Biol. Cybern. 81, 515–532 (1999)

Biological Cybernetics

Interaction of ON and OFF pathways for visual contrast measurement

Heiko Neumann[1], **Luiz Pessoa**[2], **Thorsten Hanse**[1]

[1] Universität Ulm, Fakultät für Informatik, Abteilung Neuroinformatik, Oberer Eselsberg, D-89069 Ulm, Germany
[2] Universidade Federal do Rio de Janeiro, Programa de Engenharia de Sistemas e Computação–COPPE Sistemas, Ilha do Fundâo, Rio de Janeiro, RJ-21945970, Rio de Janeiro, Brazil

Received: 5 June 1998 / Accepted in revised form: 12 May 1999

Abstract. We propose a novel model of visual contrast measurement based on segregated ON and OFF pathways. Two driving forces have shaped our investigation: (1) establishing a mechanism selective for sharp local transitions in the luminance distribution; (2) generating a robust scheme of oriented contrast detection. Our starting point was the architecture of early stages in the mammalian visual system. We show that the circuit behaves as a soft AND-gate and analyze the scale-space selectivity properties of the model in detail. The theoretical analysis is supplemented by computer simulations in which we selectively investigate key functionalities of the proposed contrast detection scheme. We demonstrate that the model is capable of successfully processing synthetic as well as natural images, thus illustrating the potential of the method for computer vision applications.

1 Introduction

An important goal of early visual information processing in man and machine is the robust and reliable detection of local luminance contrast. A discontinuity in a visual stimulus signals a causally related change in at least one physical scene parameter (Marr 1982). The development of filters to detect such discontinuities has been substantially influenced by neurophysiological investigations concerning the shape of receptive field (RF) profiles of retinal and cortical neurons. Marr and Hildreth (1980) proposed that the RF profile of mammalian retinal ganglion cells can be modeled by a Laplacian-of-Gaussian or difference of circular symmetric Gaussians, both implementing an isotropic second-order derivative filtering stage (Marr 1982). More recently, oriented linear filters built from Gabor or Gaussian derivative functions have been employed (Daugman 1985; Koenderink and van Doorn 1990). Such families of filters have been linked to the RF structure of cortical simple cells (Pollen and Ronner 1983) which sample the image in an optimal fashion and have been suggested to play a central role in human feature detection (Morrone and Burr 1988).

Initial processing stages of the retina and the lateral geniculate nucleus (LGN) utilize RFs with an isotropic center-surround structure. Antagonistic ON-center/OFF-surround and OFF-center/ON-surround cells comprise two independent parallel systems capable of signaling light increments and decrements (Schiller 1992). At the level of the striate cortex, ON-contrast and OFF-contrast cells converge onto the same target cells (Fig. 1). How do cortical simple cells acquire their polarity-sensitive elongated RFs? Hubel and Wiesel (1962) hypothesized that a simple cell sub-field is generated directly by excitatory synaptic inputs from a row of LGN neurons whose RF centers overlap the sub-field. The ON sub-field would be generated from ON-center LGN cells, and the OFF sub-field by OFF-center LGN cells. The precise mechanisms responsible for orientation selectivity remain, however, a subject of intense investigation (see a recent review by Vidyasagar et al. 1996). Nevertheless, there is evidence that an initial selectivity is brought about by the thalamo-cortical input arrangement (e.g., Ferster 1988), in line with the original Hubel and Wiesel (1962) proposal.

Local contrast information near localized luminance transitions (or boundaries) not only signals the presence of a discontinuity but also carries information about the surface qualities of neighboring image regions. In lightness computation of flat scenes, such as in the Retinex theory (Land and McCann 1971), by measuring the ratio of luminances at either side of a step transition, information about surface reflectance may be recovered. Such a class of lightness algorithms first differentiates the image and subsequently thresholds the result in order to maintain only contrast responses coinciding with sharp-edge-like transitions. Shallow gradients, i.e., those changes that are generated by a gradual illumination gradient, are thus "discounted." The edge information

Correspondence to: H. Neumann
(e-mail: hneumann@neuro.informatik.uni-ulm.de)

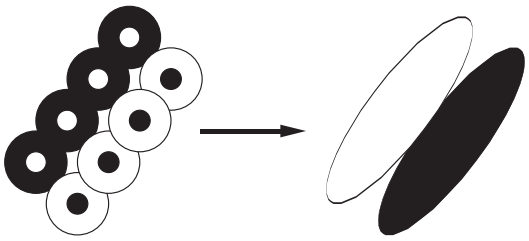

Fig. 1. Cortical simple cell receptive fields (RFs). Hubel and Wiesel (1962) hypothesized that lateral geniculate nucleus (LGN) cells with concentric, center-surround RFs project to the cortex with the proper arrangement so as to generate elongated RFs from unoriented ones. Schematic representation of how ON-center LGN cells connect with ON regions of cortical cells, and OFF-center LGN cells connect with OFF regions

left after threshold suppression is then integrated in order to generate a lightness representation (Horn 1974; Blake 1985; Arend and Goldstein 1987). A fixed threshold, however, cannot be determined for general scenes. If a low threshold is chosen, shallow gradients may be registered; if a high threshold is chosen, small contrast step-like transitions will be missed. This highlights a problem for the edge detection mechanisms mentioned above since their response amplitude is scaled with the local magnitude of the luminance gradient. What is needed is a mechanism that is selective to the shapes of the luminance transitions – i.e., whether they are abrupt or shallow.

In this paper, we develop a contrast detection circuit, referred to as a simple cell circuit, that exhibits such selectivity building upon a previous simplified implementation (Neumann and Pessoa 1994). The model uses segregated streams for ON contrasts and OFF contrasts. The proposed circuit (1) produces strong responses whenever ON and OFF retinal responses occur next to each other, such as at a luminance edge, while (2) producing minor responses when only ON or OFF responses are present, such as for gradual luminance gradients at ramp edges. To achieve this, the circuit behaves as a soft AND-gate whose responses have two components: a linear combination of ON and OFF channel contrast responses, and a multiplicative, or gating-like, component whenever inputs to the ON and OFF sub-fields are spatially adjacent, or juxtaposed. In essence, the combined computational strategy implements a robust scheme of oriented contrast detection and, in addition, establishes a mechanism that is specifically selective for sharp local luminance transitions.

In Sect. 2, we will introduce the proposed circuit for contrast detection. For this purpose, we first provide a formal description of the functionality of each model stage, including a center-surround pre-processing stage and the elements of the non-linear circuit itself. We then investigate the circuit's capability of contrast detection providing an analysis of the steady-state response. The spatial sampling of the input sub-fields varies with the parametrization of the spatial input blurring of the initial contrast measures which amounts to a multi-scale scheme of Gaussian derivative operators. Following this theoretical analysis, we show results of computer simulations in which we selectively investigate key functionalities of the proposed processing scheme. Finally, we relate the present model to other proposals that also have made use of non-linear or AND-gating processing for edge detection, most notably those of Marr and Hildreth (1980) and Iverson and Zucker (1995).

2 Neural model for oriented contrast measurement

Figure 2A shows the outline of the model. The input to the system is composed of two streams or channels, namely ON and OFF. The channels carry activations from initial processing stages having concentric center-surround RFs, as found in the retina and LGN. The model itself consists of a series of processing stages, each of which consists of a two-dimensional (2D) field (or grid) of processing units or cells. The input is also encoded as a 2D field of activity. All connections between model stages are topographically organized such that a spatial location (i, j) at a given stage connects to location (i, j) in the target field. The activation levels at individual model stages represent the output value of the respective stages.

2.1 Model stages

2.2.1 Center-surround interactions

The input luminance distribution, L, is processed by cells having center-surround, antagonistic, RFs, such as mammalian retinal ganglion cells. The model includes both ON and OFF pathways that measure the degree of

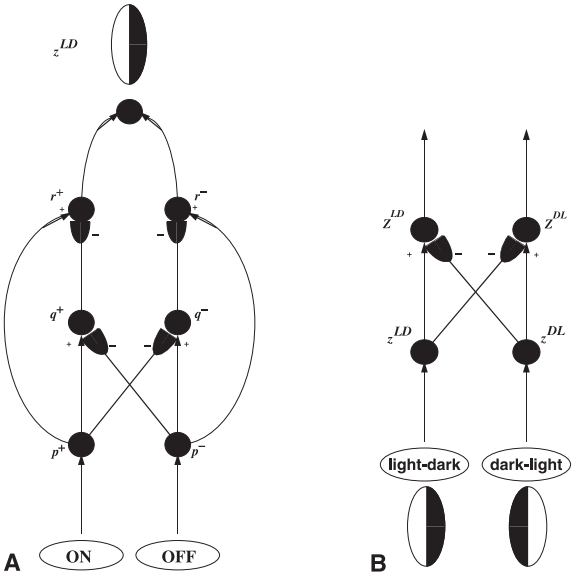

Fig. 2. **A** Proposed circuit of contrast cell sensitive to light-dark contrast polarity. **B** Final competition of cells of opposite contrast polarity at each spatial location. Two types of forward interactions have been utilized. Those denoted with an *arrow* at the end in an excitatory fashion, those with an *oval* at the ending of a link are inhibitory. *LD* Light-dark; *DL* dark-light

local luminance contrast in input images. Thus, at every spatial location there are ON and OFF cells. These two fields of cells implement lateral inhibition that contrast-enhance, or sharpen, the input luminance distribution – cells respond strongly to (unoriented) luminance discontinuities.

A mechanism of local center-surround processing can be approximated by filtering the input distribution with difference-of-Gaussian filters such as were originally proposed to model the receptive field structure of retinal ganglion cells (Rodieck 1965; Enroth-Cugell and Robson 1966):

$$\mathrm{net}^+ = L \otimes \mathrm{G}_{\sigma_+} \quad \text{and} \quad \mathrm{net}^- = L \otimes \mathrm{G}_{\sigma_-} \ , \tag{1}$$

where net^+ and net^- are net center and net surround inputs, respectively, for ON-channel retinal ganglion cells. The input luminance distribution is denoted by L, G_{σ_+} and G_{σ_-} are the respective Gaussian filters (parametrized by differently scaled space constants σ_+ and σ_-, respectively, with $\sigma_+ < \sigma_-$), and $\otimes$ implements the spatial convolution operator.

In classical proposals, such as, for example, in Rodieck (1965) or Marr and Hildreth (1980), the center and surround contributions combine linearly to determine the cell's response. Other approaches have specified visual adaptation processes that render responses sensitive to luminance ratios. Such formalisms adopt multiplicative, or shunting, mechanisms (Grossberg 1970; Hodgkin 1964; Furman 1965). Formally, the different computational approaches can be denoted in a generalized framework such that an ON-center OFF-surround interaction (generating y^+ activation) results in

$$\frac{d}{dt}y^+ = -\alpha_s y^+ + (\beta_s - \gamma_s y^+)\mathrm{net}^+ - (\delta_s + \eta_s y^+)\mathrm{net}^- \ . \tag{2}$$

Here α_s, β_s, γ_s, δ_s and η_s are constants, and net^+, net^- are the total excitatory/inhibitory inputs to y^+ (see Eq. 1). The rest of the paper focuses on the description of a simple cell circuit that is sensitive to oriented contrast. Here, the arrangement of ON/OFF -contrast is relevant and, therefore, we can restrict ourselves to the simplest case having linear ganglion cell responses that generate ON and OFF channel inputs.

The steady-state contrast input to the simple cell circuit is given by the equilibrium form of Eq. (2) (i.e., $d/dt\, y^+ = 0$). By setting $\alpha_s = \delta_s$ and $\gamma_s = \eta_s = 0$, we obtain zero DC contrast channel inputs

$$\begin{aligned} c^+ &= \frac{\beta_s}{\alpha_s} \cdot \mathscr{R}[\mathrm{net}^+ - \mathrm{net}^-] \quad \text{and} \\ c^- &= \frac{\beta_s}{\alpha_s} \cdot \mathscr{R}[\mathrm{net}^- - \mathrm{net}^+] \end{aligned} \tag{3}$$

($\mathscr{R}[x]$ denotes a rectification operation $\max[x, 0]$).[3]

[3] In the case of having imbalanced center-surround interaction, an additional stage of cross-inhibition between contrast channels helps to eliminate any residual DC-level activation (see Neumann 1996).

2.1.2 ON/OFF input to contrast cell sub-fields

The spatial branches of contrast cells receive excitatory inputs from the ON and OFF contrast channels. This input is sampled separately using elongated weighting functions with a smoothly decaying coupling strength. Before integration of contrast activity for excitatory and inhibitory sub-fields, the activities in both contrast channels undergo local competition at each spatial location. Such competitive interaction is necessary to generate sufficient initial orientation tuning of the cell. Otherwise, with separate blurring of each input channel activity, simple cells will be activated even for an orientation orthogonal to the local contrast. In formal terms, the input to either sub-field of the circuit is modeled by

$$\begin{aligned} p^+ &= \mathscr{R}[(c^+ - c^-) \otimes \lambda_\varepsilon^+] \quad \text{and} \\ p^- &= \mathscr{R}[(c^- - c^+) \otimes \lambda_\varepsilon^-] \ , \end{aligned} \tag{4}$$

(see Fig. 2), where the kernel λ determines an elongated spatial sampling function of orientation ε.

