A neural architecture of brightness perception: non-linear contrast detection and geometry-driven diffusion

Heiko Neumann\textsuperscript{a,∗}, Luiz Pessoa\textsuperscript{b}, Ennio Mingolla\textsuperscript{c}

\textsuperscript{a}Universität Ulm, Fakultät für Informatik, Abteilung Neuroinformatik, Oberer Eselsberg, D-89069 Ulm, Germany
\textsuperscript{b}Universidade Federal do Rio de Janeiro, Programa de Engenharia de Sistemas e Computação—COPPE Sistemas, Ilha do Fundão, Rio de Janeiro, RJ 21945-970, Brasil
\textsuperscript{c}Boston University, Center for Adaptive Systems and Department of Cognitive and Neural Systems, 677 Beacon Street, Boston, MA 02215, USA

Received 15 June 1996; revised 8 May 1997; accepted 22 September 1997

Abstract

A neural architecture for brightness perception is constructed in the tradition of filling-in theories. The model is developed to account for a wide variety of difficult data, including the classical phenomenon of Mach bands, low- and high-contrast missing fundamental and non-linear 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 feature signals. A new interpretation of feature signals through the explicit representation of contrast-driven and luminance-driven information is provided and directly addresses the issue of absolute brightness values. Simulations of the model implement a number of refinements with respect to the previous implementation of Grossberg and Todorović\textsuperscript{1} [S. Grossberg, D. Todorović, Neural dynamics of 1-d and 2-d brightness perception: a unified model of classical and recent phenomena. Perception and Psychophysics, 43 (2) (1988) 241-277]. These include: (a) ON and OFF channels with separate filling-in domains; (b) multiple spatial scales; (c) non-linear computations for simple and complex cells; and (d) boundary computations that engage a recurrent competitive circuit. The net effect of mechanisms involved in the computational model accomplish a unique solution for the brightness-from-luminance problem. The two parallel and topographically organized subsystems of a boundary contour (BCS) and feature contour system (FCS) are demonstrated to generate an isomorphic representation of brightness distributions. The activity of the contour-sensitive BCS regulates the process of diffusive filling-in. It realizes an adaptive form-sensitive mechanism for the control of lateral spreading of local activation in the diffusion system. It is shown that, under certain stimulus conditions and structure of the input generators to the filling-in processes, the action of BCS/FCS interaction realizes a membrane regularization of the problem of brightness reconstruction. Simulations of the present system of equations account for human perception of a wide variety of stimuli, including the ones studied by Georgeson\textsuperscript{2} [M. Georgeson, From filters to features: location, orientation, contrast and blur, Proc. CIBA Symp. on Higher-Order Processing in Vision, London, Vision, London, 19-21, 1993], whose shallow spatial gradients have posed difficulties to alternative early vision theories. Because boundary signals may undergo reorganization, including long-range grouping before feature diffusion proceeds, the proposed architecture may also serve as an alternative framework for non-linear anisotropic diffusion approaches developed recently for early processes in computer vision. © 1998 Elsevier Science B.V. All rights reserved.

Keywords: shunting networks; neural architecture; ON- and OFF-pathways; luminance pathway; brightness anchoring; filling-in; geometry-driven diffusion; regularization

1. Introduction

In this paper we develop a neural model of brightness perception in the tradition of filling-in theories ([1,3,4], see also [5]). The output generated by the model is an activity profile that is, ideally, isomorphic with a human’s brightness distribution in response to the corresponding luminance stimulus (see [5]). The neural network developed here uses the boundary/feature contour system (BCS/FCS) framework proposed by [4] and [6,7] to explain brightness data. The monocular BCS/FCS defines a basic building block of the full FACADE theory developed by Grossberg [8-10] proposed to account for a much broader set of empirical findings of form, color, brightness and depth perception.

Relative to earlier developments (such as [1]), the present model includes several modifications and extensions, notably (a) ON and OFF channels with separate filling-in domains; (b) a new non-linear circuit for simple cells for oriented contrast detection; (c) boundary computations that

\* Corresponding author.
engage a recurrent (feedback) competitive circuit in the BCS used for boundary sharpening; (d) a new treatment for the FCS, including contrast-driven and luminance-driven channels; and (e) multiple spatial scales. The combination of the non-linear simple cell with the feedback circuit in the BCS produces qualitatively different boundary signals for shallow (e.g. ramp-like) and sharp (e.g. step-like) luminance transitions. Depending on the input luminance distribution very localized or spatially extended boundaries may result. The extension of the FCS to include two parallel, ON and OFF, contrast-driven channels for filling-in and a separate luminance-driven channel to ‘anchor’ the level of brightness activity, modifies the role diffusion mechanisms have in predicting brightness appearance.

