ESANN'1996 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 24-25-26 April 1996, D-Facto public., ISBN 2-9600049-6-5, pp. 193-198

Neural Model for Visual Contrast Detection

Enno Littmann*, Heiko Neumann*, and Luiz Pessoa†

*Abt. Neuroinformatik, Universität Ulm, D–89069 Ulm, Germany
†Universidade Federale, Rio de Janeiro, RJ 21945-970, Brazil

Abstract. Most approaches that model biological early vision systems perform at the cortical level of simple cells a *linear* integration of the activity from visual ON and OFF pathways. Based on empirical as well as theoretical investigations we propose a *nonlinear* neural network model that is selectively responsive to contrast magnitude as well as to the sharpness of luminance transition. The nonlinear circuit allows for accurate and reliable detection of contrast changes even in noisy images. Simulations with artificial and camera images show higher positional selectivity for local contrasts than an equivalent linear device. Furthermore, in a multiscale hierarchy the nonlinear circuit produces a unique maximum response in scale-space where scale directly relates to the width of the luminance transition.

1. Introduction

Processing of visual stimuli begins with the segregation of data streams selectively sensitive to light-on and light-off signals, respectively. These ON- and OFF channels are segregated at the retinal bipolar cell level (feeding ganglion cells of opposite antagonistic center-surround organization) and run separately to the first stages in primary visual cortex (V1). It is still unclear how these pathways are recombined at the cortical level of simple cells and what kind of functionality is precisely supported by the pooling of transferred activity.

An overview of empirical background and the discussion of simple cell functionality has already been given elsewhere [9]. Computational investigations have supported evidence of nonlinear combination of activity in ON and OFF subfields [6]. Within the broader context of unified contour and brightness perception ([4]), the need for bifurcating the response properties in a nonlinear feedback system has been justified [10].

In this paper, based on previous work of [9], we propose a neural circuit for nonlinear integration of ON and OFF data that is based on evidence derived from empirical investigation as well as functional modeling. Specifically, the model *i)* makes explicit local contrast changes of specific polarity, *ii)* shows linear and nonlinear response properties to make the processing architecture selectively responsive to stimulus features such as local contrast and sharpness (abruptness) of contrast change, and *iii)* incorporates multiple spatial frequency (scale) selectivity operating on a single scale input. The gating of juxtaposed activity from both ON and OFF pathways extends the functionality of local contrast "detection". In addition, a new feature dimension is incorporated that denotes the scale-dependent sharpness of local transitions. The nonlinear approach is compared to the associated first order linear model in which the input from excitatory and inhibitory subfields is combined by *linear* integration only.

ESANN'1996 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 24-25-26 April 1996, D-Facto public., ISBN 2-9600049-6-5, pp. 193-198

2. Model Circuit for Pooling ON and OFF Data Streams

Center-surround Antagonism and ON- and OFF-channels. The initial processing of the input luminance distribution utilizes isotropic antagonistic center-surround interaction of both polarity. The responses are transmitted by segregated ON and OFF channels. Local contrast information is generated by cross-channel inhibition, yielding activations $c_i^+ = [x_i^+ - x_i^-]^+$ and $c_i^- = [x_i^- - x_i^+]^+$ ($x_i^\pm$ denote ON- and OFF-channel activation of the model retina[1]; $[.]^+ = \max(., 0)$).

Simple Cell Responses. The input to subfields of simple cells is generated by blurring the activity distribution in the ON and OFF contrast channel with elongated Gaussian weighting functions. Multiple spatial frequency selectivity is achieved by using weighting functions of different spatial extent in length and width. According to [5], (blurred) activities in the segregated ON and OFF channels compete at each spatial position before they are integrated in simple cell subfields. This stage is modelled by computing $p_{i\varepsilon}^{S+} = [\sum_j (c_j^+ - c_j^-)\lambda_{ji\varepsilon}^S]^+$ and $p_{i\varepsilon}^{S-} = [\sum_j (c_j^- - c_j^+)\lambda_{ji\varepsilon}^S]^+$, where $\lambda_{ji\varepsilon}^S$ denotes an elongated Gaussian weighting function of scale S and orientation ε.