These inputs are used to detect local contrast changes of a given polarity, such as the presence of edges or other abrupt luminance variations. In our simulations, we used elongated 2D Gaussian weighting functions for modeling the sensitivity profiles of $\lambda_\varepsilon^\pm$ (see Fig. 3), with space constants σ_m and σ_M to denote the minor axis and the elongation axis, respectively. In formal terms, the anisotropic weighting kernel can be described by

$$\lambda_\varepsilon = N_\sigma \exp\left[-\tfrac{1}{2}\left(\vec{\mathbf{x}}^T \mathbf{R}_\varepsilon^T \mathbf{C}_{mM} \mathbf{R}_\varepsilon \vec{\mathbf{x}}\right)\right] \ , \tag{5}$$

where N_σ is a scaling constant, $\vec{\mathbf{x}}^T = (x\ y)$ denotes the (transposed) vector of the spatial location, $\mathbf{C}_{mM}$ is the diagonal matrix

$$\mathbf{C}_{mM} = \begin{pmatrix} 1/\sigma_\mathrm{m}^2 & 0 \\ 0 & 1/\sigma_\mathrm{M}^2 \end{pmatrix} \ ,$$

determining the spreading along the principal axes, and $\mathbf{R}_\varepsilon$ determines the matrix of rotation in 2D (x, y) space

$$\mathbf{R}_\varepsilon = \begin{pmatrix} \cos\varepsilon & \sin\varepsilon \\ -\sin\varepsilon & \cos\varepsilon \end{pmatrix} \ .$$

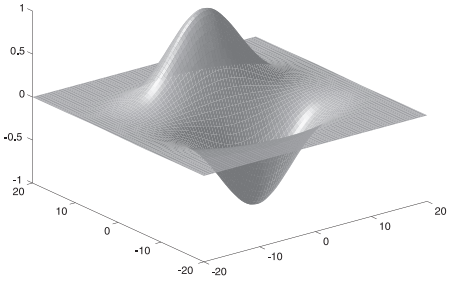

Fig. 3. Effective spatial sensitivity profiles of oriented contrast cells with odd-symmetric ON and OFF input. The model uses elongated Gaussian profiles, λ_ε^+ and λ_ε^-, both with space constants $\sigma_\mathrm{m} = 3$ and $\sigma_\mathrm{M} = 2\sigma_\mathrm{m}$; relative offset has been set to $\tau = 3$, orientation is $\varepsilon = 0°$ (see text)

The axis of elongation indicated by ε determines the preferred orientation of the cell.

In all cases, we used balanced weighting functions for excitatory and inhibitory sub-fields, such that $\lambda_{\varepsilon}^{+} = \lambda_{\varepsilon}^{-}$. By varying the space constants σ_{m} and σ_{M} of the elongated weighting functions, one obtains kernels of different spatial extent in length and width. This defines a family of kernels denoted as $\lambda_{\varepsilon}^{S+}$ and $\lambda_{\varepsilon}^{S-}$, where S denotes the spatial scale (see the notion of scale below in Sect. 3). The relative spatial offset τ of these sub-fields (see Fig. 3) varies as a function of the space constant σ_{m}, such that we have $\tau(\sigma) = \sigma_{\mathrm{m}}$.

2.1.3 Sub-field opponent interactions

The second stage of the circuit receives direct linear excitatory inputs from each ON or OFF sub-field, as well as an opponent divisive interaction between channels. The steady-state activity of the ON sub-field is given by

$$q^{+} = \frac{p^{+}}{\alpha_c + \beta_c p^{-}} \quad , \tag{6}$$

where α_c and β_c are constants of this shunting inhibitory interaction. For the OFF sub-field, we assume the same setting of constants such that at equilibrium the q^{-}-response is generated by having the p^{+} and p^{-} inputs interchanged in the right-hand side of Eq. (6).

2.1.4 Direct input and post-opponency signals combined

The third stage receives channel-specific inputs from both the first (excitatory) and second (inhibitory) stages. Formally, ON-channel equilibrium activation is

$$r^{+} = \frac{p^{+}}{\gamma_c + \delta_c q^{+}} \quad , \tag{7}$$

where γ_c and δ_c are constants of this second stage of the circuit. Again, we assume the same setting of constants in both channels. The OFF-channel activations r^{-} are then obtained by exchanging '+' and '−' indices in the right-hand side of Eq. (7).

The functionality of the sub-circuits realizing each separate sub-field channel is to generate a self-inhibition, thus normalization, of the input activity distribution. The opponent interaction between these channels and the corresponding overall net activity will be discussed below under functional aspects.

The final response for a light-dark (LD) cell is obtained by pooling ON and OFF activities. We formalize this as a simple linear combination by summing both r activities and get

$$z^{\mathrm{LD}} = r^{+} + r^{-} \quad . \tag{8}$$

The dark-light (DL) response, z^{DL}, is obtained in a corresponding manner, such that the respective ON and OFF input streams are reversed in their spatial arrangement.

2.1.5 Mutual inhibition of cells

Cells sensitive to opposite contrast polarity at the same spatial position finally undergo mutual inhibition (see Fig. 2B). In functional terms, such a competitive stage helps to sharpen the activity profile generated by each cell sensitive to opposite contrast polarity alone. The final cell responses are computed as

$$\begin{aligned} Z^{\mathrm{LD}} &= \mathscr{R}[z^{\mathrm{LD}} - z^{\mathrm{DL}}] \quad \text{and} \\ Z^{\mathrm{DL}} &= \mathscr{R}[z^{\mathrm{DL}} - z^{\mathrm{LD}}] \quad . \end{aligned} \tag{9}$$

2.2 Relationship to cortical simple cell physiology

Several design issues and functional properties of the model have strong relationships to findings about simple cell physiology. For example, we utilize oriented spatial weightings to model elongated sub-fields for integrating non-oriented input from cells with a concentric RF outline (see Fig. 1). This is consistent with evidence that an initial selectivity is brought about by the thalamocortical input arrangement (e.g., Ferster 1988), in line with the original Hubel and Wiesel proposal (1962). Furthermore, we have utilized a local competition of contrast channels before integrating sub-field activities. This approach implements the opponent inhibition mechanism for simple cell sub-fields as described in Ferster (1989; see also Tolhurst and Dean 1990). Also, using partially overlapping spatially offset sub-field weighting functions has been suggested by the findings of Heggelund (1981, 1986) as well as Ferster (1988).

Our circuit integrates the activity from the corresponding sub-field (ON or OFF) in an excitatory fashion, while the opponent influences this by inhibitory action. This scheme now realizes a mechanism of opponent inhibition between sub-fields, similar to those in other simple cell models (see Pollen and Ronner 1983). The mutual inhibition of cells with opposite polarity selectivity follows the description in Liu et al. (1992), who suggested the mutual inhibition between members of pairs of phase-related opposite contrast simple cells. Also, Ferster (1988) suggested that competition between simple cells of opposite contrast polarity occurs at each spatial location. In all, although we do not attempt to explicitly model the underlying mechanisms of cortical simple cells (e.g., orientation selectivity and spatial input sampling), our model incorporates key features of simple cell physiology. We therefore suggest that our model circuit also serves as an abstract description of information processing at early stages in the mammalian visual system.

2.3 Functionality of the circuit

2.3.1 Equilibrium response

It is possible to gain insight into the precise functionality of the model, especially with regard to its AND-gate behavior, by combining the corresponding expressions. The activation r in the ON and OFF branches can thus be computed analytically. An assumed symmetry relationship between both channels can be achieved by setting the identity $\delta_c = \beta_c \gamma_c$. In this case, for a LD cell, we get

$$r^+ = \frac{p^+}{\alpha_c\gamma_c + \delta_c(p^+ + p^-)}(\alpha_c + \beta_c\, p^-) \quad \text{and}$$
$$r^- = \frac{p^-}{\alpha_c\gamma_c + \delta_c(p^+ + p^-)}(\alpha_c + \beta_c\, p^+) \ . \tag{10}$$

The final response for a contrast cell selective to LD polarity is computed as $z^{\mathrm{LD}} = r^+ + r^-$. Using Eq. (10) we get

$$z^{\mathrm{LD}} = \frac{1}{\alpha_c\gamma_c + \delta_c(p^+ + p^-)}[\alpha_c(p^+ + p^-) + 2\beta_c\, p^+p^-] \ . \tag{11}$$

The response of a DL selective cell, z^{DL}, is obtained in a corresponding manner. In the final steady-state response, however, the roles of p^+ and p^- are interchanged.

This demonstrates the non-linear interaction of activity between the two branches. The basic functional properties of the model could be highlighted by the following more compact notation of Eq. (11) that reads

$$z = \mathscr{L}\cdot(p^+ + p^-) + \mathscr{N}\cdot(p^+\cdot p^-) \ . \tag{12}$$

Here, $\mathscr{L}$ and $\mathscr{N}$ denote transformations of the initial contrast activities in the ON and OFF channels that generate the inputs to the circuit. In particular, we have $\mathscr{L} = \alpha_c\cdot\Gamma(c^+, c^-)$ and $\mathscr{N} = 2\beta_c\cdot\Gamma(c^+, c^-)$ with $\Gamma(c^+, c^-) = [\alpha_c\gamma_c + \delta_c(p^+ + p^-)]^{-1}$. In $\Gamma(\cdot)$, the input activities $p^{\pm}$ for each individual lobe are themselves functions of c^+ and c^- (see Eq. 4). Input to the model circuit is integrated linearly from both channels. Spatially adjacent activity (in the ON and OFF pathways) is signaled by an additional correlational (gating-type) component. The relative contribution of additive and gated activities is controlled by the (shunting) parameters α_c and β_c in Eq. (6). Moreover, the activity self-normalizes with respect to the total input activity from the ON and OFF channels [function $\Gamma(c^+, c^-)$ above]. The model thus relates to the scheme proposed by Carandini and Heeger (1994) in which the activity of cortical neurons is normalized through division of pooled activity from a large number of cells.

2.3.2 Linear and non-linear responses

In all, the model behaves as a soft AND-gate that combines linear and non-linear (gating) components of sub-field activity. Adjacent ON and OFF signal configurations generate an extra boosting activation in addition to responses of reduced magnitude that are generated for other signal configurations. In terms of the circuit shown in Fig. 2A, it is the opponent inhibition of the second stage (signals q^+ and q^-) associated with the within-channel inhibition (from the second to the third stages) that implements the soft AND-gate behavior. These interactions produce a mechanism of disinhibition that generates large outputs only when both ON and OFF channel inputs are large. To understand this, consider the case where only ON signals are input to the model; the OFF pathways are thus shut down. Responses will be small since the (within channel) inhibition from the second to the third stages will attenuate the input signal via a normalization mechanism. Now consider the case when there are potent inputs to both ON and OFF channels. The cross-inhibition between the first and second stages will largely reduce second stage activities (signals q^+ and q^-). These, in turn, will not be able to inhibit stage three signals and the original inputs will be able to combine at the last stage (z^{LD} and z^{DL}) since they reach it via the "side pathways" from the first to third stages. We see that the presence of inputs in both channels leads to a disinhibition in the circuit and, hence, powerful responses. In all, the circuit detects when there are adjacent ON and OFF signals, hence generating the soft AND-gate behavior.

The relative proportion of linear and non-linear responses is controlled by the parameters α_c and β_c. The net input scaling for the initial ON and OFF channel processing also directly contributes to the contrast cell response in a systematic fashion. ON and OFF -channel input to oriented sub-fields is given by

$$p^{\pm} \propto \frac{\beta_s}{\alpha_s}\cdot\mathscr{R}[(\mathrm{net}^{\pm} - \mathrm{net}^{\mp}) \otimes \lambda_{\varepsilon}^{\pm}] \tag{13}$$

(compare Eqs. 3 and 4). We observe for a given net input that the relative proportion of the non-linear contribution in Eq. (11) is $\beta' = \beta_s\beta_c/\alpha_s$. For any further analysis, one must therefore take into account the range of activities generated in the initial center-surround preprocessing stage. For convenience, an activity range of [0..1] from initial center-surround filtering has been ensured in all our simulations.

The key idea behind our model is that contrast changes are better localized by a circuit that is selectively sensitive to abrupt luminance transitions. These transitions will be invariably associated with adjacent ON and OFF contrast signals (for some spatial scales; see below). A linear combination scheme is, of course, sensitive to such transitions, but not sensitive enough such that it shows only differential variations in response once previously adjacent activities in the ON and OFF channels have been shifted farther apart (see also du Buf 1994).