The functionality of the model, viz. its components and their interactions, has been motivated by several empirical findings in contrast and brightness perception and related physiology. In particular, the initial segregation of processing streams into separate ON and OFF pathways [9,11–14] provides a unified mechanism for contrast detection and transmission of (compressed) luminance information [15–17]. The non-linear circuit proposed for modeling simple cell functionality takes into account various non-linear properties in the generation of contrast responses [18–23] that cannot be solely accounted for by purely linear models. The non-linear recurrent field enables orientational as well as spatial sharpening by feedback interactions, and has been used, among other things, to account for binocularly fused cell responses [24], and for the modeling of cortical orientational selectivity [25–27]. In combination with the simple-cell circuit, the recurrent field is a major component of a new explanation of the phenomenon of Mach bands—the account can also rationalize why Mach bands exist for shallow, ramp-like, luminance transitions, while they are non-existent at sharp transitions [28–30]. The action of a boundary-gated filling-in mechanisms has been hypothesized to account for several completion phenomena, including the bridging of blind spot and retinal veins [31], and the filling-in of cortical scotomas [32]. It has been suggested that such mechanisms are also involved in the fading of stabilized images. Finally, recent results on the temporal dynamics of brightness perception by Paradiso and his colleagues have been interpreted in terms of brightness filling-in mechanisms that are consistent with the presented proposal (see below).

This paper provides an overview of the proposed architecture that emphasizes the non-linear interactions between different model components. For further details concerning the psychophysical results being modeled, as well as a comparison with alternative computational theories, the reader is referred to Ref. [33]. The behavior of the model is illustrated through a series of simulations that address classic brightness phenomena. New simulation results of recent brightness data by Georgeson [2] are used to question the widespread assumption of early vision models that a discrete, symbolic, description is employed to code important visual ‘events’. The dynamic nature of brightness perception is also discussed as we focus on new experimental findings by Rossi and Paradiso [34] and how the presented model may be able to account for them.

One of the central aims of the present paper is to relate our brightness filling-in framework to classical regularization theory. In particular, we show that for certain stimulus conditions and input configurations active filling-in generates a regularized solution for the reconstruction of brightness data from initial estimates triggered at boundary locations. In this context, we show that our architecture, which has been motivated by empirical data, provides an alternative framework to geometry-driven diffusion for the non-linear processing of image data as used in recent computational vision approaches. The presented proposal separates the processes dedicated to oriented contrast detection and contour grouping from processes involved in feature reconstruction and representation; see Refs [6,7] for related empirical evidence (see also Ref. [35]). This should be contrasted with recent approaches to non-linear diffusion that treat both types of processes within a single framework (e.g. [36–38]).

2. Boundary computation, feature signals and brightness gradients within the BCS/FCS theory

2.1. Overview of FACADE theory

During the last two decades Grossberg and his colleagues have proposed several components of a neural network architecture for general-purpose preattentive vision. This ongoing work culminated in the FACADE theory\(^1\) described in Refs [8–10]. The present section provides a brief overview of the key concepts of the theory, emphasizing the principles underlying monocular visual processing.

FACADE theory stresses the highly context-dependent and globally coherent nature of visual perception. It has attempted to analyze a large number of seemingly contradictory data in an attempt to identify fundamental constraints and basic mechanisms underlying biological visual information processing. This biological-evolutionary approach should be contrasted to other contributions to the understanding and modeling of visual processes, such as the more modular theories of Marr [39] and Ballard [40]. Based on the analysis of a large body of data encompassing results from psychophysics, physiology and anatomy, the existence of two interacting, but functionally autonomous, perceptual sub-systems has been proposed (see e.g. [6–10]):

- The boundary contour system (BCS), which is primarily responsible for the processing of oriented luminance

\(^1\) FACADE stands for ‘Form-And-Color-And-DEpth’. It characterizes the coherence of representations and processes in biological vision in contrast to modular architectures of specialized and functionally autonomous ‘experts’ or value-encoding units.
contrasts, generates stimulus-dependent segmentations as it generates form-sensitive scien
copic groupings.

- The feature contour system (FCS) generates surface properties such as brightness, color or depth.

Fig. 1 shows a macro-circuit of the BCS/FCS architecture. The architecture consists of the two major sub-systems introduced above, namely the BCS and the FCS. Both are fed by activities generated in response to the initial input luminance distribution. Both context-dependent image segmentations (BCS) and form-sensitive featural processing (FCS) can be influenced by previously learned patterns that are stored in the long-term memory traces of the object recognition system (ORS). The proposal presented here only focuses on model components that are driven by the bottom-up sensory pathway, i.e. no attentive top-down components are considered.