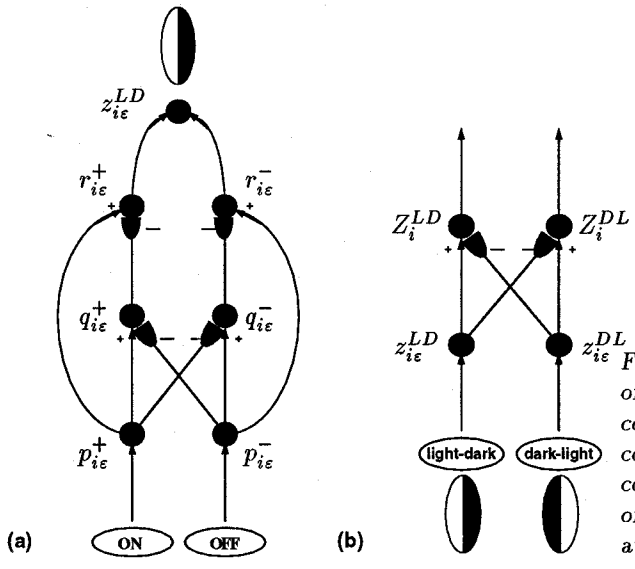

Fig. 1: (a) Proposed circuit of an odd-symmetric simple cell sensitive to light-dark contrast polarity, (b) final competition of simple cells of opposite contrast polarity at each spatial location.

Figure 1 (a) sketches the circuit of a simple cell with light-dark contrast sensitivity. The circuit itself has 4 stages and contains two streams (ON and OFF). The first stage receives the blurred contrast activities from (scale dependent) offset positions relative to the spatial reference location. The second stage contains an opponent interaction of the opposite channel activity as well as direct excita-

[1] The equilibrium activity is determined by shunting interaction, yielding e.g. an ON channel response $x_i^+ = (B\text{net}_i^+ - C\text{net}_i^-)/(A + \text{net}_i^+ + \text{net}_i^-)$. Thus the activity is always normalized with respect to the local average luminance resulting in an asymmetry of ON and OFF channel activity for odd-symmetric intensity profiles.

ESANN'1996 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 24-25-26 April 1996, D-Facto public., ISBN 2-9600049-6-5, pp. 193-198

tory inputs from each subfield. The third receives channel-specific inputs from both the first (excitatory) and second (inhibitory) stages. Finally, ON and OFF channels are pooled producing the final response in stage four. The opponent inhibition associated with the within-channel inhibition provides a mechanism for disinhibition. As a consequence, the output of the circuit will be large only when the inputs from both channels are strong, thus allowing for disinhibition of activation in each subfield.

The model employs both light-dark and dark-light odd-symmetric simple cells. These are obtained by collecting contrast information from spatially different branches. For a light-dark cell at position i, ON information originates from the "left" and OFF information from the "right" relative to the symmetry axis of the local reference coordinate system (rotation angle ε; "left" and "right" spatial offsets denoted by $(^+) = i - \tau_S$ and $(^-) = i + \tau_S$, τ_S scale-dependent constant). Simple cell responses are computed in two steps, with intermediate variables $q_{i\varepsilon}^+$ (stage 2) and $r_{i\varepsilon}^+$ (stage 3) for the ON subfield and $q_{i\varepsilon}^-$ and $r_{i\varepsilon}^-$ for the OFF subfield (Fig.1 (a)). The equations for the ON channel activations are

$$\dot{q}_{i\varepsilon}^+ = -\alpha q_{i\varepsilon}^+ + p_{i\varepsilon}^+ - \beta q_{i\varepsilon}^+ p_{i\varepsilon}^- \quad \text{and} \quad \dot{r}_{i\varepsilon}^+ = -\gamma r_{i\varepsilon}^+ + p_{i\varepsilon}^+ - \delta r_{i\varepsilon}^+ q_{i\varepsilon}^+, \qquad (1)$$