3 Families of RFs and selectivity to spatial scale

3.1 Elementary properties

3.1.1 Motivation

The visual system is faced with the problem of measuring the relevant structure of the outside world. Real-world objects only exist as meaningful entities over certain ranges of measurement scales. An example stressed by Marr and Hildreth (1980) is the appearance of the fine-textured veins of a leaf on a small scale but the overall shape and luminance variation appearing on a much coarser scale. Furthermore, more gradual intensity variations often relate to illumination variation or changes in surface orientation, whereas abrupt changes in the luminance distribution are often caused by transitions in surface properties, such as reflectance

(Horn 1974). A vision system, in general, has no advanced knowledge about the appropriate scales for analysis. Therefore, in order to deal with such a multitude of properties of real-world objects, the machinery of a vision system has to process the input luminance distribution on different scales (see, for example, Witkin 1983; Koenderink 1984; Lindeberg and ter Haar Romeny 1994; and for a short tutorial, Lindeberg 1996).

The raw luminance distribution is processed by an initial center-surround mechanism. After opponent inhibition, this activity is integrated by oriented offset weighting functions which define the excitatory and inhibitory sub-fields to an oriented cell. The signals for ON and OFF contrast responses are rectified such that the contrast channels carry a single positively bounded data representation from the initial filtering stage. By utilizing weighting functions of different sizes, λ_ε^{S+} and λ_ε^{S-}, these contrast activities are blurred and sampled at increased relative spatial offsets; thus a family of cells sensitive to different spatial scales is introduced. The spatial scale attribute S itself is a function of the space constant σ, thus $S = S(\sigma)$ (see Eq. 5).

3.1.2 Generation of scale-space derivative kernels

A contrast edge is processed by an initial center-surround filtering stage. The results are half-wave rectified and segregated into two separate representations for ON and OFF contrast, respectively. The results of individual ON and OFF responses are blurred by oriented spatially scaled weighting functions to generate the input for the opponent contrast cell sub-fields of different sizes. This non-linear processing cascade is shown to be identical to a first-order derivative kernel applied to the activation generated by the center-surround processing.

To simplify the mathematical analysis we approximate the initial center-surround filtering by a Laplacian-of-Gaussian (LoG) operator. Consider a unit step function as input that is modeled as a Heaviside function $\mathscr{H}$ (Bracewell 1978). Input processing results in a profile of a first-order derivative of a Gaussian, $D1\mathrm{G}$, which is an odd-symmetric function with positive and negative lobes. For a representation to carry positive signal responses only, the positive and negative lobes are half-wave rectified. We get

$$c^+(u) = \mathscr{R}[\mathrm{LoG}_{\sigma_1}(u) \otimes \mathscr{H}(u)] = \mathscr{R}\left[\frac{d}{du}\mathrm{G}_{\sigma_1}(u)\right] \quad \text{and}$$

$$c^-(u) = \mathscr{R}[-\mathrm{LoG}_{\sigma_1}(u) \otimes \mathscr{H}(u)] = \mathscr{R}\left[-\frac{d}{du}\mathrm{G}_{\sigma_1}(u)\right] \ ,$$

with $\mathscr{R}[x] = \max[x, 0]$. These responses are subsequently filtered by a Gaussian blurring function whose results are then combined to generate the edge response of the circuit. We get $s(u) = \mathscr{R}[d/du\mathrm{G}_{\sigma_1}(u)] \otimes \mathrm{G}_{\sigma_2}(u) + \mathscr{R}[-d/du\ \mathrm{G}_{\sigma_1}(u)] \otimes \mathrm{G}_{\sigma_2}(u)$. Since $\mathscr{R}[d/du\mathrm{G}_{\sigma_1}(u)] = d/du\mathrm{G}_{\sigma_1}(u) \cdot [1 - \mathscr{H}(u)]$ and $\mathscr{R}[-d/du\mathrm{G}_{\sigma_1}(u)] = -d/du\mathrm{G}_{\sigma_1}(u) \cdot \mathscr{H}(u)$; we can rewrite the final response to get

$$s(u) = \frac{d}{du}\mathrm{G}_{\sigma_1}(u) \otimes \mathrm{G}_{\sigma_2}(u) = \frac{d}{du}\mathrm{G}_{\sqrt{\sigma_1^2+\sigma_2^2}}(u) \ . \qquad (14)$$

This demonstrates the equivalence of our hierarchical processing scheme based on segregated representations of ON and OFF activations and an approach that utilizes filters derived from an analytic description of first-order derivative operations. It also verifies that the subsequent blurring of separated contrast channel responses of one frequency band – generated by the initial center-surround processing stage – produces results equivalent to those achieved with a bank of scaled center-surround filters.

3.1.3 Opponent offset blurring approximates first-order derivative filtering

The kernels $\lambda^{\pm}$ were spatially offset by an amount τ relative to the reference location of the target cell. They collect responses in the ON and OFF contrast channels to be further processed in the circuit. The difference of spatial offset Gaussians resembles the profile of a first-order derivative operation. Consider the model of a parametrized ramp edge transition in a local (u, v)-gauge coordinate system where the u-axis is oriented along the contrast profile. Formally, the difference of offset Gaussians can then be written as a convolution of the $D1\mathrm{G}$ profile with a unit impulse of width $\Delta = 2\tau$, then taking the limit, $\Delta \to 0$. We get

$$\lim_{\Delta\to 0}\frac{1}{\Delta}[\mathrm{G}(u) - \mathrm{G}(u-\Delta)] = \lim_{\Delta\to 0}\left[\frac{d}{du}\mathrm{G}_\sigma(u) \otimes \Pi_\Delta(u)\right] = \frac{d}{du}\mathrm{G}_\sigma(u) \ ,$$

where $\Pi_\Delta(u) = \mathscr{H}(u + \frac{\Delta}{2}) \cdot \mathscr{H}(\frac{\Delta}{2} - u)$.

3.1.4 Responses as a function of scale

We now consider the processing of a luminance transition with a ramp profile of a priori unknown width (compare du Buf 1993, 1994). The suggested model is sensitive to the spatial adjacency of retinal ON and OFF signals. Spatial proximity, however, is a relative measure that will depend on the spatial scale of processing. What may be spatially proximal for a large scale may be distant for a small one. Sensitivity to spatial scale endows model cells to selectively process a given visual structure, such as the ramp transition shown in Fig. 4. For a wide transition ramp, the initial center-surround processing generates isolated ON and OFF contrast responses which are located at the shoulders ("knees") of the ramp transition. A cell with a RF selective to odd symmetric contrast variations located at position $x = 0$ at the center of the ramp for the smallest scale receives no input from either blurred ON or OFF contrast channels (compare scale 1 in Fig. 4 middle, bottom). Cells slightly offset to either side of this central location receive input from only one sub-field (lobe) and thus generate normalized linear responses. As the blurring proceeds, more diffuse distributions of ON and OFF contrast responses appear, which eventually meet at the center of the ramp such that they appear to be juxtaposed for the coarse scale at location $x = 0$ (scales 3 and 4 in Fig. 4, bottom). Now integrated inputs from both sub-fields add and a superimposed component is generated by the product of the magnitude of these inputs. If the product is

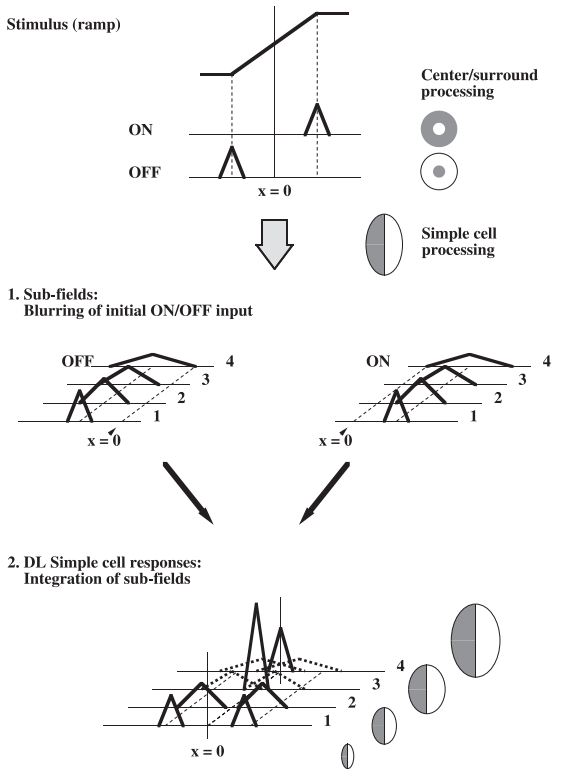

Fig. 4. Responses of a scaled family of contrast cells sensitive to an odd-symmetric dark-light (*DL*) luminance transition (ramp contrast). *Top* Input stimulus is shown together with schematic ON and OFF responses generated by the initial stage of center-surround processing. Responses of either contrast channel appear at the "knee points" of the plateaus. Depending on the width of the ramp relative to the size of the initial center-surround filter, they appear in spatial isolation (as shown in the sketch) or juxtaposed. *Middle* Increased blurring of initially isolated contrast responses for a wide transition ramp. Four different scale selectivities are shown with reference to the location at the center of the ramp ($x = 0$). These blurred activations provide the inputs to cells sensitive to DL transitions. *Bottom* Responses of contrast cells for the four scales that integrate their input to sub-fields from increasingly distant offsets [the display of responses for light-dark cells has been omitted in order to keep the sketch simple]. For a cell at a more coarser scale, ON and OFF inputs appear to be juxtaposed such that the non-linear component can contribute via its ignition. As a consequence, a unique maximum response is generated as a function of the scale

sufficiently high in amplitude, the latter contribution facilitates the target cell activation by an additional "boost." The spatial juxtaposition of ON and OFF input activation to the circuit signals a transition which is sharply localized relative to the current spatial scale.

The circuit is thus selective to inputs with sharp contrast transitions. The quantitative properties of the cell responses in scale-space can be derived from the activation given in Eq. (11). Below, we present the results of this investigation.

3.2 Quantitative analysis of scale-space response

In this subsection, we summarize the quantitative findings for responses generated by scaled contrast cells which are applied to different idealized luminance transitions. We first show the results derived for step edge responses. This motivates the subsequent derivation for the results processing the gradual transition of a ramp as discussed above. In both studies, we compare the results for the non-linear circuit with those generated by a scheme of linear filtering. Detailed results of investigations can be found in Appendices 1 and 2.

We utilize the compact notation of a non-linear response of the circuit that is given by

$$z = \mathscr{L} \cdot \left[(p^+ + p^-) + \frac{\mathscr{N}}{\mathscr{L}} \cdot (p^+ \cdot p^-) \right] \tag{15}$$

(compare with Eq. 12). For quantitative analysis, we omit the gain $\mathscr{L}$ and only consider the expression given in brackets. For the gain of the non-linear component, we define

$$\mathscr{N}' \equiv \frac{\mathscr{N}}{\mathscr{L}} (\beta_s + \delta_s) = 2 \frac{\beta_c}{\alpha_c} (\beta_s + \delta_s) \tag{16}$$

(see Appendix 2). Based on these definitions, we relate the linear and non-linear response components.

3.2.1 Step edge response

A step edge profile of height h is defined by a Heaviside function,

$$f_{\text{step}}(u) = h \cdot \mathscr{H}(u) \ , \tag{17}$$

in which the u-axis is taken in the direction orthogonal to the local orientation tangent to the luminance contrast. Processing this input first by a center-surround filter (approximated by a Laplacian-of-Gaussian) subsequently followed by a first-order (Gaussian) derivative filtering generates a response profile with a unique maximum at the step location $u = 0$. For the non-linear circuit, the segregated contrast responses always appear to be juxtaposed at either side of the contrast step. The relevant component of response (contribution given in brackets in Eq. 15) of the non-linear circuit at the location of the luminance step is given by

$$s^{\text{nl}}_{\text{step}}(u = 0) = \frac{h}{\sqrt{2\pi}\sigma^3} \left(1 + \mathscr{N}' \frac{1}{\sqrt{2\pi e}\sigma} \right) \ . \tag{18}$$

This shows that for a linear as well as a non-linear cell, the response is maximal for the smallest scale, i.e., it smoothly drops as blurring increases. Furthermore, it shows that the non-linear contribution only becomes significant for a gain $\mathscr{N}' \gg \sqrt{2\pi e}\sigma \approx 7\sigma$. Considering the definition of the gain in Eq. 16, it is reasonable to treat the parameter β_c as a non-zero function of σ.