The functionality of the BCS and FCS helps to hierarchically compensate for a number of perceptual uncertainties. Based on noisy and fragmented local contrast measurements, object/surface outlines must be segmented that may be partly obscured by other objects in the scene. The measurements themselves are influenced by the uncertainty of a finite spatial processing aperture. The detection processes of these surface regions in turn must compensate for variable illumination conditions. In order to account for these and other constraints, the BCS and FCS employ complementary processes [41]. Phenomena related to oriented contrast perception, such as grouping and illusory contour effects, constrain the design of the BCS and suggest mechanisms that are orientation selective, insensitive to local contrast polarity (i.e. light–dark and dark–light luminance transitions are pooled), and that inwardly group local contrast inducing measurements. Such inducing elements are combined by a co-operative–competitive circuit that groups fragmented local measurements in order to generate an emergent context-sensitive boundary segmentation [6,7]. At the same time, featural qualities, such as a brightness, are reconstructed in the FCS by utilizing mechanisms of isotropic activation spreading that are dependent on local contrast polarity. The context-sensitive segmentations generated in the BCS are signalled to the FCS such that the topographically organized activations in the BCS structure the FCS processes by forming compartments that control the isotropic activation spreading process. In other words, the orientation insensitive activation spreading process of the FCS is thereupon modulated by the activity patterns in the BCS.

2.2. Brightness gradients

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 predicted brightness, 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’), in Ref. [42] it was shown that ‘boundary webs’ can form in regions of luminance gradients, whereby the process of diffusion may be totally or partially blocked within extended regions, yielding a percept of spatially gradual changes in

![Diagram of functional organization of BCS and FCS subsystems](image)

Fig. 1. General overview of the functional organization of the BCS and FCS subsystems (for details see text; adopted from Ref. [6]).

![Diagram of brightness gradients](image)

Fig. 2. Two ways that filling-in can result in perceived smooth gradients of brightness. (a) This panel shows a one-dimensional cross-section of boundary and feature signals before diffusion. (b) Conservative diffusion by the heat equation would result in a flat featural profile at equilibrium. (c) Non-conservative diffusion by BCS/FCS mechanisms can strike a balance between sources (initial featural distributions) and sinks (passive decay of featural activity), resulting in a modulated distribution of brightness at equilibrium. (d) Continuous curves that denote boundary activity can be thought of as interpolated distributions between discrete network nodes (whose activity strengths are denoted by the lengths of vertical lines). (e) Spatially extended (fuzzy) distributions of boundary signals may overlap with distributions of featural signals. (f) Boundary signals can ‘trap’ featural quality over extended areas, resulting in smooth modulation of activity.
brightness. Boundary signals work to contain, or block, 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 (fuzzy) boundary signals, little diffusion of featural quality throughout a large area may occur (see Fig. 2). 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 as print signals, that is, signals that simply replicate the filtered distribution to the filling-in stage.

Fig. 2 shows two distinct ways in which a smooth or gradual modulation of perceived brightness may be realized as the outcome of the interaction of BCS and FCS mechanisms. The first is the non-conservative nature of the hypothesized diffusion itself, whereby a non-homogeneous equilibrium exists that balances the injection of new energy into the diffusion process (from earlier processing stages) with the tendency for the activity of cells at the diffusion layer to spontaneously decay away. The second mechanism can be thought of as the spatial ‘trapping’ of an entire signal profile from earlier processing stages by an extended region of strong BCS signals. Note that the need for the outcome of BCS/FCS interactions to be a smooth brightness profile for some inputs (e.g., sinusoidal waves) must be balanced against the need to generate flat brightness plateaus separated by sharp transitions for other inputs (e.g., luminance step).

The results of a given initial boundary filter will be some distribution of activity in a network of nodes with overlapping receptive fields. Some of these initial boundary signals will need to be sharpened, while others will need to be left fuzzy. Stimuli with abrupt luminance transitions (e.g., luminance steps) generally require sharp boundary signals. Such boundaries create spatially abrupt barriers to diffusion between regions, resulting in discretely differing output levels (corresponding to the model’s predicted brightness distribution). Stimuli containing smooth luminance modulations (e.g., sine waves), on the other hand, typically require smoother boundaries to be able to contain (at least some of) the modulation that is present in the featural quality to create smoothly varying brightness distributions.

How then can the visual system, or a model of it, decide (without a homunculus) 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. Specifically, the natural pairing of closely juxtaposed ON and OFF responses characteristic of luminance edges will not tend to occur, so that a peak in, for example, ON activity may be relatively isolated. The essential steps in achievement of sharp vs extended boundary signals are illustrated in Fig. 3.

The above analysis indicates that ON- and OFF-responses can be used to guide the computation of boundaries: ON/ OFF spatial juxtaposition [17, 19, 43–46] should eventually lead to sharp localized boundaries while ON/OFF separation should produce spatially broad and 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 (i.e., those that are spatially localized) are sharpened while those that are small (i.e., spatially distributed) are not. It is demonstrated that the combination of a non-linear simple-cell circuit with a recurrent network can automatically ‘switch’ its behavior to provide appropriate output in each case.