where α, β, γ, and δ are constants (OFF channel activations are obtained by exchanging '+' and '−' indices). In order to achieve a symmetry in the response properties of both channels the constants must obey the identity $\delta = \beta\gamma$. The output response of a light-dark simple cell is computed by $\dot{z}_{i\varepsilon}^{LD} = -z_{i\varepsilon}^{LD} + r_{i\varepsilon}^+ + r_{i\varepsilon}^-$. All of the above processes are assumed to reach equilibrium fast. Thus the overall response of the circuit may be evaluated at steady-state, yielding

$$z_{i\varepsilon}^{LD} = \frac{1}{\alpha\gamma + \delta(p_{i\varepsilon}^+ + p_{i\varepsilon}^-)} \left(\alpha(p_{i\varepsilon}^+ + p_{i\varepsilon}^-) + 2\beta\, p_{i\varepsilon}^+ p_{i\varepsilon}^-\right) \qquad (2)$$

The dark-light response, $z_{i\varepsilon}^{DL}$, is obtained in a similar manner. This demonstrates the nonlinear interaction of activity between the two branches. Input is integrated linearly from both channels; juxtaposed activity in the ON *and* OFF pathway is signalled by an additional correlational (gating-type) component. The relative contribution of additive and gated activity is controlled by the shunting parameters α and β in Eqn. 1. The activity self-normalizes with respect to the total input activity from the ON and OFF channel. The model thus resembles properties of the scheme proposed by [1] to normalize activity of cortical neurons through division of pooled activity from a large number of cells.

Mutual Inhibition of Cells. Simple cells of opposite polarity and same spatial location are postulated to undergo mutual inhibition ([2, 7], see Fig.1 (b)). The final step to compute light-dark responses is given by

$$Z_i^{LD} = \left[z_i^{LD} - z_i^{DL}\right]^+ \quad \text{and} \quad Z_i^{DL} = \left[z_i^{DL} - z_i^{LD}\right]^+. \qquad (3)$$

Complex Cell Responses. Complex cell responses are, *in vivo*, insensitive to direction of contrast [3] and are obtained in the model, for simplicity, by pooling light-dark and dark-light simple cell responses.

ESANN'1996 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 24-25-26 April 1996, D-Facto public., ISBN 2-9600049-6-5, pp. 193-198

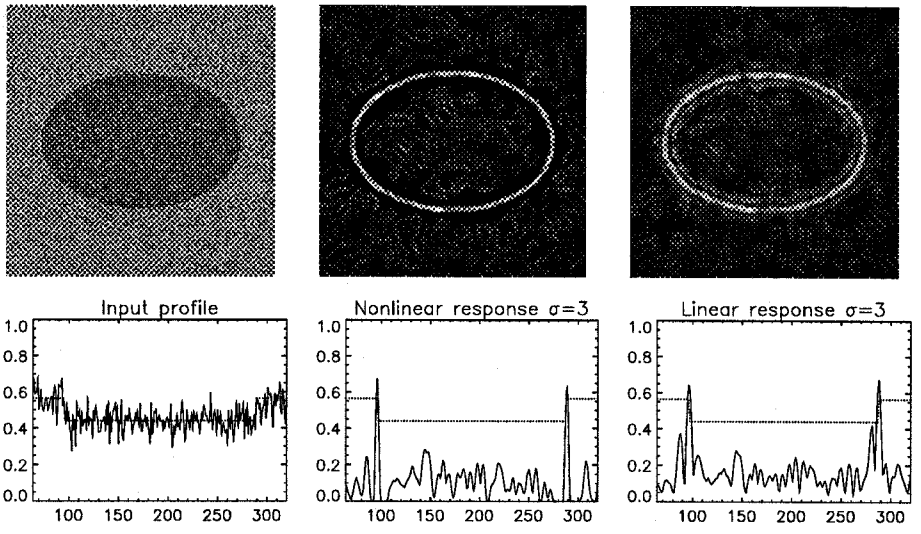

Fig. 2: Elliptic region with step contrast: Upper row: Input luminance distribution corrupted by additive Gausian noise with 50% amplitude of contrast height (a); pooled activity of all orientation fields generated by the nonlinear circuit (b) and the linear model (c). Bottom row: Luminance profile of the input image (d); profiles of activation generated by the nonlinear circuit (e) and those generated by a linear cell (f).

