1994-02

A Multi-Scale Network Model of Brightness Perception

Mingolla, Ennio

Boston University Center for Adaptive Systems and Department of Cognitive and Neural Systems

http://hdl.handle.net/2144/2148

Boston University
A MULTI-SCALE NETWORK MODEL OF BRIGHTNESS PERCEPTION

Ennio Mingolla, Heiko Neumann, and Luiz Pessoa

February 1994

Technical Report CAS/CNS-94-010

Permission to copy without fee all or part of this material is granted provided that: 1. the copies are not made or distributed for direct commercial advantage, 2. the report title, author, document number, and release date appear, and notice is given that copying is by permission of the BOSTON UNIVERSITY CENTER FOR ADAPTIVE SYSTEMS AND DEPARTMENT OF COGNITIVE AND NEURAL SYSTEMS. To copy otherwise, or to republish, requires a fee and/or special permission.

Copyright © 1994

Boston University Center for Adaptive Systems and
Department of Cognitive and Neural Systems
111 Cummington Street
Boston, MA 02215
A MULTI-SCALE NETWORK MODEL OF BRIGHTNESS PERCEPTION

Ennio Mingolla,†Heiko Neumann‡§ and Luiz Pessoa†

and

† Center for Adaptive Systems and
‡ Department of Cognitive and Neural Systems,
Boston University, 111 Cummingston Street, Boston, MA 02215, USA
§ Universität Hamburg, FB Informatik, AB Kognitive Systeme,
Bodenstedtstr.16, D-22765 Hamburg, Germany

Abstract

A neural network model of brightness perception is developed to account for a wide variety of difficult data, including the classical phenomenon of Mach bands and nonlinear contrast effects associated with sinusoidal luminance waves. The model builds upon previous work by Grossberg and colleagues on filling-in models that predict brightness perception through the interaction of boundary and featural signals. Model equations are presented and computer simulations illustrate the model's potential.

1 Introduction

Brightness phenomena are a rich source of information on how the visual system encodes luminance changes in the world. In this paper we develop a neural network for brightness perception in the tradition of filling-in theories (Cohen and Grossberg, 1984; Gerrits and Vendrick, 1970). Our simulations implement a number of refinements already described in the development of Grossberg's (1987, 1994) Form-And-Color-And-DEpth (FACADE) theory, which though conceived as part of the theory, were not implemented in the simulations of Grossberg and Todorovic (1988). These include: a) ON and OFF channels with separate filling-in domains; b) multiple spatial scales; c) computations for simple and complex cells; and d) boundary computations that engage a recurrent competitive circuit. Simulations of the present system of equations account for human's perception of a wide variety of stimuli, including ones whose brightness contains shallow spatial gradients.

2 Smooth Brightness Gradients within the BCS/FCS Theory

A fundamental idea of the BCS/FCS theory is that boundaries are used to generate filling-in compartments where featural quality (e.g., "brightness" in our case) is diffused, or spread. The final diffused activities in the FCS correspond to the model's predicted brightness, which is the outcome of interactions between boundaries and featural quality, whereby boundaries control the process of filling-in by forming gates of variable resistance to diffusion.

Note that while it may seem natural to assume that boundary signals only exist in locations corresponding to discontinuities of luminance ("edges"), Grossberg and Mingolla (1987) showed that "boundary webs" can form in regions of luminance gradients, whereby the process of diffusion

†Supported in part by the Air Force Office of Scientific Research (AFOSR-F49620-92-J-0334), the Northeast Consortium for Engineering Education (NCEE-A303-21-90), and the Office of Naval Research (ONR N00014-91-J-4100).
‡Supported in part by the German BMFT, grant 413-5839-01 IN 101 C/1.
§Supported in part by CNPq and NUTES/UFRJ, Brazil.
may be totally or partially blocked within extended regions, yielding a percep of spatially gradual changes in brightness. Boundary signals work to contain diffusion; large boundary values do not allow a featural value at a given spatial position to affect a neighboring one. In regions with zero boundary activity, featural quality is free to diffuse, while in regions containing spatially dense boundary signals, little diffusion of featural quality throughout a large area may occur. In other words, featural quality cannot be spread, and the corresponding predicted brightness will be similar to the profile of featural quality derived by the initial filtering of the scenic input (image) at those spatial positions. In this sense, extended boundaries of sufficient amplitude can be thought of asprint signals, that is, signals that simply replicate the filtered input to the filling-in stage.