![Fig. 3. Mach bands for ramps vs no Mach bands for step edges. Two luminance profiles are shown: (a) a ramp, and (b) a step. Panels (c) and (d) show the responses of OFF (dashed) and ON (solid) channels. Responses to the ramp are smaller in amplitude and further apart than the corresponding responses to the step. Panels (c) and (f) display a transformation, including blurring and pooling of ON and OFF signals from (c) and (d), respectively. Recurrent competition can then leave the distribution of (e) effectively unchanged, as shown in (g), while transforming the distribution of (f) into that of (h) by a ‘Choice’ (winner-takes-all) transformation. Filling-in of featural signals is differentially affected by boundary signals of (g) and (h). The spatially extended boundaries that ‘survived’ recurrent competition can trap visible featural signals, resulting in Mach bands (undershoots and overshoots), as shown in (i). Panel (j) shows how the spatially collapsed boundary distribution corresponding to the luminance step segregates the resulting brightness profile into two relatively flat plateaus, without Mach bands.](image)
3. Computational mechanisms of contrast and brightness perception

3.1. Overview of main functional components

The outline of the main functional components of the architecture is presented in Fig. 4. The sketch shows the primary information-flow and the associated computational elements which define the key stages of the model. The vertical organization is characterized by three basic levels: (1) the input at the primary level is defined by the initial luminance distribution of the stimulus; (2) the intermediate level consists of the basic mechanisms for uncertainty reduction and regularized brightness reconstruction; (3) the final level realizes a multiple-scale spatial representation of the predicted brightness percept. The intermediate level in this layout (level (2)) in turn consists of two major pathways that are both driven by different signal components of the input distribution. One pathway (Fig. 4, left) consists of a contrast-driven process that itself feeds into the two parallel topographically organized subsystems, namely the BCS (denoted here as ‘boundaries’) and the FCS (denoted here as ‘filling-in’). The boundary system consists of a hierarchically organized sequence of non-linear contrast detection and a recurrent field that automatically sharpens localized highly energetic contrast activity. Filling-in is organized in parallel ON and OFF systems that allow featural quantities—related to brightness and darkness processes, respectively—to be diffused by an active diffusion mechanism of non-conservative filling-in. The diffusion in both sub-systems is locally regulated by the activation of the corresponding topographically organized boundary system. The steady-state solution of the filling-in activity in both ON and OFF compartments generates a representation of smooth brightness reconstruction for the corresponding initial contrast measures (see below). A complementary ‘luminance-driven’ pathway (Fig. 4, right) allows the adjustment of the activity distributions—generated by the filling-in and sensitive to relative contrast magnitude—so that they correspond to absolute brightness magnitudes.

The computational mechanisms involved will be outlined in the following Section 3.2. It is then shown in Section 4 how this relates to regularization theory and anisotropic diffusion approaches.

3.2. Description of computational mechanisms

The main focus of our initial modeling attempts has been

---

Fig. 4. Key stages of the computational model. This coarse-grain layout of the architecture consists of two major parallel pathways, both of which are fed by the input luminance. One pathway (left) consists of a contrast-driven process that itself feeds into the two parallel topographically organized subsystems of the BCS (‘boundaries’) and FCS (‘filling-in’). The output is a regularized solution of relative contrast reconstruction generated at outline boundaries. Together with the non-linearly compressed activity distribution of the ‘luminance-driven’ pathway (which provides a reference level, or ‘anchor’, for absolute brightness), a brightness activation is generated at the final stage (‘brightness’).
to account for a wide variety of brightness perception phenomena, including the classical phenomenon of Mach bands, low- and high-contrast missing fundamental stimuli, luminance staircases and Craik–O’Brien–Cornsweet arrangements, and non-linear contrast effects associated with sinusoidal waveforms. As already motivated above, the model builds upon previous work on filling-in in which brightness profiles are generated through the interaction of boundary and feature signals. The mechanisms of the different processing stages will be briefly outlined in the following paragraphs. The outline in Fig. 5 contains a more detailed description of the structure of the model. For better readability, each paragraph of the more detailed description contains a label useful as a link to the associated box that denotes the processing mechanism in Fig. 5. A more complete description of the newly developed model can be found in Ref. [33]. The implementation employed is one-dimensional, i.e. stimuli of interest are actually slices through full two-dimensional stimuli with intrinsic one-dimensional symmetry.

3.2.1. Input and initial center-surround processing

3.2.1.1. Center-surround antagonism and ON- and OFF-channels (ON/OFF filter). Initially, the input luminance distribution is spatially processed, or filtered, providing the activation feeding into the two main processing streams of contrast and luminance computations. This

Fig. 5. Diagram of model components. Stages that perform multiple scale computations are specified by stacks of shifted boxes. Simple and complex cells are indicated by polarity specific and direction of contrast insensitive ‘cells’, respectively. The final output is given by the stage ‘multiple scale brightness’. The names employed are indicated in parentheses in the formal description of the model stages.
initial processing stage displays properties resembling center-surround, antagonistic, retinal processing. The center-surround organization found in the retina is composed of two parallel pathways selective to opposite polarities, namely, ON and OFF. The ON channel is driven by local isotropic on-center/off-surround interactions, whereas the OFF channel is driven by a spatial off-center/on-surround interactions.