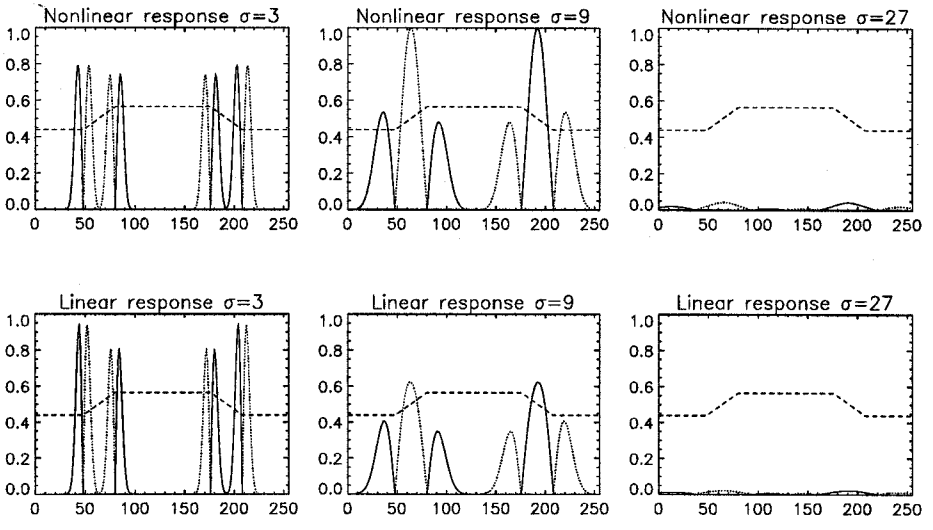

Fig. 3: Gradual transition: A ramp profile (dashed) is processed using a number of orientation selective blurrings viz scales. Responses were generated by the nonlinear circuit (upper row) and the linear model (bottom row). Boxes from left to right are generated for increasing scale parameters; responses for LD transition (solid) and DL transition (dotted). The asymmetries in response at different luminance levels occur due to initial shunting center-surround processing.

ESANN'1996 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 24-25-26 April 1996, D-Facto public., ISBN 2-9600049-6-5, pp. 193-198

Fig. 4: Processing a real camera image at scale $\sigma = 3$ (a): Pooled activity of all orientation fields generated by the nonlinear circuit (b) vs. the linear model (c).

3. Computer Simulations

A set of computer simulations demonstrate the functionality of the model. The nonlinear processing results are compared with a model that integrates activity from left and right branches in a linear fashion[2]. In all simulations, simple cell responses are shown after mutual inhibition ($Z_{i\varepsilon}^{LD}$ and $Z_{i\varepsilon}^{DL}$). The model parameters of the nonlinear circuit were set to $\alpha = 1.0$, $\beta = 2.0$, $\gamma = 0.5$, and $\delta = 1.0$. The Gaussian weighting functions were elongated by a 2:1 ratio; the variance σ is measured at the short axis (in pixels). The separation τ grows linearly with the variance. Eight discrete, equally spaced orientations were processed.

Figure 2 shows the model behavior when presented with an elliptic region. The luminance distribution has been corrupted by Gaussian noise (half width 50% contrast amplitude, Fig. 2 a). Activity of all orientations has been pooled yielding an "energy" distribution of simple cell activity (Fig. 2 b–c). The profile of the nonlinear activity distribution (Fig. 2 e) shows higher positional selectivity than in the linear case (Fig. 2 f). While the average noise activity is similar, the activity close to the contrast edge is apparently suppressed.

The multiple spatial frequency selectivity is demonstrated on the basis of processing a ramp transition. Figure 3 shows the results generated by the nonlinear circuit (top row) in comparison to the linear cell responses (bottom row). Increased scale-dependent blurring of ON and OFF input together with an increased spatial separation of input branches eventually generates a juxtaposition of ON and OFF activation for a contrast cell that is located at the center of the ramp. The maximum response of the linear model still occurs at a fine scale and has its peak at the left and right end of the ramp transition (bottom left). In contrast, the nonlinear circuit produces a unique maximum response in scale-space at the ramp edge; the corresponding scale is directly related to the width of the luminance transition (Fig. 3 (top center)).