3.2.2 Ramp edge response

A ramp edge profile of height h and width R is defined by

$$f_{\text{ramp}}(u) = h \cdot \mathscr{H}(u) \otimes \Pi_R(u) \ . \tag{19}$$

where $\Pi_R(u)$ denotes a finite pulse of width R. In the limit as $R \to 0$, the ramp converges to a step function as defined above. Again, we utilize the sequence of processing steps of center-surround filtering and the

first-order Gaussian derivative operation. The latter is approximated by the individual blurring of the contrast responses followed by a numerical differentiation for final sampling of the inputs from the sub-fields of the contrast cell. Now, depending on the ratio between the operator scale and the width of the ramp, σ/R, the ON and OFF contrast responses appear in isolation ($\sigma \ll R$) or juxtaposed similar to the case of a step transition ($\sigma \approx R$). For the position $u = 0$ at the center of the ramp, we can now treat the response as a function of scale. For the linear model, a relative maximum in response occurs at

$$\sigma_{\max} = \frac{R}{2\sqrt{3}} \quad . \tag{20}$$

For the analysis of scale-space response for the non-linear circuit, we investigate the additional contribution from (blurred) ON and OFF contrast responses sampled from offsets relative to the ramp center. The relevant component (bracket part in Eq. 15) of the corresponding response for the circuit is

$$s^{\mathrm{nl}}_{\mathrm{ramp}}(u=0)|_{\sigma=\sigma_{\max}} = \frac{48\sqrt{3}h}{\sqrt{2\pi}} \exp\left[-\frac{3}{2}\right] \cdot \frac{1}{R^3} \times \left(1 + \mathscr{N}' \frac{\sqrt{2}}{8\sqrt{3\pi}} \frac{h}{R} \exp\left[\frac{3}{2}\right]\right) \quad . \tag{21}$$

This shows that there exists an optimum scale for processing a ramp edge of a priori unknown width and height. The non-linear contribution is itself dependent on the slope of the ramp and becomes significant for a gain $\mathscr{N}' \gg 4\sqrt{6\pi}\exp\left[-\frac{3}{2}\right]m^{-1} \approx 4m^{-1}$, where $m = h/R$.

The non-linear circuit shows a multitude of desired principles. For one, the additional "boosting" of activity helps to drive the cell output to generate a response that is maximal in absolute terms. It is shown in Appendix 2 that the non-linear cell tuned to an optimum scale generates absolute maximum amplitude responses. In addition, the contribution generated by the non-linearity depends on the scale σ (step) or on the slope m (ramp) of the luminance transition. The parameter to control the efficacy of the non-linear contribution in $\mathscr{N}'$ is β_c, the gain of opponent shunting inhibition in the circuit (see Eq. 6). The gating of contrast activations – adding to the (linear) cell response – enables the circuit to preferentially respond to sharp transitions, irrespective of their height. This selective functional property was one of the principal motivations for the development of the circuit in the context of the modeling of functional mechanisms for neural contrast and brightness processing (Neumann and Pessoa 1994; Pessoa et al. 1995).

4 Simulations

In this section, we evaluate the model on the basis of a series of computer simulations that demonstrate the functionality. The section is organized so as to start with the demonstration of the usefulness of the approach for image processing purposes. We tested its robustness to noise in comparison with a corresponding pure linear model and demonstrated its selectivity in automatic scale selection, all based on synthetic test images. The computational relevance is also demonstrated by means of natural images from a test data set. Since we have also shown that the model accounts for some relevant properties of cortical simple cells that are usually believed to behave almost linearly, we conducted an experiment showing that our non-linear circuit also shows substantial linear behavior. We then finally show results that demonstrate the usefulness of the model as a first step in a processing hierarchy in which initial measurements are grouped together in order to generate meaningful pieces for shape recognition. In particular, we show the model's strength in the initial measurement for subsequent integration by long-range grouping processes.

In all simulations, responses are shown after mutual inhibition (Z^{LD} and Z^{DL}). Luminance values of input images have been normalized to the range [0..1]. The initial center-surround processing stage involves shunting interaction and subsequent half-wave rectification to generate segregated ON and OFF channels (see Sect. 2.1). Model parameters for this initial stage were set to $\alpha_s = 0.5$, $\beta_s = 1.0$, and $\delta_s = 0.1$ (as already pointed out above, we set $\gamma_s = \eta_s = 1.0$ in order to yield symmetric normalization in both channels). Parameters of isotropic Gaussians were set to $\sigma_+ = 1.0$ and $\sigma_- = 3.0$, thus obtaining a 1 : 3 center-surround ratio. The model parameters of the non-linear circuit were set to $\alpha_c = 1.0$, $\beta_c = 10000.0$, $\gamma_c = 0.01$, and $\delta_c = 100.0$. Their specific choice is non-critical as long as the linear components scaled by α_c and γ_c are small compared to the cross-channel inhibition effect. The Gaussian weighting functions were elongated by a $\sigma_{\mathrm{M}} : \sigma_{\mathrm{m}} = 2 : 1$ ratio; the variance σ_{m} is measured in pixels along the short axis. The separation τ grows linearly with the variance (scale S). Eight discrete, equally spaced orientations were processed. The corresponding linear circuit was approximated simply by eliminating the opponent interaction between the sub-fields in the non-linear contrast cell circuit (stages $q^{\pm}$ and $r^{\pm}$, see Fig. 2). Thus, the scheme used for comparison directly integrates the activity from both sub-fields, namely $z = p^+ + p^-$.

4.1 Image processing

As indicated, we demonstrate the capacity of the circuit on the basis of synthetic as well as real camera images from a test set. Test simulations run on synthetic data were to justify the robustness of the model against noise and the selectivity of the mechanisms to scale. Results generated for natural images are intended to serve for comparison of performance with other known methods.

4.1.1 Synthetic images

A first strict test of an image processing algorithm consists of probing it with noisy images. Figure 5 shows a synthetic image containing an elliptic region embedded

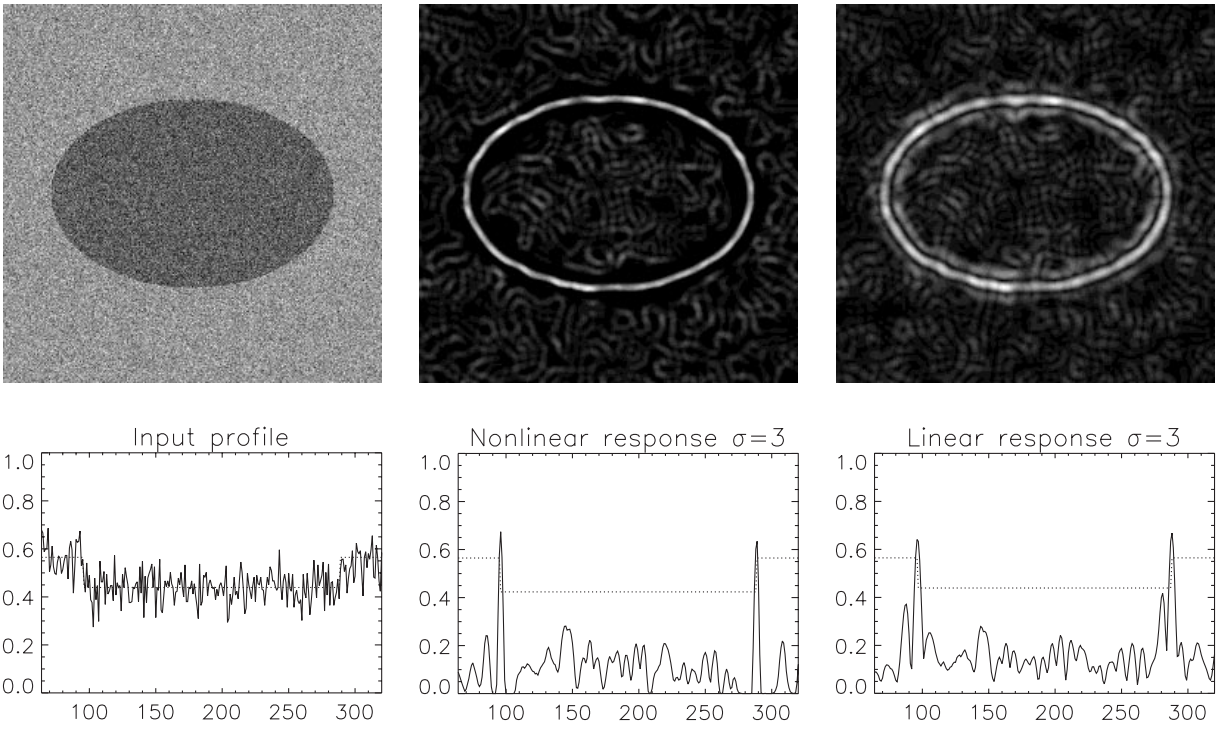

Fig. 5. Synthetic image with noise. Input luminance distribution corrupted by additive Gaussian noise with 50% amplitude of contrast height (*top left*); pooled activity for all orientation fields generated by the non-linear circuit (*top middle*) and the linear circuit (*top right*). Corresponding one-dimensional cross-sections of the above shown 2D distributions (*bottom row*)

in a lighter background (left) and a cross-section profile. The entire image was corrupted by Gaussian noise (half-width 50% amplitude). Figure 5 (middle) shows the output of the new model revealing that it is capable of accurate contour localization, even in the presence of non-trivial noise levels. It is also instructive to compare the performance of the circuit with the linear scheme. As shown in Fig. 5 (right), the linear scheme is less robust to noise as it also shows less selectivity in terms of edge localization. Note that no post-processing, such as final thresholding operations, was performed – these could be used to remove the low-intensity spurious signals due to noise.

In order to evaluate the theoretical derivations on the mechanism's selectivity to spatial scale we have conducted a series of experiments based on the processing of scaled gradual luminance transitions, namely ramps. Figure 6 illustrates the spatial frequency selectivity of the non-linear model by comparing its behavior with the analogous linear scheme. For the non-linear contrast cell, as the spatial scale of operation is increased, cells whose centers are closer to the middle of the ramp eventually become ignited by the non-linear proportion of the response. That is, by using larger elongated Gaussian operators that sample the blurred ON and OFF sub-field inputs target cells conjointly receive input from offset positions as they appear to be juxtaposed in a coarser scale – a maximum response is produced at the middle of the ramp. This behavior should be contrasted to that of the linear scheme. It can be recognized that although maxima evolve at the middle of the ramp, they are only local such that the absolute maximum response appears at a fine spatial scale, localized to the left and right of the "knee" points.

4.1.2 Natural images

Natural images provide a good test of the image processing capabilities of our model of contrast detection. In particular, we can assess the contrast localization properties of the model by again comparing its output with that produced by an analogous linear scheme. Figure 7 illustrates the better contrast localization properties of our circuit when compared to the linear scheme. We see that much sharper "edge signals" are generated by the circuit effectively registering the contour outlines present in the image.

4.2 Linear versus non-linear processing behavior

We have pointed out above (see Sect. 2.2) that our model inherits a number of simple cell properties. We developed a non-linear circuit that pronounces any input configuration of juxtaposed ON and OFF activation, thus being specifically selective to sharp luminance transitions. In order to justify the relevance of the model, it has to be demonstrated that the overall behavior does not, in general, contradict previous findings in physiology that have been taken as support for linear mechanisms driving simple cells. Candidate tests include the linear summing principle given increasing stimulus length and using eigenfunctions of linear systems such

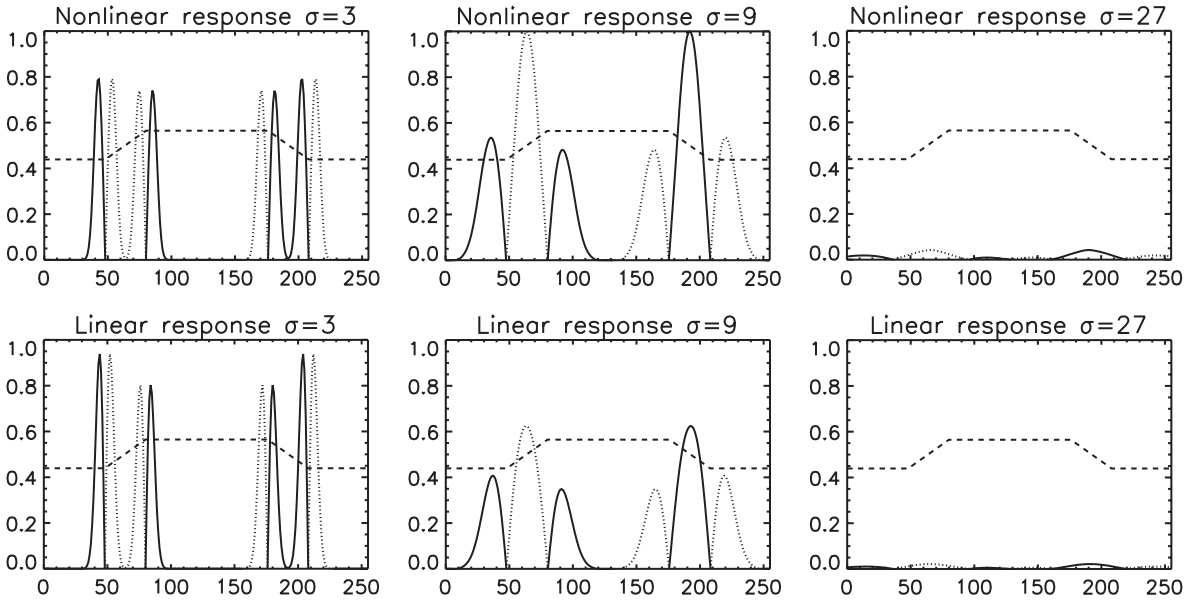

Fig. 6. Scale-space selectivity. A ramp luminance distribution is processed at small, medium and large spatial scales. For better visualization of the responses, we plot one-dimensional cross-sections of the ramp as well as of responses generated with the non-linear circuit (*top*) and linear scheme (*bottom*). Responses are shown for light-dark cells (*solid line*) and dark-light cells (*dotted line*). The ramp widths were parametrized by $R = 32$ units

Fig. 7. Processing a real camera image (*left*). Pooled activity of all orientation fields generated by the non-linear circuit (*center*) versus the linear model (*right*). Note that no further post-processing, such as local non-maximum suppression, has been applied to the data

as spatial sine-wave luminance modulations. Sine-wave stimuli have been used to characterize linear systems such as a linear filtering operation. In order to demonstrate the validity of our model in this respect, we process a spatial cosine luminance waveform (Fig. 8). The result generates a spatial distribution of activity that corresponds to the initial cosine modulation and is consistent with that predicted for a simple cell.

