberg et al., 1997).

The value of a purely computational model is to clarify the computational procedures the problem demands. The value of a neural model is to clarify the representation and algorithms by which the computational problem can be solved using principles and processes that can be realized by neuronal hardware. Each kind of model has is particular pitfalls. A computational model may come up with an elegant and efficient solution of the problem, but may be completely ignorant about how the proposed computations could be implemented by the brain. A computational model yields only one out of potentially many possible computational solutions, and the particular solution found may be fundamentally different from the solutions realized—and realizable—by neuronal interactions. A neural model, on the other hand, may result in a grab-bag of quasi-neural structures that cannot be independently motivated. We have tried to circumvent this problem by using a limited set of neural mechanisms and processes, which are well motivated and which each serve a distinct computational purpose (as detailed in Motivation of the model components and the model architecture section). These mechanisms operate within a basic architecture of two interacting regions that has been employed to model visual processes in other domains (Bayerl & Neumann, 2004, 2007; Hansen & Neumann, 2004a; Thielscher & Neumann, 2003).

Summary

We have shown that basic tasks in early vision processing, such as contour enhancement, noise suppression and preservation of multiple activities at junctions can be realized by a neural model based on principles that are consistent with empirical findings of neural processing within V1.

Appendix A

Orientation significance for a synthetic distribution

We consider an idealized distribution of orientation responses that is similar to the typical distribution resulting from the preprocessing scheme. For the purpose of analysis we define an idealized synthetic orientation response distribution W_{synth} having a response $w_{\text{opt}} > 1$ at the preferred orientation, zero response at the orthogonal orientation and unit-valued residual responses at all other orientations:

$$\begin{aligned} W_{\text{synth}} &= w_{\text{opt}} && \text{if } \theta = \theta_{\text{opt}} \\ &= 0 && \text{if } \theta = \theta_{\text{nonopt}} = \theta_{\text{opt}} + \pi/2 \\ &= 1 && \text{else.} \end{aligned} \tag{A1}$$

For an even number of orientations $O_{\max} \bmod 2 = 0$, the responses can be grouped in pairs of mutually orthogonal orientations θ and $\theta_\perp$, whose responses w_θ and $w_{\theta_\perp}$ provide antagonistic contributions to the orientation significance:

$$\begin{aligned} &w_\theta \exp(2i\theta) + w_{\theta_\perp} \exp(2i(\theta + \pi/2)) \\ &\quad = w_\theta \exp(2i\theta) + w_{\theta_\perp} \exp(2i\theta)\exp(\pi/2) \\ &\quad = (w_\theta - w_{\theta_\perp}) \exp(2i\theta). \end{aligned} \tag{A2}$$

The residual responses obey $w_\theta = w_{\theta_\perp} = 1$, such that these responses cancel each other. The only remaining term in the numerator of the significance function is then

given by the pair of optimal and non-optimal responses. The denominator is the sum of all responses. The orientation significance for the synthetic orientation response distribution W_{synth} thus reads

$$\begin{aligned}\text{osgnf}(W_{\text{synth}}) &= \frac{|(w_{\text{opt}} - w_{\text{nonopt}})\exp(2i\theta_{\text{opt}})|}{w_{\text{opt}} + (O_{\max} - 2)} \\ &= \frac{w_{\text{opt}}}{w_{\text{opt}} + O_{\max} - 2}.\end{aligned} \tag{A3}$$

For $O_{\max} = 4$, the value used in all simulations, the orientation significance function is given by

$$\begin{aligned}\text{osgnf}(W_{\text{synth}}) &= \text{osgnf}(w_{\text{opt}}) \\ &= \frac{w_{\text{opt}}}{w_{\text{opt}} + O_{\max} - 2} \\ &= \frac{w_{\text{opt}}}{w_{\text{opt}} + 2}.\end{aligned} \tag{A4}$$

Acknowledgments

We thank Bevil R. Conway and Martin Giesel for detailed and insightful comments on the manuscript. The author H. N. acknowledges the support by a grant from the European Union, ICT-project no. 215866-SEARISE.

Commercial relationships: none.
Corresponding author: Thorsten Hansen.
Email: Thorsten.Hansen@psychol.uni-giessen.de.
Address: Justus-Liebig-Universität Giessen, Abteilung Allgemeine Psychologie, Otto-Behaghel-Str. 10F, 35394 Giessen, Germany.