2.1 Boundary Computations

Analysis of several brightness stimuli indicates that stimuli with abrupt luminance transitions (e.g., luminance steps) generally require sharp boundary signals to create spatially abrupt barriers between regions of discretely differing brightness levels, while stimuli containing smooth luminance modulations will require smoother boundaries to be able to trap (at least some of) the modulation that is present in the convolved input (featural quality) to create smoothly varying brightness distributions. How then can the visual system, or a model of it, decide without a homunculus — rules invoked by other processes — whether or not to sharpen boundary signals?

What differentiates the situations requiring sharp and extended boundary signals? Consider a system where the input waveform is filtered by both ON and OFF center-surround operators. The solution originates from the observation that sharp transitions of luminance produce strong responses in both the ON and OFF channels. In other words, in the region surrounding an “edge,” there will be strong ON-activity (at the “light” region) and strong OFF-activity (at the “dark” region). On the other hand, waveforms with more gradual variations of luminance lead to a different distribution of ON and OFF responses.

The above analysis indicates that ON and OFF responses can be used to guide the computation of boundaries: ON/OFF spatial coincidence (Gove, 1993; Grossberg, 1994; Grossberg, Mingolla, and Williamson, 1993; Iverson and Zucker, 1989) should eventually lead to sharp boundaries while ON/OFF separation should produce spatially shallow signals. We propose a two stage process for achieving the computational competencies necessary to generate the proper activity distribution within the BCS: 1) initial boundary responses to stimuli with ON/OFF coincidence should be large; and 2) responses that are “large” relative to neighboring ones are sharpened while those that are (absolutely) small or near in size to neighboring ones are not.

3 Multiple Scale Contrast and Brightness System

Recently, Neumann (1993, 1994) presented an analysis that clarified how the double-opponent subtraction of ON/OFF signals for brightness computation (Grossberg, 1987, 1994; Grossberg and Wyse, 1992) can be thought of as “factored” into the computation in parallel of “contrast” and “luminance-derived” signals. The current model combines the work of Grossberg and Todorovic (1988) and of Neumann (1993, 1994) and modifies them in several ways (see Introduction). The full description of the model follows (Figure 1). The implementation employed is 1-D, i.e., stimuli of interest are actually slices through full 2-D stimuli with 1-D symmetry. As far as possible the parameters of the model have been kept constant across scales; spatial parameters that vary according to scale have an index $S$. Variables that are computed in multiple scales also contain the index $S$.

The input pattern to the model is a spatial pattern given by $I_t$. 

2
Center-surround Antagonism and On- and Off-channels (ON/OFF filter). The input pattern is processed by both ON- and OFF-cells in a way similar to retinal ganglion cells. This is accomplished with membrane equations (Grossberg and Todorovic, 1988).

Segregation of Contrast and Brightness Systems (ON/OFF contrast and luminance). Call the above ON and OFF filtering responses by $Y^+_t$ and $Y^-_t$, respectively. As proposed by Neumann (1993, 1994), the representation of “contrast” signals is obtained via inhibitory interactions between ON and OFF channels as in

$$
  c^+_t = [Y^+_t - Y^-_t]^+
$$

(1)

$$
  c^-_t = [Y^-_t - Y^+_t]^+
$$

(2)

where $[x]^+ = \max(x, 0)$. “Baseline”, or luminance-driven, activity is obtained by pooling the output of ON and OFF channels, obtaining a low-pass filtered and non-linearly compressed version of the input

$$
  \delta_t = Y^+_t + Y^-_t.
$$

(3)

Simple and Complex Cell Responses. Simple and Complex cells are the first major stage leading to the computation of boundaries. Before feeding into simple cells, ON- and OFF-contrast signals ($c^+_t$ and $c^-_t$) are first blurred by Gaussian kernels producing the signals $p^+_t$ and $p^-_t$, where $S = 1, \ldots, k$, and $k$ is the number of scales.