This demonstration shows that our circuit, although intrinsically non-linear, shows significant linear behavior especially demonstrated for test cases similar to experimental studies with sinusoidal luminance modulations.

4.3 Initial measurement and grouping

Glass patterns (Glass 1969) illustrate how the visual system is capable of employing local information to generate global structure. The perception of structure generated by a Glass pattern (or a variation of it) is based on two types of interaction: a local operation of feature measurement or token extraction and a long-range integration of local items (Sagi et al. 1993). A test of the initial mechanisms underlying Glass pattern perception was provided by a recent investigation by Brookes and Stevens (1991). In their study, they generated Glass patterns using oriented white dot pair items which were either radially or circularly oriented. Due to the contrast sensitivity of oriented cells, the authors hypothesized that the inclusion of an opposite polarity dark dot between the two white dots of each local pair (''dot-between'' arrangement) would disrupt the perception of the structure seen without the distractor. In particular, the experiment was designed such that the disruption of a radial organization caused the pattern to be more likely perceived as circular. In contrast, adding a black dot to

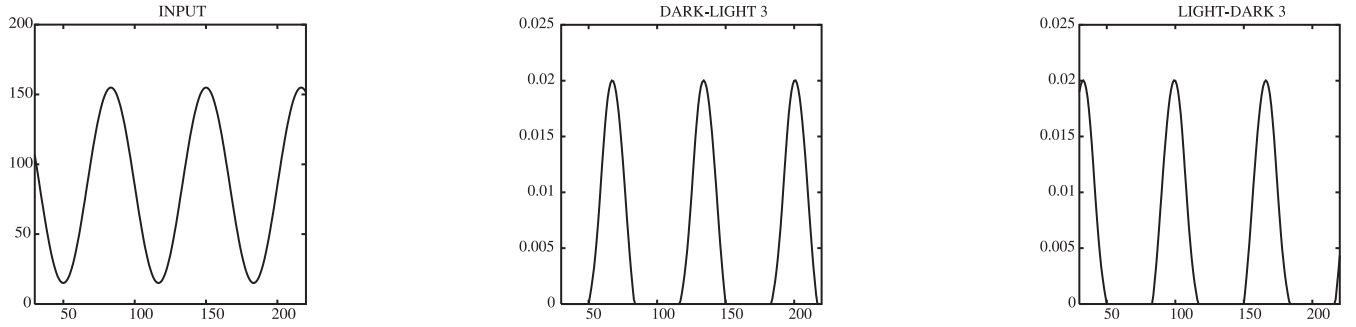

Fig. 8. Processing results of the non-linear circuit generated for a one-dimensional spatial luminance sine wave. The input signal is shown on the left, the responses for two simple cells sensitive to opposite contrast (dark-light, light-dark) are shown in the *middle* and in the *right panel*, respectively. As predicted for a pure linear cell, symmetric responses occur for opposite contrast sensitivities at the positive and negative slopes of the sine wave

disrupt a circular organization made by white dot pair items increases the tendency to perceive a radial organization. As a control, the "dot-between" case was modified so that the disrupting black dot was placed beside one of the white dots ("dot-adjacent" case). Figure 9 shows the Glass patterns that are composed by the local arrangements of Fig. 10.

Brookes and Stevens (1991) reasoned that if the perception of structure intimately depends on localized mechanisms sensitive to local contrast direction, then the perceived structure should be lost since the local oriented "dot-between" arrangement largely disrupts responses for the orientation of the white dot pair item. Even more, in the "dot-between" case, a strong response in the orientation orthogonal to the item is predicted while the "dot-adjacent" case should have no disturbing effect (it may be even more supportive). The perceptual results confirmed this prediction. In an alternative-forced-choice judgement, there was a significant increase in the subjects' responses to orthogonal apparent organization (reversing concentric and radial) when the black dot was placed in the middle between the pairs of white dots. The placement of the black dot adjacent to a white dot had no influence on the correct judgement of perceptual organization.

The experimental setting described above demonstrates that small differences in the local arrangement of structure can cause categorical changes in global structure. Such behavior is suggestive of mechanisms that are sensitive to precise contrast arrangements. Motivated by this observation, we stimulated our circuit with the local dot arrangements shown in Fig. 10 and compared the results with those generated by a linear model. Each pattern was pre-processed by the initial stage of center-surround interaction, segregating representations of ON and OFF contrast. Figure 11 displays the results. In order to gain more insight into the orientation specificity of the generated activations, we produced "needle" diagrams to encode individual response magnitudes for the different orientations. The certainty of orientation selectivity serves as an indicator for the clarity of subsequent long-range integration mechanisms for grouping [Sagi et al. (1993); see computational mechanisms by, for example, Grossberg and Mingolla (1985)]. For the white dot pair item (Fig. 11, left), the non-linear response shows a clear dominance along the orientation of the dot

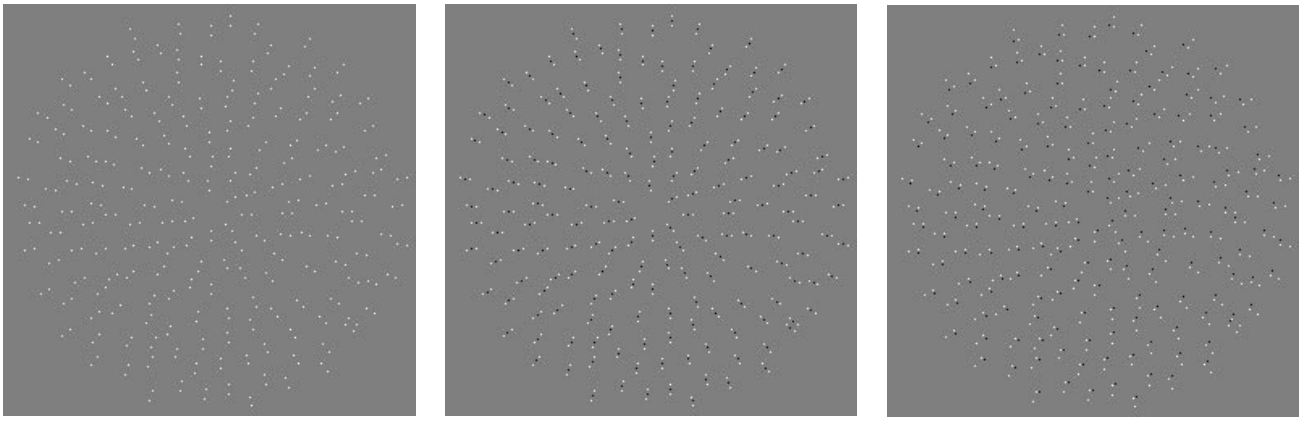

Fig. 9. Displays of the different variants of Glass patterns from the study of Brookes and Stevens (1991). Patterns with white dot-pair items, "dot-between" and "dot-adjacent" are shown from left to right. Each item in the radially oriented patterns is composed of local pairs of equal polarity and local triples of mixed polarity dots (see figure below). Due to reduced resolution and additional discretization, the perceptual effect may be significantly reduced. However, in our original displays and the plates shown by Brookes and Stevens (1991), the variations produce significant effects

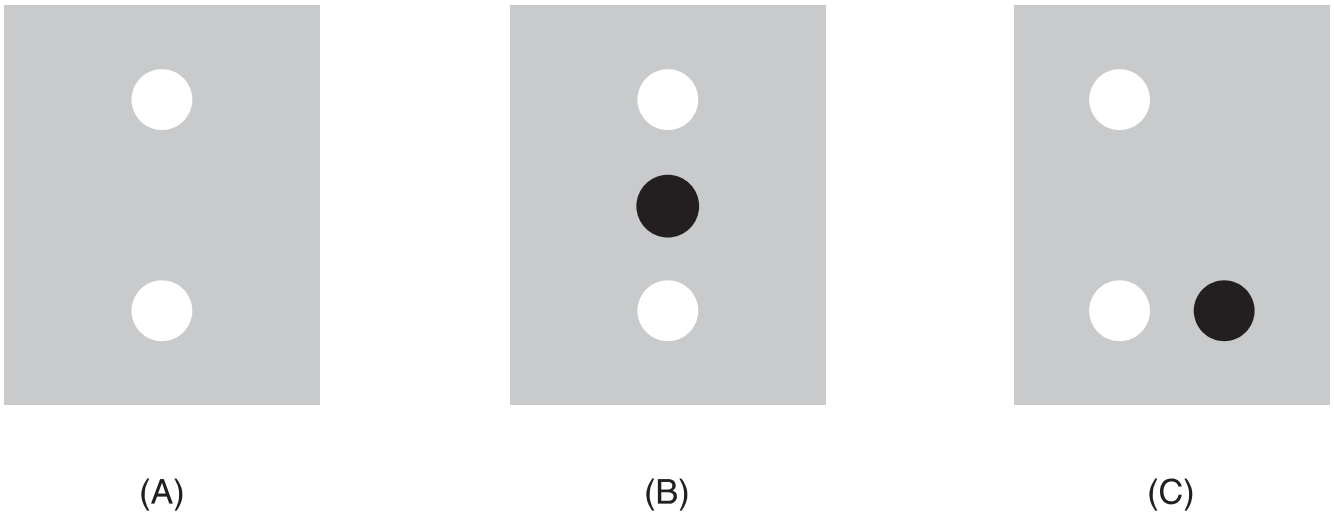

Fig. 10A–C. Local dot arrangements used in the Brookes and Stevens (1991) experiments to demonstrate categorical changes in the appearance of Glass patterns. **A** A local pair of white dots is shown as it appears in the original version of a Glass pattern. Stimulus variants were obtained by including a third dot of opposite polarity (*black*). The dot was either placed **B** in between the two white dots (''dot-between''), or **C** beside one of the white dots (''dot-adjacent'')

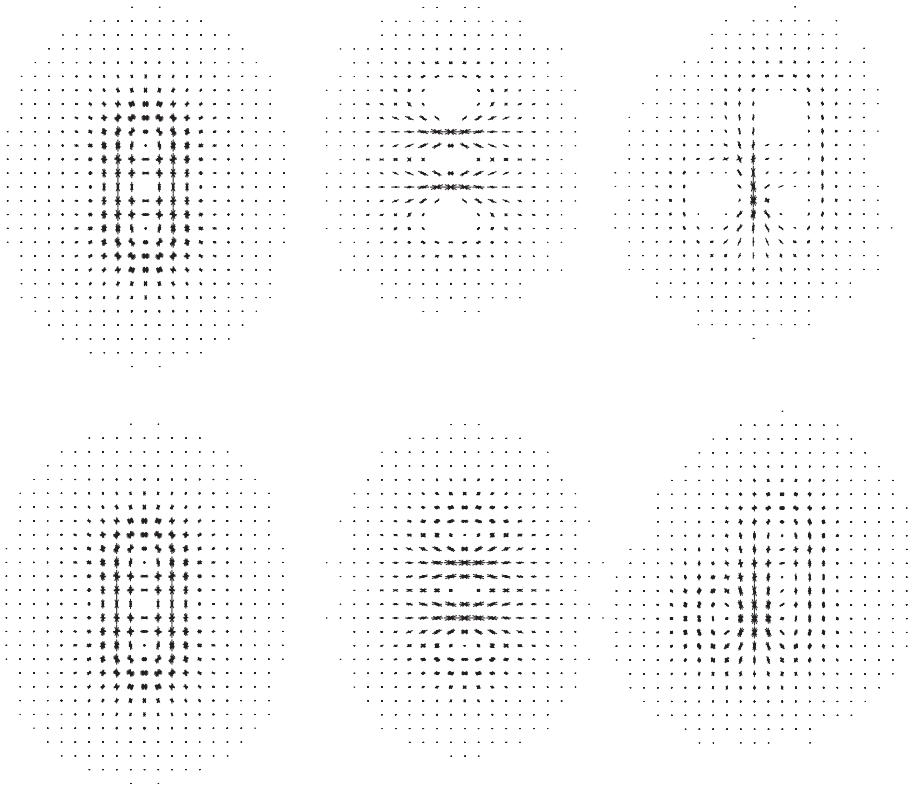

Fig. 11. Results of processing the local items used by Brookes and Stevens (1991) to generate the Glass pattern variants. Results of processing for the three items (white dot pair, ''dot-between'', ''dot-adjacent'') are shown for the non-linear circuit (*top row*) and for the linear model (*bottom row*). For a better visualization of individual contributions for different orientations, we show ''needle'' diagrams in which the activities for the spatial orientations are plotted at each location. The length of each oriented needle encodes the relative magnitude of response (the magnitudes are normalized with respect to the maximum response that appears in the display). The results indicate the individual dominances in orientation response that correspond to that of the perceptual appearance of the variants of Glass stimuli (see text)

pair. The linear model also shows a certain preference in that direction but the response is much more blurred and the orientations are much more uncertain. For the ''dot-between'' stimulus (Fig. 11, middle), the non-linear circuit generates a categorically different pattern with a strong orthogonal contrast orientation. This coincides with the observation in psychophysical experiments that the correct organization is disrupted and a perception of the orthogonal orientation appears. In the ''dot-adjacent'' case (Fig. 11), the correct organization is not largely disturbed and the orientation is mostly along the same polarity dot pair; between the pair of white and black dots, contrast responses of high amplitude are generated which support the overall orientation preference. In comparison to the result of the white dot pair alone, the spatial distribution of responses appears more blurred. We suggest that the increased blurring weakens the response in the subsequent stage of integration for grouping and perceptual organization. The responses generated by the linear model also show preferences

corresponding to those of the non-linear circuit, although weaker. In addition, as in the dot pair item, the distribution appears much more blurred and more uncertain in orientation space.