The properties of localized receptive field processing have been modeled in classical computer vision approaches by linear schemes of Gaussian weighted inputs to a target cell (Difference-of-Gaussians, [47]; Laplacian-of-Gaussian, [17]). However, it remains a matter of debate how to appropriately model the encoding and transmission of both contrast and overall luminance information [29]. In order to address this issue Neumann [12–14] presented an analysis of how the opponent interactions of ON/OFF signals for brightness computation [10] can be seen as ‘factored’ into the computation of parallel ‘contrast’ and ‘luminance-derived’ signals. As shown below, this factorization is a key element of our model.

The spatial input pattern \( L_i \) is processed by both ON- and OFF-cells in a way similar to retinal ganglion cells. This is accomplished with steady-state solutions of membrane equations of the form \( \dot{x} = -Ax_i + (B - x_i)net_i^+ - (C + x_i)net_i^- \) where \( net_i^+ \) and \( net_i^- \) are low-pass filtered versions of the input luminance \( L \), generated by \( net_i^+ = \Sigma L_j W_{ij} \) and \( net_i^- \) as constants, see e.g. Ref. [11]. Membrane equations of the type employed have been extensively studied before (e.g. [48–51]). These investigations have shown that shunting networks are mechanisms capable of automatically adjusting their sensitivity to inputs in a way that renders them sensitive to relative luminance values (see also Ref. [52]).

3.2.2. Segregation of contrast and brightness systems (ON/OFF contrast; luminance-driven). Representing the above ON and OFF filtering responses by \( x_i^+ \) and \( x_i^- \), respectively, as proposed in [12–14], the representation of contrast signals is obtained via inhibitory interactions between ON and OFF channels, as in

\[
\begin{align*}
   c_i^+ &= \frac{(B + C)x_i^+ - net_i^+}{A + net_i^+ + net_i^-} \\
   c_i^- &= \frac{(B + C)x_i^- - net_i^-}{A + net_i^+ + net_i^-},
\end{align*}
\]

where \( [x]^+ = \max(x, 0) \) (rectification). ‘Baseline’, or luminance-driven, activity is obtained by pooling the output of ON and OFF channels, \( s_i = x_i^+ + x_i^- \), generating a low-pass filtered and non-linearly compressed version of the input that yields:

\[
s_i = \frac{(B - C)(net_i^+ + net_i^-)}{A + net_i^+ + net_i^-},
\]

3.2.2.3. Simple and complex cell responses. Simple and complex cells are the first major stages leading to the computation of boundaries. As indicated above, ON and OFF responses can be used to guide the computation of boundaries: close spatial juxtaposition of ‘preboundary’ (contrast) signals generated by both ON and OFF signals should eventually lead to sharp boundaries, while relatively isolated signals from either channel should produce spatially shallow boundaries. To realize this competency in the present model requires two stages. First, a non-linear design for a contrast-sensitive ‘simple cell’ (so named in analogy with visual cortical cells sensitive to oriented contrast) is used. Simple cells are often diagrammed as elongated ellipses, with two subfields separated by the major axis. The second mechanism for insuring that sharp luminance transitions lead to sharp boundaries, while shallow transitions lead to fuzzy boundaries, is the use of a recurrent competitive circuit for processing the output of the non-linear simple cell (see the following paragraph on boundaries generated by feedback competition).

The behavior of the model simple cell is such that one half of the cell’s receptive field is excited and the other half inhibited by increases in luminance. Our proposed circuit employs this typical arrangement, but with an important modification. Suppose that a simple cell receives input from both ON and OFF center-surround cells of the earlier processing levels—corresponding to light increments and decrements, respectively. Then one half of the simple cell would be excited by ON input and the other half would be excited by OFF input. The simple cell circuit is such that the summation of a small amount of ON signal to the appropriate side and an equal OFF signal to the appropriate side results in a much larger response than, for example, doubling the initial input to the ON side and setting the OFF side to zero. As shown in Fig. 6, the proposed mechanism favors the appropriate spatial juxtaposition of ON and OFF signals to the appropriate sides of the simple cell by ‘penalizing’ input distributions that excite only one pathway. The penalty for the ON pathway is in the form of a tonic ‘self-inhibition’ along the pathway that can be shut off.

\(^2\)This operation on discretized data points corresponds to a projection of the sensory input to a spatial kernel (receptive field). Since the spatial coupling that implements the weighting function is only fed by a feed-forward data flow, the net response at this level could also be described by a spatial convolution of the input against the weighting function (kernel), such that \( net = L * W \) (‘*’ denotes the convolution operator).

\(^3\)The present implementation only employs odd-symmetric simple and complex cells. Even-symmetric cells (53) have not been included for simplicity since they were not needed in order to address the current data set. However, they can be easily incorporated into the present scheme by pooling the results of both even- and odd-symmetric mechanisms together.
only by the simultaneous inhibition of the inhibitor (disinhibition) from the OFF pathway. The simple cell thereby acts as a 'soft' edge detector, that is, a unit that is exquisitely sensitive even to small amplitude, sharp, spatial transitions at edges—by virtue of its amplified response to close ON/OFF juxtapositions—but which also responds (albeit in a reduced form) to smooth gradients of luminance, such as are found in the transition from 'flat' to 'ramp' portions of the luminance profiles of Mach band stimuli (see Fig. 3).