The model employs both light-dark and dark-light simple cells. These are obtained by collecting contrast information from spatially different branches. Consider a simple cell at position $i$. For a light-dark cell, ON-information originates from the “left” and OFF-information from the “right” (with respect to position $i$). To simplify the notation below we use the following convention:
$l = i - \sigma_S$ and $r = i + \sigma_S$, where $l$ and $r$ are the “left” and “right” spatial offsets and $\sigma_S$ is a scale-dependent constant. Responses are assumed to reach equilibrium fast and are computed in two steps. A light-dark cell involves the following computations for the left branch (fed by the ON channel)

$$q_i^{S+} = \frac{p_i^{S+}}{\alpha + \beta q_i^{S-}}$$

$$z_i^{S+} = \frac{p_i^{S+}}{\gamma + \delta q_i^{S+}}$$  \hspace{1cm} (4)

where $\alpha$, $\beta$, $\gamma$, and $\delta$ are constants. Similarly for the right branch (fed by the OFF channel) the equilibrated values are

$$q_i^{S-} = \frac{p_i^{S-}}{\alpha + \beta q_i^{S+}}$$

$$z_i^{S-} = \frac{p_i^{S-}}{\gamma + \delta q_i^{S-}}$$  \hspace{1cm} (5)

The final simple cell response is $z_i^{S+} + z_i^{S-}$. A dark-light cell is obtained by reversing the inputs for the “left” and “right” branches.

Complex cell responses, $x_i^S$, are insensitive to direction of contrast and are obtained by summing light-dark and dark-light simple cell responses. It was assumed that the complex cell output was a scaled-thresholded version as in $X_i^S = \kappa[x_i^S - T]^+$, where $[x]^+ = \max(x, 0)$; $\kappa$ and $T$ are constants.

**Boundaries:** Feedback Competition of Complex Cell Responses (boundary). Boundaries are obtained by processing complex cell responses through a recurrent competitive network. The system sharpens strong inputs and leaves small signals largely unmodified. The nonlinear feedback network is modeled after the work of Grossberg and Marshall (1989) and is given by

$$\frac{dw_i^S}{dt} = -Aw_i^S + (L - w_i^S)(F_i^{S+} + B_i^{S+}) - (M + w_i^S)(F_i^{S-} + B_i^{S-}).$$  \hspace{1cm} (6)

where $A$, $L$ and $M$ are constants; see Grossberg and Marshall (1989) for more details. Equation 6 is solved using fourth order Runge-Kutta until the activities equilibrate; the computations are similar to those of the “first competitive stage” of the Cooperative-Competitive (CC) loop of Grossberg and Mingolla (1985).

**On and Off Feature Filling-in** (ON/OFF filling-in). Filling-in is performed for both ON and OFF domains (Gerrits and Vendrick, 1970; Grossberg, 1987, 1994; Grossberg and Wyse, 1992). Diffusive filling-in for the ON domain is implemented as

$$\frac{dv_i^{S+}}{dt} = -Kv_i^{S+} + \sum_{j \in N_i} (v_j^{S+} - v_i^{S+})P_{ji}^S + c_i^{S+}$$  \hspace{1cm} (7)

where $K$ is a constant and $N_i$ specifies the neighborhood of influence of node $i$. Term $c_i^{S+}$ is the on contrast that is supplied to the diffusion stage; a similar equation regulates the OFF domain ($v_i^{S-}$), with $c_i^{-}$ as the contrast input. Diffusion is limited to nearest neighbors so that $N_i = (i - 1, i + 1)$. The diffusion coefficients, $P_{ji}^S$, regulate the magnitude of cross influence of location $j$ on location $i$ and depend on boundary signals as

$$P_{ji}^S = \frac{\rho}{1 + \epsilon(w_i^S + w_j^S)},$$  \hspace{1cm} (8)

where $\rho$ and $\epsilon$ are constants. For simulations, Equation 7 is solved with fourth order Runge-Kutta until activities equilibrate.

**Single Scale Brightness Prediction** (single scale brightness). The filled in activities in the ON and OFF domains are used in conjunction with the low pass luminance to determine the single scale brightness prediction, $u_i^S$. The basic idea is that the low pass luminance ($s_i$ in Equation 3)
provides a baseline of activity that can be modified by the equilibrated \( \text{ON} \) and \( \text{OFF} \) contrasts (\( v_i^{S+} \) and \( v_i^{S-} \) in Equation 7). The interactions are governed by the following differential equation

\[
\frac{du_i}{dt} = -F u_i^S + s_i + G v_i^{S+} - H u_i^S v_i^{S-}
\]

which is computed at equilibrium; \( F, G, \) and \( H \) are constants.