References

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A, Optics and Image Science, 2,* 284–299. [PubMed]

Allman, J., Miezin, F., & McGuinness, E. (1985). Stimulus specific responses from beyond the classical receptive field: Neurophysiological mechanisms for local-global comparisons in visual neurons. *Annual Review of Neuroscience, 8,* 407–430. [PubMed]

Amir, Y., Harel, M., & Malach, R. (1993). Cortical hierarchy reflected in the organization of intrinsic connections in macaque monkey visual cortex. *Journal of Comparative Neurology, 334,* 19–46. [PubMed]

Batschelet, E. (1981). *Circular statistics in biology*. London: Academic Press.

Bayerl, P., & Neumann, H. (2004). Disambiguating visual motion through contextual feedback modulation. *Neural Computation, 16,* 2041–2066. [PubMed]

Bayerl, P., & Neumann, H. (2007). A neural model of feature attention in motion perception. *Biosystems, 89,* 208–215. [PubMed]

Ben-Shahar, O., & Zucker, S. (2004). Geometrical computations explain projection patterns of long-range horizontal connections in visual cortex. *Neural Computation, 16,* 445–476. [PubMed]

Biederman, I. (1985). Human image understanding: Recent research and a theory. *Computer Vision, Graphics, and Image Processing, 32,* 29–73.

Bolz, J., Gilbert, C. D., & Wiesel, T. N. (1989). Pharmacological analysis of cortical circuitry. *Trends in Neurosciences, 12,* 292–296. [PubMed]

Bosking, W. H., Zhang, Y., Schofield, B., & Fitzpatrick, D. (1997). Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *Journal of Neuroscience, 17,* 2112–2127. [PubMed] [Article]

Buzás, P., Eysel, U. T., & Kisvárday, Z. F. (1998). Functional topography of single cortical cells: An intracellular approach combined with optical imaging. *Brain Research, 3,* 199–208. [PubMed]

Cardanobile, S., Cohen, M., Corchs, S., Mugnolo, D., & Neumann, H. (2008). Investigation of input–output gain in dynamical systems for neural information processing. In W. Arendt & W. Schleich (Eds.), *Mathematical analysis of evolution, information and complexity* (pp. 169–184). Weinheim, Germany: Wiley-VCH.

Chen, T., Lu, W., & Chen, G. (2005). Dynamical behaviors of a large class of general delayed neural networks. *Neural Computation, 17,* 949–968.

Cohen, M., & Grossberg, S. (1983). Absolute stability and global pattern formation and parallel memory storage by competitive neural networks. *IEEE Transactions on Systems, Man, and Cybernetics, 13,* 815–826.

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

Enroth-Cugell, C., & Robson, J. G. (1966). The contrast sensitivity of retinal ganglion cells of the cat. *The Journal of Physiology, 187,* 517–552. [PubMed] [Article]

Fechner, G. T. (1889). *Elemente der Psychophysik*. Leipzig, Germany: Breitkopf & Härtel.

Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: Evidence

for a local "association field." *Vision Research, 33,* 173–193. [PubMed]

Finkel, L. H., & Edelman, G. M. (1989). Integration of distributed cortical systems by reentry: A computer simulation of interactive functionally segregated visual areas. *Journal of Neuroscience, 9,* 3188–3208. [PubMed] [Article]

Finkel, L. H., & Sajda, P. (1992). Object discrimination based on depth-from-occlusion. *Neural Computation, 4,* 901–921.

Fischer, S., Cristóbal, G., & Redondo, R. (2006). Sparse overcomplete Gabor wavelet representation based on local competitions. *IEEE Transactions on Image Processing, 15,* 265–272. [PubMed]

Fischer, S., Cristobal, G., & Redondo, R. (2007). Sparse approximation of images inspired from the functional architecture of the primary visual areas. *EURASIP Journal on Advances in Signal Processing, 90727,* 1–16.