Before feeding into simple cells, ON- and OFF-contrast signals (c_i^+ and c_i^-) are first blurred by Gaussian kernels producing the signals p_i^S+ and p_i^S-, where S = 1, ..., k is the index of spatial scale, and k is the number of scales. Note that through different widths in the spatial coupling, different scales of frequency selective processing are defined. On the other hand, the generation of initial contrast responses x_i^S in the segregated channels is achieved by a scale-specific (fixed) parametrization for center-surround coupling. 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' (with respect to position i). The 'left' and 'right' spatial offsets are denoted by l = i - σ_S and r = i + σ_S (σ_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):

\[ q_i^{S+} = \frac{p_i^{S+}}{\alpha + \beta p_i^{S-}}, \]

\[ z_i^{S+} = \frac{p_i^{S+}}{\gamma + \delta q_i^{S+}}, \]

where α, β, γ and δ are constants. Similarly for the right branch (fed by the OFF channel) the equilibrated values are:

\[ q_i^{S-} = \frac{p_i^{S-}}{\alpha + \beta p_i^{S+}}, \]
The final simple cell response is $Z_i^S = z_i^S + z_i^S^-$ (see Fig. 6). A dark–light cell is obtained by reversing the inputs for the 'left' and 'right' branches.

Complex cell responses, $y_i^S$, 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 $Y_i^S = \kappa[y_i^S - T]^+$, where $\kappa$ and $T$ are constants.

### 3.2.2.4. Boundaries: feedback competition of complex cell responses (Boundary)

Boundaries are obtained by processing complex cell responses through a recurrent competitive network. The competitive circuit, in current parlance, is capable of automatically switching—that is, without a change of parameters—between 'winner-takes-all' and 'leader-takes-most' modes of behavior, or even, for relatively homogeneous input patterns, to have little effect at all on its input. Grossberg [50] called these modes 'choice' and 'contrast-enhancing', respectively. Note that the case of 'winner-takes-all' is obtained when sharp boundary transitions are desired (at luminance edges), while the 'softer' outputs occur just in those cases where fuzzy boundaries are needed (see Fig. 7).

The system sharpens strong inputs and leaves small signals largely unmodified—the latter thus generate proper activation distributions to produce 'print signals', or spatially dense boundaries. The non-linear feedback network is modeled after the work of Grossberg and Marshall [24], and is given by:

$$
\begin{align*}
\dot{w}_i^S &= -Dw_i^S + (L - w_i^S)(f_i^S + \beta_i^S) \\
&
-(M + w_i^S)(\gamma_i^S + \beta_i^-),
\end{align*}
$$

where $D$, $L$, and $M$ are constants. $f_i^S$ and $\gamma_i^S$ denote net feedforward and feedback activations in a recurrent center-surround anatomy (see [24] for more details). Feedforward net activations are generated by integrated complex cell responses $Y_i$, more specifically by:

$$
\begin{align*}
\gamma_i^S &= \sum_j Y_j W_{ij}^S, \\
\gamma_i^- &= \sum_j Y_j W_{ij}^-.
\end{align*}
$$

Feedback net activations are generated by integration of non-linearly transformed contour responses, namely by:

$$
\begin{align*}
\beta_i^S &= \omega \sum_j h(w_j) W_{ij}^S, \\
\beta_i^- &= \psi \sum_j h(w_j) W_{ij}^-.
\end{align*}
$$

with the constants $\omega$, $\psi$ and $h(x) = \max[x - T, 0]^{a}$ a signal function with 'faster-than-linear' above-threshold characteristics. The spatial kernels used in the above equations, $W_{ij}$, were all defined as Gaussians whose spatial parameters

---

**Fig. 7.** In a recurrent competitive field (a), strengths of lateral excitation (plus signs, solid arrow heads) and inhibition (minus signs, open arrow heads) fall off as a function of distance from a central node. Such a circuit can produce 'switching behavior', whereby distributions of responses with relatively sharp peaks are further sharpened (b), while those distributions where all values are (absolutely) small or where many values very similar in size to neighboring ones are essentially left intact (c).
depend on spatial scale. These computations are similar to those of the ‘first active stage’ of the co-operative–competitive (CC) loop described in Ref. [7].

3.2.2.5. ON and OFF feature filling-in (ON/OFF filling-in).
Filling-in is performed for both ON and OFF domains [3,9]. Diffusive filling-in for the ON domain is described as:

\[ v_i^+ = -Kv_i^+ + \rho \sum_{j \in \mathcal{N}_i} (v_j^+ - v_i^+) P_{ji}^+ + c_i^+ , \quad (9) \]

where \( K \) and \( \rho \) are constants and \( \mathcal{N}_i \) specifies the neighborhood of influence of node \( i \) (with nearest neighbor coupling \( \mathcal{N}_i = \{i - 1, i, i + 1\} \)). Term \( c_i^+ \) is the ON contrast that is supplied to the diffusion stage of the ON sub-system. A similar equation regulates the OFF domain:

\[ v_i^- = -Kv_i^- + \rho \sum_{j \in \mathcal{N}_i} (v_j^- - v_i^-) P_{ji}^- + c_i^- , \quad (10) \]

with \( c_i^- \) as the contrast input that acts as filling-in signal in the OFF domain. The diffusion coefficients, \( P_{ji}^- \), are chosen to ensure that the magnitude of cross influence of location \( j \) on location \( i \) and depend on boundary signals as \( P_{ji}^\pm = 1/(1 + \alpha e\mathcal{N}^\pm) \), where \( e \) is a constant, and \( w \) is the location boundary activation generated by the recurrent field (see Eq. (8)).