5 Discussion

We have introduced a circuit for non-linear contrast detection. The development was motivated by an analysis of the process of local structure measurement within luminance distributions. The proposed mechanism incorporates the local detection of contrast by measuring the amplitude of luminance changes, while at the same time accounting for the sharpness of any such change. The importance of the latter functionality was motivated by our earlier investigations on brightness perception (Pessoa et al. 1995). However, as discussed in Sect. 1, such functionality is also important for robust contrast estimation in Retinex-like algorithms, where shallow gradients need to be discounted. In fact, we claim that processes of contrast measurement need to be considered from a broader perspective, one in which issues such as the determination of surface qualities, ratio processing, to name a couple, are adequately taken into account. Mechanisms only based on contrast amplitude measurements may be too limited to account for general processing needs.

The circuit for contrast detection produces non-zero responses whenever ON or OFF responses from initial center-surround processing are present. This allows the circuit to signal shallow luminance gradients as in the case of ramps or harmonic modulations. The sharpness of a luminance transition, for example, at a luminance edge, is encoded by the spatial juxtaposition of such ON and OFF responses. Relative to the sampling scale of the proposed mechanism, the circuit registers any juxtaposed input configuration producing an extra strong response component with respect to the individual linear contributions.

The response properties of the circuit were embedded within the framework of scale-space processing, in particular the mechanism of scale selection. With reference to linear models of contrast processing based, for instance, on Gaussian derivatives, we have specified conditions for the generation of absolute maximum scale-space responses and verified these results through specific computer experiments. We suggest that the representation generated by the circuit is uniquely selective to intrinsic parameters of the luminance transition of a priori unknown width and height.

5.1 Relation to other proposals

Our scheme shares important features with other contrast measurement proposals including the well-known Marr and Hildreth (1980) edge detection proposal. Marr and Hildreth's main idea is that it is possible to detect edges by linking the outputs of ON and OFF center-surround cells (such as LGN cells) through a logical AND gate. Such a scheme is based on the idea that spatially adjacent ON and OFF responses represent the inflection point of a sharp transition – thus the zero-crossings of its second-order derivative indicate the edge transition (Poggio 1983). A closer inspection of the process of zero-crossing detection indicates that it utilizes an oriented multiplicative combination of initial ON/OFF responses based on a simple first-order difference mechanism. The zero-crossing detection can be formalized using our previous notation used in Sect. 2. A LD (DL) transition is detected via

$$z^{\mathrm{LD}} = \max\left[c^{+}_{i,j}\cdot c^{-}_{i+1,j}, c^{+}_{i,j}\cdot c^{-}_{i,j+1}\right] \quad \text{and}$$
$$z^{\mathrm{DL}} = \max\left[c^{-}_{i,j}\cdot c^{+}_{i+1,j}, c^{-}_{i,j}\cdot c^{+}_{i,j+1}\right] \tag{22}$$

(compare with Eqs. 3, 4 and 11).

While our model provides a circuit capable of realizing an AND-gate-like behavior, it does not compute a logical AND gate, but instead a soft gate. This property is one of the key differences between the present proposal and the Marr and Hildreth method. Although important, edge detection should not be the sole aim of contrast measurement. As stated above, what is needed is a general, robust method of contrast measurement that incorporates not only physical contrast assessment, but also other factors, such as transition form. In addition, image contours are present not only at edges but at other luminance distributions [see, for example, Fig. 8 of Pessoa (1996)]. Moreover, utilizing a numerical difference scheme, simple Fourier interpretation renders such a mechanism highly sensitive to noise. This is not the case in our model as we use Gaussian low-pass filters before sampling ON and OFF inputs.

The analysis of the differential structure of image curves and contrast outlines led Zucker and colleagues (Dobbins et al. 1990; Iverson and Zucker 1995) to propose a syntactic scheme defining a language of logical/linear operators. In this formalization, operators composed of tangentially separated sub-fields were designed to selectively respond to contrast (''edge'') and line features while enabling the operators to automatically suppress ''false'' responses. The characterizing feature of Zucker's logical/linear scheme is that it selectively suppresses any response that is generated by a luminance structure to which a contrast or line operator does not fit. Yet, it generates the responses in a linear fashion – but only at locations that match the filter's structure. In our case, the circuit signals the presence of a relevant ''sub-optimal'' structure, i.e., a structure that does not match the spatial arrangement the circuit is most sensitive to. In the case of optimal structures (e.g., step transitions), however, the presence of such a sharp contrast is signaled by an extra amount of high amplitude response that raises the activation over that of the linear response alone. This allows the decision as to what constitutes significant structure to be postponed, allowing more global stages of visual integration to selectively focus and enhance the visual structure that initially may not appear as dominant (in terms of amplitude response). In all, the non-linearities in Zucker's and in our

scheme differ: in Zucker's approach, the non-linearity functions as an early local decision mechanism, while in our model it serves as an emphasizing, or "boosting," mechanism to generate a high amplitude or energy response.

While other computational models have not explicitly discussed any scale-space-related processing properties, our scheme nicely generalizes to the multi-scale framework of edge detection. We have demonstrated how the non-linearity in the response to juxtaposed ON and OFF input can be utilized to signal the optimal operator scale for a gradual luminance transition of a certain spatial extent. The multi-scale scheme approximates a Gaussian first-order derivative scheme that makes use of the non-linearity in a scale-dependent input combination of adjacent contrast arrangement.

5.2 Simple cell physiology

Originating from Hubel and Wiesel's proposal (1962), a long-standing view of simple cell response is that it depends on the linear sum of ON and OFF LGN signals (for a review see von der Heydt 1987). Simple cells linearly sum (or pool) all of their inputs – weighted by their effective strength which is defined by the RF response profile. Data supporting the view that simple cells behave as linear devices come from a study by Schumer and Movshon (1984) showing that simple cells integrate their inputs in a sum-to-threshold linear manner. In addition, stimulus-response measurements of cells using spatial frequency modulated luminance gratings reveal a basically linear relationship whose sensitivity profile can be closely approximated by, among others, Gabor/Wavelet profiles generated by a Gaussian weighted sinusoidal spatial modulation (Pollen and Ronner 1983; see Daugman 1985, for a mathematical description), derivative of Gaussian profiles, or offset Gaussian profiles (Heggelund et al. 1983).

Non-linearities utilized in several computational schemes, like ours, have also been observed in cortical simple cell behavior, thus indicating that the view of such cells as linear devices is, in general, untenable (see von der Heydt 1987). For example, Hammond and MacKay (1983) investigated the length-summation characteristics and probed cells with bars composed of opposite polarity segments (i.e., light and dark segments). An OFF sub-field was probed by a dark bar that included two light ends, and the same was done for the ON sub-field using light bars with dark segments at their ends. The inclusion of inverse-polarity segments largely suppressed cell responses instead of only causing a reduction of response that is proportional to the length of the opposite polarity segments used – as predicted by linearity. A more recent investigation by von der Heydt et al. (1991) has demonstrated the existence in areas V1 and V2 of cells selectively responsive to oriented high spatial frequency gratings. Cells of this type vividly respond to gratings composed of alternating light and dark bars while remaining silent when stimulated by isolated bars of either contrast polarity. This has been interpreted as evidence for non-linear sub-field integration, since the responses could not be reconciled with a linear filter model responsive to gratings of the given spatial frequency. The responses were critically dependent on the precise stimulus periodicity, suggesting that the cells integrate spatially aligned arrangements of alternating ON and OFF LGN input in a non-linear way [see von der Heydt (1987) for an overview and discussion of previous findings in that direction, and Petkov and Kruizinga (1997) for an approach to model the key behavior].

These and several other observations suggest that the view of cortical contrast selective cells as being basically linear devices is too limited. In addition, there is growing evidence that they are sensitive to more than just local contrast magnitude. We suggest that an analysis based on a broader view of the functionality of early stages of visual processing may guide the development of models of visual contrast measurement.

Acknowledgements. We are grateful to the anonymous reviewers for their critical comments that helped to improve the manuscript. The collaborative research is supported by the German-Brazilian Academic Research Collaboration Program (DAAD-CAPES, Probral 1997, Project No. 70712). The research work by H.N. and T.H. is performed in the Collaborative Research Center on the "Integration of symbolic and subsymbolic information processing in adaptive sensory-motor systems" (SFB-527), which is located at the University of Ulm and is funded by the German Science Foundation (DFG). L.P. is also supported in part by a CNPq/Brazil grant (520419/96-0).

Appendix 1: A linear Gaussian edge detector in scale-space

This section draws upon previous work by Neumann and Ottenberg (1992a,b) who investigated the optimum scale-space response for one-dimensional odd-symmetric intensity transitions based on a linear pre-processing stage. In particular, we utilize here the cascade processing of the initial center-surround processing stage followed by an operation that realizes a low-pass filtered first-order derivative. We use an oriented gauge coordinate system (u, v) as introduced above in order to express the derivative attributes in terms of matched directional derivative operations. We utilize this framework to relate the findings on the evolution of scale-space response to the concept of automatic scale selection, which has only recently been suggested by Lindeberg (1996).[4] It should be noted that for the analysis of the scale-space response, we only analyze here the numerator term of the normalized response. Since we further investigate the relation between linear and non-linear responses in relation to their scale-selection properties (see Appendix 2), the denominator component for normalization was omitted here. We start with the analysis of the response to a step edge, then proceed to the more general case of a ramp transition.

Response for a step edge luminance transition

A step edge profile of height h is defined by an amplitude-scaled Heaviside function $\mathscr{H}(.)$, thus having

[4] Thanks to Maria Quiteria (Inst. of Oceanographic Studies, Rio de Janeiro) for providing some of the insights and a fresh view on recent modeling approaches.

$$f_{\text{step}}(u) = h \cdot \mathscr{H}(u) \ .$$

The model consists of a sequence of principal processing stages. The first step consists of a center-surround mechanism defining an isotropic band-pass filtering operation. We approximate the (normalized) difference-of-Gaussian by a Laplacian-of-Gaussian (LoG) operation. The response is generated by the spatial convolution of the step profile with the impulse response of the LoG filter, $f_{\text{step}}(u) \otimes \text{LoG}_\sigma(u)$. We get[5]

$$\text{LoG}_{\text{step}}(u) = h \frac{d}{du} \text{G}_\sigma(u) \ .$$

The LoG-filtered profile is processed subsequently by a first-order (Gaussian) derivative ($D1$G) filter. At its limit, for $\sigma \to 0$, the Gaussian derivative will converge to $d/du\ \delta(u)$. The final spatial profile of contrast cell responses for its minimum spatial scale limits is given by

$$s_{\text{step}}(u) = h \cdot \frac{d^2}{du^2} \text{G}_\sigma(u) = \frac{h}{\sigma^2} \left[1 - \frac{u^2}{\sigma^2} \right] \text{G}_\sigma(u) \ .$$

The response amplitude at the step edge location, $u = 0$, for a normalized Gaussian weighting function is then

$$s_{\text{step}}(u = 0) = \left| \frac{h}{\sigma^2} \ \text{G}_\sigma(u) \right| = \left| \frac{h}{\sqrt{2\pi} \cdot \sigma^3} \right| \ .$$

The result shows that the response taken as a function of scale σ is strong monotonically decreasing such that the maximum response appears for the smallest scale.