Gilbert, C. D. (1992). Horizontal integration and cortical dynamics. *Neuron, 9*, 1–13. [PubMed]

Gilbert, C. D. (1993). Circuitry, architecture, and functional dynamics of visual cortex. *Cerebral Cortex, 3,* 373–386. [PubMed]

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

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

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

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

Graham, D. J., Chandler, D. M., & Field, D. J. (2006). Can the theory of "whitening" explain the center-surround properties of retinal ganglion cell receptive fields? *Vision Research, 46,* 2901–2913. [PubMed] [Article]

Grossberg, S., (1973). Contour enhancement, short term memory, and constancies in reverberating neural networks. *Studies in Applied Mathematics, 52,* 217–257.

Grossberg, S., (1981). Adaptive resonance in development, perception and cognition. In S. Grossberg (Ed.), *Mathematical psychology and psychopysiology*. Providence, RI: American Mathematical Society.

Grossberg, S., & Hong, S. (2006). A neural model of surface perception: Lightness, anchoring, and filling-in. *Spatial Vision, 19,* 263–321. [PubMed]

Grossberg, S., Kuhlmann, L., & Mingolla, E. (2007). A neural model of 3D shape-from-texture: Multiple-scale filtering, boundary grouping, and surface filling-in. *Vision Research, 47,* 634–672. [PubMed]

Grossberg, S., & Mingolla, E. (1985a). Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading. *Psychological Review, 92,* 173–211. [PubMed]

Grossberg, S., & Mingolla, E. (1985b). Neural dynamics of perceptual grouping: Textures, boundaries, and emergent segmentations. *Perception & Psychophysics, 38,* 141–171. [PubMed]

Grossberg, S., Mingolla, E., & Ross, W. D. (1997). Visual brain and visual perception: How does the cortex do perceptual grouping? *Trends in Neurosciences, 20,* 106–111. [PubMed]

Grossberg, S., Mingolla, E., & Williamson, J. R. (1995). Synthetic aperture radar processing by a multiple scale neural system for boundary and surface representation. *Neural Networks, 8,* 1005–1028.

Grossberg, S., & Raizada, R. D. (2000). Contrast-sensitive perceptual grouping and object-based attention in the laminar circuits of primary visual cortex. *Vision Research, 40,* 1413–1432. [PubMed]

Grossberg, S., & Williamson, J. R. (2001). A neural model of how horizontal and interlaminar connections of visual cortex develop into adult circuits that carry out perceptual grouping and learning. *Cerebral Cortex, 11,* 37–58. [PubMed] [Article]

Guo, S., & Huang, L. (2006). Stability analysis of Cohen-Grossberg neural networks. *IEEE Transactions on Neural Networks, 17,* 106–117. [PubMed]

Guy, G., & Medioni, G. (1996). Inferring global perceptual contours from local features. *International Journal of Computer Vision, 20,* 113–133.

Hansen, T., & Neumann, H. (2004a). A simple cell model with dominating opponent inhibition for robust image processing. *Neural Networks, 17,* 647–662. [PubMed]

Hansen, T., & Neumann, H. (2004b). Neural mechanisms for the robust representation of junctions. *Neural Computation, 16,* 1013–1037. [PubMed]

Hansen, T., Sepp, W., & Neumann, H. (2001). Recurrent long-range interactions in early vision. In S. Wermter, J. Austin, & D. Willshaw (Eds.), *Emerging neural architectures based on neuroscience, LNCS/LNAI 2036* (pp. 127–138). Berlin, Germany: Springer.

Hegdé, J., & Van Essen, D. C. (2000). Selectivity for complex shapes in primate visual area V2. *Journal of Neuroscience, 20,* RC61. [PubMed] [Article]

Hegdé, J., & Van Essen, D. C. (2003). Strategies of shape representation in macaque visual area V2. *Visual Neuroscience, 20,* 313–328. [PubMed]

Hegdé, J., & Van Essen, D. C. (2007). A comparative study of shape representation in macaque visual areas V2 and V4. *Cerebral Cortex, 17,* 1100–1116. [PubMed] [Article]

Heitger, F., von der Heydt, R., Peterhans, E., Rosenthaler, L., & Kübler, O. (1998). Simulation of neural contour mechanisms: Representing anomalous contours. *Image and Vision Computing, 16,* 407–421.