3.2.3. Luminance-driven pathway and multiple-scale brightness

3.2.3.6. Single scale brightness computation and brightness pooling (Single scale brightness; multiple scale brightness).
The filled-in activities in the ON and OFF domains are used in conjunction with activities in the segregated luminance-driven pathway to determine the single scale brightness prediction, \( u_i^b \). The basic idea is that the luminance-driven channel (\( u_i \) in Eq. (3)) provides a baseline of activity that can be modified by the equilibrated activity in the ON and OFF filling-in domains (\( v_i^+, v_i^- \) in Eqs. (9) and (10), respectively). The interactions are governed by the following differential equation:

\[ u_i^b = -Fu_i^b + s_i + Gu_i^+ + Hu_i^- , \quad (11) \]

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

The final brightness percept is obtained by averaging the outputs of the different spatial scales, \( b_i = (1/k) \sum u_i^b \), where \( k \) is the number of scales being employed. The activity \( b_i \) defines the final output of the model at location \( i \). This completes the description of the computational mechanisms of the model.

4. Filling-in homogeneous regions and its relation to regularization theory

In this section the filling-in process that is regulated by BCS signals will be analyzed. We are specifically interested in the relationship between the action of filling-in and the finding of a regularized solution for the brightness reconstruction problem. Utilizing elements of variational calculus it will be shown that there are particular conditions which have direct relations to regularizing the inverse problem of brightness reconstruction, given a particular structure of the input data. Under those conditions and for certain parameter settings, the ON and OFF filling-in processes (see Eqs. (9) and (10)) generate a smooth brightness function that is reconstructed from initial contrast measures at abrupt luminance transitions. The action of the parallel topographically organized boundary system can be thought of as an ‘external’ form-sensitive system that generates compartments of filling-in domains segregated by sharp contrast edges. Within these compartments the filling-in process can generate a regularized solution of the inverse problem of brightness reconstruction, given a dense field of filling-in generators. At spatial locations where there exist gradually changing transitions, the action of the boundary system generates local print signals to block activations in the diffusion system. Here, the logic of the strategy does not in general follow an optimization scheme in the framework of regularization theory.

4.1. Filling-in in continuous form

The notations for the diffusive filling-in process given in Eqs. (9) and (10) use an already discretized format in order to denote the evolution of activity at different sites of the network. However, in order to relate the filling-in equations to the framework of variational calculus for the regularization of ill-posed problems, we utilize a spatially continuous format to denote a partial differential equation:

\[ \bar{v}(x) = -Kv(x) + c(x) + \rho \partial^2 \bar{v}(x;w) \Delta v(x), \quad (12) \]

where \( \Delta \) denotes the Laplacian operator of diffusion equations (\( \partial^2/\partial x^2 \) in the one-dimensional case and \( \partial^2/\partial x^2 + \partial^2/\partial y^2 \) in the two-dimensional case). The Laplacian term is biased on the function \([4]

\[ \bar{P}(x; w) = \frac{1}{1 + \phi f(w(x))}, \quad (13) \]

which specifies the efficacy of the diffusion to be inversely proportional to the boundary activation at each spatial location. The type of function utilized here is similar to one also proposed by Perona and Malik [36] for local gradient control in anisotropic diffusion. The diffusion equation is biased on a source, \( c(x) \) that denotes the filling-in of local contrast activation in the ON or the OFF channel, and a sink, \(-v(x)\), that—in the absence of any input—forces a decay of the activation at rate \( K \).

For the proper scaling of the input activation \( \bar{c}(x) = Kc(x) \) (the filling-in signal)\footnote{This scaling is attained by the multiplication of the upper and lower saturation levels of the initial contrast signal.}, we can normalize Eq. (12) w.r.t. the
decay rate. We thus obtain
\[
\int \psi(x) = c(x) - \nu(x) + \rho \xi(x) \Delta \nu(x),
\]  
(14)
with \( \Psi(x) = (1/K) \mathcal{P}(x; w) \) as a variable modulation function. The steady-state solution of Eq. (14) identifies a state in the diffusion system in which the driving input and the spontaneous decay are balanced without further lateral spreading of site activation.