Response for a ramp edge luminance transition

A ramp edge profile of width R and height h can be generated by the convolution of an amplitude-scaled Heaviside function (see above) with a unit-pulse function $\Pi(.)$ of the corresponding width. The profile is thus described analytically as

$$f_{\text{ramp}}(u) = h \cdot \mathscr{H}(u) \otimes \Pi_R(u) \ .$$

The response of the model is generated by the spatial convolution of the ramp profile with the impulse response of the LoG-filter, $f_{\text{ramp}}(u) \otimes \text{LoG}_\sigma(u)$. We get

$$\text{LoG}_{\text{ramp}}(u) = \frac{h}{R} \left[\text{G}_\sigma\left(u + \frac{R}{2} \right) - \text{G}_\sigma\left(u - \frac{R}{2} \right) \right] \ .$$

Corresponding to the procedure we have adopted for the step case, further processing of the LoG-filtered profile by a first order (Gaussian) derivative ($D1$G) filter yields

$$s_{\text{ramp}}(u) = \frac{h}{R} \left[\frac{d}{du} \text{G}_\sigma\left(u + \frac{R}{2} \right) - \frac{d}{du} \text{G}_\sigma\left(u - \frac{R}{2} \right) \right] \ .$$

The response consists of a sum of two Gaussian derivative profiles of opposite sign shifted to either side of the central location of the ramp transition. For completeness, we can confirm the result to include the special case of a step transition. By taking the limit $R \to 0$, the response $s_{\text{ramp}}(u)$ converges to that derived for a step, $\lim_{R \to 0} s_{\text{ramp}}(u) = s_{\text{step}}(u)$ (see above).

[5] For simplicity, we are neglecting the fact that, for this 2D isotropic pre-filtering stage, the effective space constant is $\sigma_e = \sqrt{2}\ \sigma$. Since this refers only to a rescaling of coordinate axes, we further deal with the original scale constant σ only.

Scale-space processing and responses at ramp edge location

Depending on the width of the ramp and the relative scaling of the cascaded pre-processing stage, the response profile appears in different spatial arrangement. We consider a case that has been sketched in Fig. 4. For a scale that is small relative to the transition width ($\sigma \ll R$), Gaussian derivatives appear as isolated ON and OFF channel responses at the "knees" where the plateaus meet the ramp transition. For a coarser scale after blurring, the central lobes of either ON or OFF responses (depending on a light-dark or a dark-light transition, respectively) meet and overlap at the center of the ramp and appear as juxtaposed input contrast responses. Analytically, for a condition $R \ll \sigma$ (corresponding to $R \to 0$), we will get the second-order derivative of a Gaussian profile that peaks at the center of the ramp (compare with the results of mathematical investigation above and see Fig. 6).

In order to evaluate the behavior of a scale-space edge detector for a gradual luminance transition, we evaluate its response at the center of the ramp ($u = 0$). We make use of the even-symmetry of the Gaussian, such that we finally get the amplitude

$$s_{\text{ramp}}(u = 0) = 2 \frac{h}{R} \cdot \text{G}'_\sigma\left(\frac{R}{2} \right) \ .$$

We now treat s_{ramp} as a function of scale σ such that we derive the magnitude function

$$s_{\text{ramp}}(\sigma)|_{u=0} = \frac{h}{\sqrt{2\pi}\sigma^3} \exp\left[-\frac{R^2}{8\sigma^2} \right] \ .$$

Figure 12 shows profiles of scale-space responses of a linear cell for variable parameter settings of ramp transition widths R and constant height $h = 1$. The responses for a ramp transition shown in Sect. 4 (Fig. 6, bottom) represent a coarse sample of the analytic profile presented in Fig. 12 (right). For a mechanism capable of selecting an appropriate scale it needs a proper selection and decision criterion. A useful candidate for this selection appears to be the maximum scale response along $u = 0$. The unique peak response along the scale is determined by the zero-crossing of the slope function for $s_{\text{ramp}}(\sigma)|_{u=0}$. We get

$$\frac{d}{d\sigma} s_{\text{ramp}}(\sigma) = -\frac{h}{\sqrt{2\pi}\sigma^4} \left(\frac{R^2}{4\sigma^2} - 3 \right) \exp\left[-\frac{R^2}{8\sigma^2} \right] = 0 \ .$$

The solution for the scale-space location of the peak response is given by

$$\sigma_{\max} = \frac{R}{2\sqrt{3}} \ ,$$

which indicates an operator scaling slightly less than half the ramp transition width. A mechanism reading out such maximum responses must rely on a robust selection criterion that guarantees the correspondence with a representation of variable-width contrast edges. We investigate this topic below in terms of absolute maximum contrast responses suggesting a winner-take-all type mechanism.

Scale selection for a ramp contrast

In order to investigate the relevant conditions for selecting maximal responses corresponding to a ramp of certain width, we compare the activations generated for $s_{\text{ramp}}(\sigma_{\max})|_{u=0} \equiv s_{\text{ramp}}(u = 0)|_{\sigma=\sigma_{\max}}$ with the absolute height of one lobe in the profile of $D1$G. In particular, we investigate the response for $s_{\text{ramp}}(u = -R/2 + \sigma_0)$. The latter corresponds to the case of isolated contrast responses at the "knees" of the ramp that are processed by the $D1$G filter. For the condition $\sigma_0 \ll R$, the response profile in the u direction shows

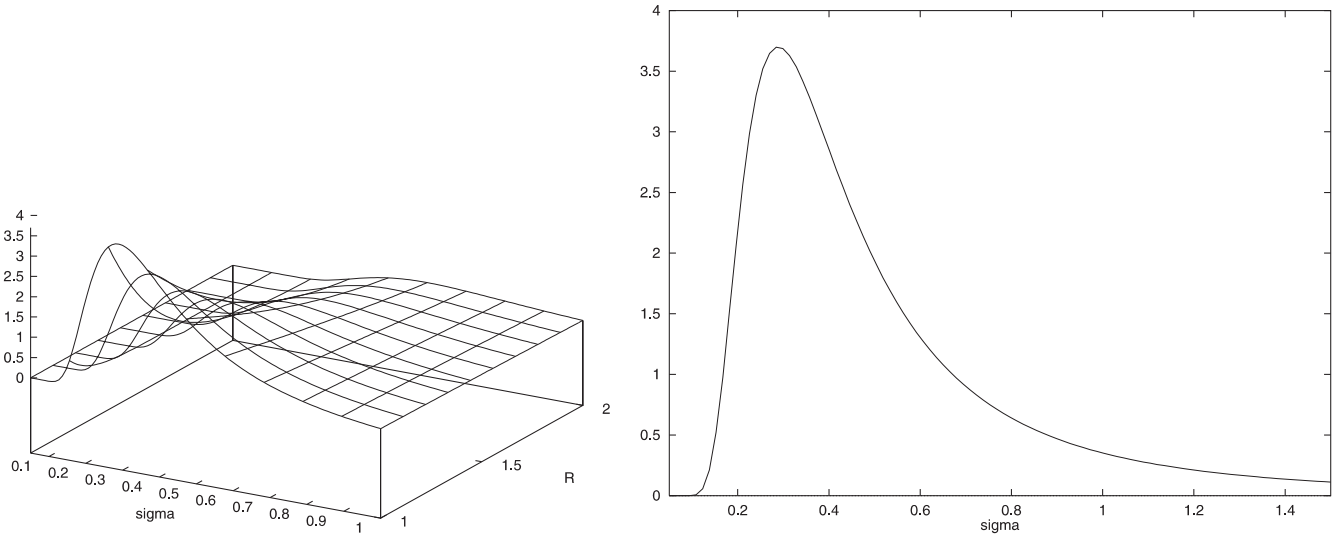

Fig. 12. Linear contrast cell response at ramp location $u = 0$ treated as a function of scale (σ). Processing results of a family of ramp profiles of unit height ($h = 1$) is shown as a surface for different widths R (*left*); the shape of the response profile for one fixed setting of R shows a unique maximum in response as a function of scale (*right*). As it is predicted by the equation for $s_{\text{ramp}}(\sigma)|_{u=0}$, a unique shape of scale-space profiles is shifted in peak location for increasing values of R accompanied by a drop in overall response amplitude

$$s_{\text{ramp}}(u = -R/2 + \sigma_0)|_{\sigma_0 \ll R}$$
$$= \frac{h}{R}\left[\frac{d}{du}\mathrm{G}_\sigma(\sigma_0) - \underbrace{\frac{d}{du}\mathrm{G}_\sigma(-R + \sigma_0)}_{\approx 0}\right]$$

(see Fig. 6 for comparison). Since the second derivative term almost vanishes, we can simplify the equations by only considering the dominant component for further analysis. This condition corresponds to a situation where $\sigma_0 < \sigma_{\max}$ holds. If we compare the responses of the two selected cases we get

$$s_{\text{ramp}}(u = 0)|_{\sigma=\sigma_{\max}} = \frac{48\sqrt{3}h}{\sqrt{2\pi}}\exp\left[-\frac{3}{2}\right]\cdot\frac{1}{R^3}$$
$$s_{\text{ramp}}(u = -R/2 + \sigma_0)|_{\sigma_0 \ll R} \approx \frac{h}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}\right]\cdot\frac{1}{R\sigma^2}$$

With the ratio of these two expressions, we get a condition for absolute maximum response. For this purpose we introduce the normalized scale-space variable $\hat{\sigma} = \sigma/R$. The condition then finally reads

$$\frac{48\sqrt{3}}{e}\hat{\sigma}^2 > 1 \qquad \text{or, equivalently}$$
$$\hat{\sigma} > \sqrt{\frac{e}{48\sqrt{3}}} \approx 0.18, \qquad \text{with } e = \exp[1]\ .$$

We left the first version of the inequality for direct comparison with the result derived for the non-linear circuit (Appendix 2).

The result for the linear model demonstrates that the identified optimum scale is not guaranteed to yield maximum responses in absolute terms as related to the overall field of activities. In particular, if $\hat{\sigma}$ becomes less than approximately 1/5, the response $s_{\text{ramp}}(u = 0)|_{\sigma=\sigma_{\max}}$ declines to a value lower than the maximum response for the smallest scale, as demonstrated in Fig. 6 (bottom). If, however, the necessary condition is met, the maximum scale response determines the overall maximum amplitude in the array of scale-space filter responses.

Appendix 2: Scale-space response for the non-linear circuit

It has been shown in Appendix 1 that the linear model already shows a selectivity for an optimal scale. The response taken as a function of scale $S(\sigma)$ yields a unique maximum value, pre-requisite for automatic scale selection. However, it was shown that this value cannot be guaranteed to be the maximum response in absolute terms – as also demonstrated by the simulations shown in Fig. 6.

A compact notation of the response properties of the non-linear circuit is given by

$$z = \mathscr{L}\cdot\left[(p^+ + p^-) + \frac{\mathscr{N}}{\mathscr{L}}\cdot(p^+ \cdot p^-)\right]$$

(compare with Eq. 12). This demonstrates that the non-linear circuit in its linear processing behavior inherits all the qualitative properties shown in Appendix 1. Moreover, for configurations approaching the optimum scale conditions, i.e., configurations with juxtaposed ON and OFF activations, the results of the linear processing component are superimposed by the correlation (gating) between ON and OFF responses.

The function $\mathscr{L}$ only defines a proportionality for both linear and non-linear components. For any further analysis of the response z, it is therefore sufficient to consider the terms in large brackets (see equation above). The non-linear contribution is scaled by the factor $\mathscr{N}/\mathscr{L}$. Since we want to relate the cell response to the activations generated by the initial center-surround processing stage, we have to scale this ratio accordingly. Based on the dependency described in Eq. (3), we arrive at

$$\mathscr{N}' \equiv \frac{\mathscr{N}}{\mathscr{L}}\frac{\beta_s}{\alpha_s} = 2\frac{\beta_c}{\alpha_c}\frac{\beta_s}{\alpha_s}\ ,$$

in which β_s/α_s is the gain of the initial center-surround processing stage, denoted by $\text{net}^{\pm} - \text{net}^{\mp}$.