Hirsch, J. A., & Gilbert, C. D. (1991). Synaptic physiology of horizontal connections in the cat's visual cortex. *Journal of Neuroscience, 11,* 1800–1809. [PubMed] [Article]

Hong, S., & Grossberg, S. (2004). A neuromorphic model for achromatic and chromatic surface representation of natural images. *Neural Networks, 17,* 787–808. [PubMed]

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology, 160,* 106–154. [PubMed] [Article]

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology, 195,* 215–243. [PubMed] [Article]

Hupé, J. M., James, A. C., Payne, B. R., Lomber, S. G., Girard, P., & Bullier, J. (1998). Cortical feedback improves discrimination between figure and background by V1, V2 and V3 neurons. *Nature, 394,* 784–787. [PubMed]

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

Kapadia, M. K., Westheimer, G., & Gilbert, C. D. (2000). Spatial distribution of contextual interactions in primary visual cortex and in visual perception. *Journal of Neurophysiology, 84,* 2048–2062. [PubMed] [Article]

Kisvárday, Z. F., Kim, D. S., Eysel, U. T., & Bonhoeffer, T. (1994). Relationship between lateral inhibitory connections and the topography of the orientation map in cat visual cortex. *European Journal of Neuroscience, 6,* 1619–1632. [PubMed]

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

Kuffler, S. W. (1953). Discharge patterns and functional organization of mammalian retina. *Journal of Neurophysiology, 16,* 37–68. [PubMed]

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

Lamme, V. A., & Spekreijse, H. (2000). Modulations of primary visual cortex activity representing attentive and conscious scene perception. *Frontiers in Bioscience, 5,* D232–D243. [PubMed]

Li, W., Piëch, V., & Gilbert, C. D. (2006). Contour saliency in primary visual cortex. *Neuron, 50,* 951–962. [PubMed] [Article]

Li, Z. (1998). A neural model of contour integration in the primary visual cortex. *Neural Computation, 10,* 903–940. [PubMed]

Li, Z. (1999a). Contextual influences in V1 as a basis for pop out and asymmetry in visual search. *Proceedings of the National Academy of Sciences of the United States of America, 96,* 10530–10535. [PubMed] [Article]

Li, Z. (1999b). Visual segmentation by contextual influences via intra-cortical interactions in the primary visual cortex. *Network, 10,* 187–212. [PubMed]

Li, Z. (2000a). Can V1 mechanisms account for figure-ground and medial axis effects. In *Advances in neural information processing systems 12 (NIPS'99).* Cambridge, MA: MIT Press.

Li, Z. (2000b). Pre-attentive segmentation in the primary visual cortex. *Spatial Vision, 13,* 25–50. [PubMed]

Lindeberg, T. (1998). Feature detection with automatic scale selection. *International Journal of Computer Vision, 30,* 77–116.

Lu, W., & Chen, T. (2003). New conditions on global stability of Cohen-Grossberg neural networks. *Neural Computation, 15,* 1173–1189. [PubMed]

Malach, R., Amir, Y., Harel, M., & Grinvald, A. (1993). Relationship between intrinsic connections and functional architecture revealed by optical imaging and in vivo targeted biocytin injections in primate striate cortex. *Proceedings of the National Academy of Sciences of the United States of America, 90,* 10469–10473. [PubMed] [Article]

Marr, D. (1982). *Vision.* San Francisco: W. H. Freeman.

Marr, D., & Hildreth, E. (1980). Theory of edge detection. *Proceedings of the Royal Society of London B: Biological Sciences, 207,* 187–217. [PubMed]

Mokhtarian, F., & Suomela, R. (1998). Robust image corner detection through curvature scale space. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 20,* 1376–1381.

Neumann, H., & Mingolla, E. (2001). Computational neural models of spatial integration and perceptual grouping. In T. F. Shipley & P. J. Kellman (Eds.), *From fragments to objects: Segmentation and grouping*

in vision: *Vol. 130. Advances in psychology* (chap. 12, pp. 353–400). Amsterdam, The Netherlands: Elsevier Science.

Neumann, H., Pessoa, L., & Hansen, T. (1999). Interaction of ON and OFF pathways for visual contrast measurement. *Biological Cybernetics, 81,* 515–532. [PubMed]

Neumann, H., & Sepp, W. (1999). Recurrent V1–V2 interaction in early visual boundary processing. *Biological Cybernetics, 81,* 425–444. [PubMed]

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

Peterhans, E., & Heitger, F. (2001). Simulation of neuronal responses defining depth order and contrast polarity at illusory contours in monkey area V2. *Journal of Computational Neuroscience, 10,* 195–211. [PubMed]

Peterson, M. A., & Rhodes, G. (Eds.) (2003). *Perception of faces, objects and scenes: Analytic and holistic processes*. New York: Oxford University Press.