**Multiple Scale Brightness Pooling** (multiple scale brightness). As implemented in Grossberg, Mingolla, and Williamson (1993), the final brightness percept is obtained by averaging the outputs of the different spatial scales:

\[
U_i = \frac{1}{k} \sum S u_i^S
\]

where \( k \) is the number of scales being employed. \( U_i \) is the final output of the model.

4 Simulations

The model correctly predicts the appearance of Mach bands on a trapezoidal wave and Figure 2 shows several model stages (3 spatial scales were used). In the bottom row the input \( (I_i) \) is shown to the left of the final multiple scale brightness prediction \( (U_i) \). The input luminance is initially filtered by \( \text{ON} \) and \( \text{OFF} \) channels and the results generate “contrast” and “luminance” information; the second row shows \( \text{ON/OFF} \) filtering \( (Y_i^+ \) and \( Y_i^-) \), \( \text{ON/OFF} \) contrasts \( (c_i^+ \) and \( c_i^-) \), and the low-pass luminance \( (s_i) \). Contrasts are used for multi-scale computations of complex cells and boundaries. At the same time they are used as featural inputs to \( \text{ON/OFF} \) diffusion leading to single scale brightness predictions. The contrasts fed to diffusion are trapped by the spatially extended boundaries and provide the activities that will generate the corresponding bands for the single scale predictions; for scale 1 some of the complex cell responses are sharpened, though. Rows 3-5 show these multi-scale computations for the trapezoidal wave. Column 1: complex cells \( (X_i^S) \); column 2: boundaries \( (W_i^S) \); column 3: equilibrated \( \text{ON} \) and \( \text{OFF} \) filling-in \( (v_i^{S+} \) and \( v_i^{S-}) \); column 4: single scale brightness \( (u_i^S) \). The final multiple scale brightness prediction is obtained by averaging the single scale results and, as mentioned, is shown in the first row \( (U_i) \). The model also correctly predicts the existence of Mach bands on triangular waves (see Ross et al., 1989).

Other stimuli that the model can account for include (Figure 3) a) the square-wave; b) the high contrast missing fundamental (MF), which is perceived more or less veridically; c) the low contrast MF, which is perceived as a square wave; and d) a high contrast sine wave, perceived with compression. This set of data was chosen so as to illustrate the model’s ability to generate sharp or extended boundaries as a function of the input luminance; the square wave and low contrast MF necessitate sharp boundaries while the high contrast MF and sine wave require extended signals. Finally, it should be emphasized that all of our simulations, including ones not shown in this paper, employ the same set of parameters.

5 Conclusion

We have presented an implementation of a neural network model of brightness perception that can account for some challenging data involving slow variations in brightness; as well as sharp transitions. The model can account for the classical phenomenon of Mach bands as well as other stimuli (e.g., brightness contrast). Finally, the model should be compared to alternative approaches attempting to explain similar sets of brightness data. For example, while the local energy model of Burr and Morrone obtains rather good quantitative fits to Mach bands (Ross et al., 1989),
it does not account for several of the stimuli simulated here (e.g., sine wave). The MIDAAS model of Kingdom and Moulden (1992) accounts for the 1-D phenomena investigated here but does so by employing symbolic interpretation rules (homunculus) that, we feel, are bound to yield contradictions requiring appeal to other rules in a 2-D implementation. The present implementation of ideas from FACADE theory (Grossberg 1987, 1994), on the other hand, has a natural extension to a 2-D implementation. Indeed an implementation of a related multi-scale network for processing large images has already been described in Grossberg, Mingolla, and Williamson (1993).

Reference


Figure 2: Simulation of a trapezoidal wave. Graphs show activity (in arbitrary units) as a function of spatial position \( i \). Bottom row: Input and multi-scale brightness. Second row: ON/OFF filtering, ON (solid) /OFF (dotted) contrast, and low-pass (LP) "luminance". Rows 3-5: Multi-scale computations (3 scales); OFF filling-in (dotted) shown with negative values for illustration only.
Figure 3: Simulations of square wave (left), high contrast missing fundamental (second column), low contrast missing fundamental (third column), and high contrast sine wave (right). All four simulations display the input luminance (top) and the final multiple scale brightness (bottom). All simulations use the same set of parameters.