Non-linear response for step edges

In Appendix 1, we have analyzed the responses of a linear mechanism processing a step edge profile. The non-linear part of the total response is contributed by $(p^+ \cdot p^-)$. We determine the magnitudes from the maxima in the $D1G$ profile. These correspond to the offset points for sub-field integration given by $\tau = \pm\sigma$. For the profile of a first-derivative of a Gaussian we get

$$h\frac{d}{du}\,\mathrm{G}_\sigma(u=\sigma) = -\frac{h}{\sqrt{2\pi}\sigma^2}\exp\left[-\frac{1}{2}\right] \ .$$

We introduce the functions $s^{\mathrm{lin}}_{\mathrm{step}} \equiv s_{\mathrm{step}}$ (see Appendix 1) and $s^{\mathrm{nl}}_{\mathrm{step}}$ and show the non-linear response in terms of the linear response derived in Appendix 1. The significant part of a step edge response (proportion of response z given in brackets, see above) generated by the non-linear circuit is then given by

$$\begin{aligned} s^{\mathrm{nl}}_{\mathrm{step}}(u=0) &= \frac{h}{\sqrt{2\pi}\sigma^3} + \mathscr{N}'\frac{1}{\sqrt{2\pi}e\sigma}, \\ &= s^{\mathrm{lin}}_{\mathrm{step}}(u=0)\cdot\left(1+\mathscr{N}'\frac{1}{\sqrt{2\pi}e\sigma}\right) \ . \end{aligned}$$

Two observations can be made from this result: (1) as in the linear case, the response for a step edge is a monotonically decreasing function of scale in which the maximum response is generated at the smallest scale, and (2) the contribution of the non-linearity dominates for a gain factor

$$\mathscr{N}' \gg \sqrt{2\pi}e\sigma \ .$$

Optimum input conditions for non-linear processing

Non-zero correlation between ON and OFF channel activation is available when both sub-fields of a cell integrate positive input activation. The initial center-surround responses at the "knees" of the ramp (see Appendix 1) therefore have to be progressively blurred to meet at the center of the ramp. For each Gaussian bump in the initial $\mathrm{LoG}_{\mathrm{ramp}}(u)$ profile we get

$$\frac{h}{R}\left[\pm\mathrm{G}_{\sigma_0}\left(u\pm\frac{R}{2}\right)\ \otimes\ \mathrm{G}_\sigma(u)\right] = \pm\frac{h}{R}\ \mathrm{G}_{\sqrt{\sigma_0^2+\sigma^2}}\left(u\pm\frac{R}{2}\right) \ .$$

Proper scaling of the blurred profile approaches the value determined for the optimum scale, such that $\sqrt{\sigma_0^2+\sigma^2} = \sigma_{\max}$. The profile showing the difference of two broadly tuned offset Gaussians assembles a $D1G$ filter.

Input integration and maximum scale response for ramp edges

Since offsets were symmetric around the center of the ramp, we sample each shifted Gaussian $\mathrm{G}_{\sigma_{\max}}(u \pm R/2)$ at $u = \pm\tau = \pm R/2$. For the gating-type input to the non-linear response we therefore get

$$(p^+ \cdot p^-)|_{\sigma=\sigma_{\max}} = \frac{h^2}{R^2}\cdot\frac{1}{2\pi\cdot\sigma^2_{\max}} = \frac{6h^2}{\pi\cdot R^4} \ ,$$

with $\sigma_{\max} = R/(2\sqrt{3})$. We can now write the response amplitude of the non-linear circuit at the optimum scale condition, $\sigma_{\max}$, as a combination of the linear (additive) and non-linear (correlational) input contribution. We introduce the functions $s^{\mathrm{lin}}_{\mathrm{ramp}} \equiv s_{\mathrm{ramp}}$ (see Appendix 1) and $s^{\mathrm{nl}}_{\mathrm{ramp}}$ and present the non-linear response in terms of the linear response derived in Appendix 1. We get

$$\begin{aligned} s^{\mathrm{nl}}_{\mathrm{ramp}}(u=0)|_{\sigma=\sigma_{\max}} &= \frac{48\sqrt{3}h}{\sqrt{2\pi}}\exp\left[-\frac{3}{2}\right]\cdot\frac{1}{R^3} + \mathscr{N}'\frac{6h^2}{\pi\cdot R^4}, \\ &= s^{\mathrm{lin}}_{\mathrm{ramp}}(u=0)|_{\sigma=\sigma_{\max}} \\ &\quad\times\left(1+\mathscr{N}'\frac{\sqrt{2}}{8\sqrt{3\pi}}\frac{h}{R}\exp\left[\frac{3}{2}\right]\right) \ . \end{aligned}$$

Several observations can be made from this result: (1) the response of the circuit for a ramp edge transition can be predicted for an optimum scale, (2) the contribution of the non-linearity depends on the slope of the ramp transition denoted as $m = h/R$, (3) the plain non-linearity of the circuit (without the gain $\mathscr{N}'$) contributes with a 26% increase – times the slope – in output magnitude. In the case of discrete signals and images, any transitions are of unit width at minimum, i.e., $R \geqslant 1$. For a ramp of unit slope, and the non-linear component to dominate the response, the gain factor $\mathscr{N}'$ has to be high enough to "boost" the output activation. Moreover, with such a gain factor it can be guaranteed that the magnitude of response is maximal in absolute terms (compare the analysis for a linear filter in Appendix 1). A dominance of the response caused by the non-linear, i.e., multiplicative, contribution from both ON and OFF subfield is achieved for

$$\mathscr{N}' \gg 4\sqrt{6\pi}\frac{R}{h}\exp\left[-\frac{3}{2}\right] \ .$$

Since the signal parameters h and R are not known in advance, the gain $\mathscr{N}'$ may be chosen high enough to guarantee the dominance of any non-linear contribution.

References

Arend L, Goldstein R (1987) Lightness models, gradient illusions, and curl. Percept Psychophys 42:65–80

Blake A (1985) Boundary conditions for lightness computation in Mondrian world. Comput Vis Graph Image Process 32:314–327

Bracewell RN (1978) The Fourier transform and its applications 2nd edn. McGraw-Hill, New York

Brookes RA, Stevens KA (1991) Symbolic grouping versus simple cell models. Biol Cybern 65:375–380

Buf J du (1993) Responses of simple cells: events, interferences, and ambiguities. Biol Cybern 68:321–333

Buf J du (1994) Ramp edges, Mach bands, and the functional significance of the simple cell assembly. Biol Cybern 69:449–461

Carandini M, Heeger DJ (1994) Summation and division by neurons in primate visual cortex. Science 264:1333–1336

Daugman J (1985) Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. J Opt Soc Am A 2:1160–1169

Dobbins AL, Iverson L, Zucker SW (1990) A logical/linear model of cortical subunit interaction. Proc ARVO'90 (Invest, Opthalmol Vis Sci 31: 397, Sarasota, Florida, USA, April 29 – May 4

Enroth-Cugell C, Robson JG (1966) The contrast sensitivity of retinal ganglion cells of the cat. J Physiol (Lond) 187:517–552

Ferster D (1988) Spatially opponent excitation and inhibition in simple cells of the cat visual cortex. J Neurosci 8:1172–1180

Ferster D (1989) The synaptic inputs to simple cells in the cat visual cortex. In: Lam D, Gilbert C (eds) Neural mechanisms of visual perception. Portfolio Publishing, The Woodlands, Tex, pp 63–85

Furman GG (1965) Comparison of for subtractive and shunting lateral-inhibition in receptor-neuron fields. Kybernetik 2:257–274

Glass L (1969) Moiré effect from random dots. Nature 223:578–580

Grossberg S (1970) Neural pattern discrimination. J Theor Biol 27:291–337

Grossberg S, Mingolla E (1985) Neural dynamics of form perception: boundary completion, illusory contours, and neon color spreading. Psychol Rev 92:173–211
Hammond P, MacKay D (1983) Influence of luminance gradient reversal on simple cells in feline striate cortex. J Physiol (Lond) 337:69–87
Heggelund P (1981) Receptive field organization of simple cells in cat striate cortex. Exp Brain Res 42:89–98
Heggelund P (1986) Quantitative studies of enhancement and suppression zones in the receptive field of simple cells in cat striate cortex. J Physiol (Lond) 373:293–310
Heggelund P, Krekling S, Skottun BC (1983) Spatial summation in the receptive field of simple cells in the cat striate cortex. Exp Brain Res 52:87–98
Heydt R von der (1987) Approaches to visual cortical function. Rev Physiol Biochem Pharmacol 108:69–150
Heydt von der R, Peterhans E, Dürsteler MR (1991) Grating cells in monkey visual cortex: Coding texture? In: Blum B (ed) Channels in the visual nervous system: neurophysiology, psychophysics and models. Freund, London, pp 53–73
Hodgkin AL (1964) The conduction of the nervous impulse. Liverpool University Press, Liverpool
Horn BKP (1974) Computing lightness from an image. Comput Graph Image Process 3:277–299
Hubel DH, Wiesel TN (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J Physiol 160:106–154
Iverson LA, Zucker SW (1995) Logical/linear operators for image curves. IEEE (Inst Electr Electron Eng) Trans Pattern Analysis and Machine Intelligence 17:982–996
Koenderink JJ (1984) The structure of images. Biol Cybern 50:363–370
Koenderink JJ, van Doorn A (1990) Receptive field families. Biol Cybern 63:291–297
Land EH, McCann JJ (1971) Lightness and retinex theory. J Opt Soc Am 61:1–11
Lindeberg T (1996) Scale-space: a framework for handling image structures at multiple scales. Lecture notes from the CERN school of computing 1996. Egmond aan Zee, The Netherlands, 8–21 September
Lindeberg T, ter Haar Romeny B (1994) Linear scale-space I: basic theory. In: ter Haar Romeny B (ed) Geometry-driven diffusion in computer vision, Freund, London, pp 1–38
Liu Z, Gaska JP, Jacobson LD, Pollen DA (1992) Interneuronal interaction between members of quadrature phase and antiphase pairs in the cat's visual cortex. Vision Res 32:1193–1198
Marr D (1982) Vision. Freeman, San Francisco
Marr D, Hildreth E (1980) Theory of edge detection. Proc R Soc Lond Ser B 207:187–217
Morrone MC, Burr DC (1988) Feature detection in human vision: a phase-dependent energy model. Proc Royal Soc Lond Ser B 235:221–245
Morrone MC, Burr DC, Maffei L (1982) Functional implications of cross-orientation inhibition of cortical visual cells. I. Neurophysiological evidence. Proc R Soc Lond Ser B 216:335–354
Neumann H (1996) Mechanisms of neural architecture for visual contrast and brightness perception. Neural Netw 9:921–936
Neumann H, Ottenberg K (1992a) Estimating ramp-edge attributes from scale-space. In: Vandevalle J, Boite R, Moonen M, Oosterlinck A (eds) Signal Processing VI: theories and applications vol I–III (6th Eur Signal Proc Conf ICPR-92), Brussels, Belgium. Elsevier, Amsterdam, pp 603–607
Neumann H, Ottenberg K (1992b) Estimating attributes of smooth signal transitions from scale-space. Proc 11th Int Conf Pattern Recog vol I–IV (ICPR-92), The Hague IEEE Press, Los Alamitos, California, pp 754–758
Neumann H, Pessoa L (1994) A simple cell model with multiple spatial frequency selectivity and linear/non-linear response properties. Proc of the World Congr on Neural Networks (WCNN-94), vol I–IV, San Diego, USA. LEA, Hillsdale, New Jersey, pp 290–298
Pessoa L (1996) Mach bands: how many models are possible? Recent experimental findings and modeling attempts. Vision Res 36:3205–3227
Pessoa L, Mingolla E, Neumann H (1995) A contrast- and luminance-driven multiscale network model of brightness perception. Vision Res 35:2201–2223
Petkov N, Kruizinga P (1997) Computational models of visual neurons specialized in the detection of periodic and aperiodic oriented visual stimuli: bar and grating cells. Biol Cybern 76:83–96
Poggio T (1983) Visual algorithms. In: Braddick OJ, Sleigh AC (eds) Physical and biological processing of images. (Springer series in information sciences, vol 11) Springer, Berlin Heidelberg Newy York, pp 128–153
Pollen DA, Ronner SF (1983) Visual cortical neurons as localized spatial frequency filters. IEEE (Inst Electr Electron Eng) Trans Systems, Man, and Cybernetics 13:907–916
Rodieck RW (1965) Quantitative analysis of cat retinal ganglion cell responses to visual stimuli. Vision Res 5:583–601
Sagi D, Kovács L, Eötvös L (1993) Long range processes involved in the perception of Glass patterns. Proc. ARVO'93 (Invest Ophthalmol Vision Sci 34. Sarasota, Florida, USA, May 2–7, p 1130
Schiller P (1992) The ON and OFF channels of the visual system. Trends Neurosci 15:86–91
Schumer RA, Movshon JA (1984) Length summation in simple cells in cat striate cortex. Vision Res 24:565–571
Tolhurst DJ, Dean AF (1990) The effects of contrast on the linearity of spatial summation of simple cells in the cat's striate cortex. Exp Brain Res 79:582–588
Vidyasagar TR, Pei X, Volgushev M (1996) Multiple mechanisms underlying the orientation selectivity of visual cortical neurones. Trends Neurosci 19:272–277
Witkin AP (1983) Scale-space filtering. Proc 8th Int Joint Conf Artif Intell, IJCAI-83. vol I–II, Karlsruhe, Morgan Kaufman Publ., Palo Alto, California pp 1019–1022