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

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

Ringach, D. L., Hawken, M. J., & Shapley, R. (1997). Dynamics of orientation tuning in macaque primary visual cortex. *Nature, 387,* 281–284. [PubMed]

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

Ross, W. D., Grossberg, S., & Mingolla, E. (2000). Visual cortical mechanisms of perceptual grouping: Interacting layers, networks, columns, maps. *Neural Networks, 13,* 571–588. [PubMed]

Salin, P. A., & Bullier, J. (1995). Corticocortical connections in the visual system: Structure and function. *Physiological Reviews, 75,* 107–154. [PubMed]

Sandell, J. H., & Schiller, P. H. (1982). Effect of cooling area 18 on striate cortex cells in the squirrel monkey. *Journal of Neurophysiology, 48,* 38–48. [PubMed]

Schmidt, K. E., Goebel, R., Löwel, 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. [PubMed]

Schmidt, K. E., & Löwel, S. (2002). Long-range intrinsic connections in cat primary visual cortex. In B. Payne & A. Peters (Eds.), *The cat primary visual cortex* (pp. 387–428). Orlando, FL: Academic Press.

Sepp, W., & Neumann, H. (1999). A multi-resolution filling-in model for brightness perception. In *Proceedings of the ICANN,* Edinburgh, UK.

Sincich, L. C., & Blasdel, G. G. (2001). Oriented axon projections in primary visual cortex of the monkey. *Journal of Neuroscience, 21,* 4416–4426. [PubMed] [Article]

Stettler, D. D., Das, A., Bennett, J., & Gilbert, C. D. (2002). Lateral connectivity and contextual interactions in macaque primary visual cortex. *Neuron, 36,* 739–750. [PubMed] [Article]

Thielscher, A., Kölle, M., Neumann, H., Spitzer, M., & Grön, G. (2008). Texture segmentation in human perception: A combined modeling and fMRI study. *Neuroscience, 151,* 730–736. [PubMed]

Thielscher, A., & Neumann, H. (2003). Neural mechanisms of cortico-cortical interaction in texture boundary detection: A modeling approach. *Neuroscience, 122,* 921–939. [PubMed]

Thielscher, A., & Neumann, H. (2005). Neural mechanisms of human texture processing: Texture boundary detection and visual search. *Spatial Vision, 18,* 227–257. [PubMed]

Thielscher, A., & Neumann, H. (2007). A computational model to link psychophysics and cortical cell activation patterns in human texture processing. *Journal of Computational Neuroscience, 22,* 255–282. [PubMed]

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

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

von der Heydt, R., Peterhans, E., & Baumgartner, G. (1984). Illusory contours and cortical neuron responses. *Science, 224,* 1260–1262. [PubMed]

Weber, E. H. (1846). Tastsinn und Gemeingefühl. In R. Wagner (Ed.), *Handwörterbuch der Physiologie. Reprinted in W. Ostwald (1905), Klassiker der exakten Wissenschaften* (vol. 149). Leipzig, Germany: Engelmann.

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

Wersing, H., Steil, J. J., & Ritter, H. (2001). A competitive-layer model for feature binding and sensory segmentation. *Neural Computation, 13,* 357–387. [PubMed]

Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt. II. *Psychologische Forschung, 4,* 301–305. Translated as "Laws of organization in perceptual forms." In W. D. Ellis (Ed.) (1955), *A source book of Gestalt psychology* (pp. 274–282). London: Routledge & Kegan Paul. Also available online from http://psychclassics.yorku.ca

Williams, L. R. (1998). Fruit and texture images. http://www.cs.unm.edu/~williams/saliency

Würtz, R. P., & Lourens, T. (2000). Corner detection in color images through a multiscale combination of end-stopped cortical cells. *Image and Vision Computing, 18,* 531–541.

Yen, S. C., & Finkel, L. H. (1998). Extraction of perceptually salient contours by striate cortical networks. *Vision Research, 38,* 719–741. [PubMed]