Similar results are achieved for processing a camera image (Fig. 4). The response of the nonlinear model (at the scale of $\sigma = 3$) shows again higher positional selectivity than the response of the corresponding linear model. Edges are detected sharper and more accurately, and the overall noise level appears to be reduced.

[2] This model can be realized by replacing eq. (2) with $z_{i\varepsilon}^{LD} = p_{i\varepsilon}^{+} + p_{i\varepsilon}^{-}$. The linear integration model with elongated Gaussian lobes for ON and OFF subfields approximates a first order derivative operation [8].

ESANN'1996 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 24-25-26 April 1996, D-Facto public., ISBN 2-9600049-6-5, pp. 193-198

4. Summary

A neural network model is described that realizes nonlinear pooling of visual ON and OFF pathways. The investigation serves as a framework for the functional modeling of early stages in mammalian visual information processing. It is argued that the role of pooling ON and OFF activation in static form perception goes beyond local linear contrast detection. The experiments show that the nonlinear pooling leads to higher positional selectivity of local contrast changes and, in a multi-scale hierarchy, to a unique, scale-dependent maximum response where scale directly relates to the width of the luminance transition. Our simulations indicate that the nonlinear circuit is more robust given noisy images. This property has to be verified by further experiments.

The functionality of the corresponding 1D model has also been justified in a broader context of grouping and brightness perception. The measurement of contrast magnitude (strength) and sharpness (correlation) allows for categorial switching in the control of brightness perception [10]. Thus, the model includes several functional aspects that have been treated separately in previous contributions to model simple cell behavior. The approach will be further pursued to incorporate the processing of even-symmetric luminance variations and end-stop characteristics for proper processing of corners and curvature.

References

[1] M. Carandini and D.J. Heeger. Summation and division by neurons in primate visual cortex. *Science*, 264:1333 – 1336, 1994.

[2] D. Ferster. Spatially opponent excitation and inhibition in simple cells of the cat visual cortex. *Journal of Neuroscience*, 8(4):1172 – 1180, 1988.

[3] K.H. Foster, J.P. Gaska, M. Nagler, and D.A. Pollen. Spatial and temporal frequency selectivity of neurones in visual cortical areas V1 and V2 of the macaque monkey. *Journal of Physiology*, 365:331 – 363, 1985.

[4] S. Grossberg. 3-D vision and figure-ground separation by visual cortex. *Perception and Psychophysics*, 55(1):48 – 121, 1994.

[5] P. Heggelund. Receptive field organization of simple and complex cells. In D. Rose and V.G. Dobson, eds., *Models of the Visual Cortex*, p.358 – 365. Wiley, 1985.

[6] L.A. Iverson and S.W. Zucker. Logical/linear operators for measuring orientation and curvature. TR TR-CIM-90-06, McGill University, McGill Research Centre for Intelligent Machines, Computer Vision and Robotics Laboratory, 1990.

[7] Z. Lui, J.P. Gaska, L.D. Jacobson, and D.A. Pollen. Interneuronal interaction between members of quadrature phase and anti-phase pairs in the cat's visual cortex. *Vision Research*, 32(7):1193 – 1198, 1992.

[8] H. Neumann and K. Ottenberg. Estimating attributes of smooth signal transitions from scale-space. In *11th IAPR Int. Conf. on Pattern Recognition (ICPR-92)*, pages (III) 754 – 758, Le Hague, NL, 1992. IEEE Computer Society Press.

[9] H. Neumann and L. Pessoa. A simple cell model with multiple spatial frequency selectivity and linear/non-linear response properties. In *Proc. WCNN-94*, pages (IV) 290–298, San Diego (CA/USA), June 4 - 9 1994. LEA.

[10] L. Pessoa, E. Mingolla, and H. Neumann. A contrast- and luminance-driven multiscale network model of brightness perception. *Vis. Res.* 35(15):2201–2223,1995.