4.2. Inverse problems and regularization

Brightness reconstruction from local (contrast) estimates can be described as the general problem to find a solution for \( u = A^H d \) with \( A^H = (A^T A)^{-1} A^T \) given the initial mapping \( A \) and \( d = 0 \) of a function \( u \) of \( X \) into an observation \( d \) of \( Y \). Since the existence and uniqueness of a solution and its continuous dependence on the data cannot be guaranteed in any case, the inverse problem is ill-posed in the sense of Hadamard [54,55]. The solution to the problem has to be regularized such that proper constraints have to be imposed on the possible set of candidates in the function space of solutions. We select the function \( \hat{u} \) as the solution that minimizes the norm \( \| \Delta \hat{u} - d \| \) subject to the additional constraint of smoothness of \( \hat{u} \). Smoothness is characterized by a minimal derivative, e.g. of first order. Therefore, the overall goal is to minimize the constraint given by the local differences between the measured data and the reconstructed function values (data term) and the stabilizing functional imposed on the function (smoothness term). This results in the goal of minimizing the quadratic functional
\[
\| \Delta u - d \|^2 + \lambda \| P u \| ^2 \rightarrow \min,
\]  
(15)
where \( P \) denotes a 'constraint operator' (a mapping \( P: X \rightarrow Z \)) that stabilizes the solution for the inverse problem and \( \| \cdot \| \) represents a proper norm in the spaces \( Y \) and \( Z \), respectively [54]. For a detailed discussion of the formal mathematical background we refer to [55,56].

Utilizing the calculus of variations [57] the minimization of a functional, \( E(u) = \int_D E(u; du) + \lambda E_p(u) \rightarrow \min \), yields a solution function that realizes a compromise between the data term (‘closeness’ to the data) and the smoothness (or model) term. The model term represents the \( a \) \_priori model of the function searched for in the reconstruction process. Optimal functions have been identified as generalized splines [58]. A sub-class of this set of functions which allows for a direct physical interpretation of the behavior and result of the underlying system consists of elastic membranes and thin plates [59]; elastic membranes minimize the first variation in the function used for the reconstruction process.

4.3. Diffusive filling-in related to membrane regularization

The input to the diffusion system is provided by the segregated contrast measures generated by the initial center-surround processing stage that utilizes a shunting mechanism. We are specifically interested in the relation between the (continuous) filling-in equation and the minimization of quadratic functionals for regularization purposes. Therefore, for the moment, we assume the existence of a solution for the inverse problem of reconstructing brightness data from local contrast measures. The resulting functional for a membrane regularization of reconstruction from sparse generators alone reads for the one-dimensional case:\footnote{Since the model outlined in this article has been defined in one dimension, we consider the corresponding equations only. In general, however, the filling-in equation as well as the functionals subject to variations can be canonically extended to higher dimensions.}

\[
e(v) = \int F[v(x), v_x(x), x] \rightarrow \min,
\]  
(16)
with \( F = \kappa (v(x) - c(x))^2 + \lambda \Delta v(x) \); \( \kappa = \) local estimated data, \( v(\cdot) \) and its first-order derivative \( v_x(\cdot) \) is the function sought in the function space of the given inverse problem; \( \kappa \) is a local confidence measure with \( \kappa = 0 \) where there is no input generator from the initial contrast processing (see [60,61]); \( \lambda = \) regularization parameter. The necessary condition for the existence of a solution for this minimization problem is denoted by the corresponding Euler equation, which for \( F \) results in:

\[
k(x) (v(x) - c(x)) - \lambda \Delta v(x) = 0,
\]  
(17)
with \( \Delta = d^2/dx^2 \). The non-conservative diffusion \( (1/K) \Delta v(x,t) = c(x,t) - v(x,t) + \rho \xi(x)v(x,t) \) (Eq. (14)) defined for the filling-in of brightness equilibrates to a steady-state solution:

\[
v(x) - c(x) - \rho \xi(x) \Delta v(x) = 0.
\]  
(18)
Thus, the general outline of this equation is similar to the one derived for the linear quadratic functional in Eq. (17) setting \( \lambda = \rho \). However, the contribution of the Laplacian diffusion term in Eq. (14) is regulated by the external boundary system, for which modulation activity has been denoted by \( \xi(x) \). Furthermore, the term in Eq. (17) that measures data compatibility contains a confidence measure that is variable over the image domain. We discuss these issues below.

In order to properly deal with the modulation activity represented by \( \xi(x) \), we adapt the format of the function \( F \). We rewrite to get \( F = \kappa (v(x) - c(x))^2 + \rho \xi(x)v(x,t)^2 \). Now the variational problem of minimizing the corresponding functional results in:

\[
e(v) = \int F(v(x), v_x(x), x) \rightarrow \min
\]  
(19)
(compare with Ref. [62] for the route using controlled-continuity stabilizers). The corresponding Euler equation

\footnote{The notation of the data similarity term \( (v(x) - c(x)) \) can be generalized using a measurement functional \( \mathcal{L} \) represented by a pointwise \( k \)-th-order derivative of the reconstructed function \( v \).}
reads:
\[ k_i(v(x) - c_i(x)) - \rho(\xi_i v_i(x)) + \xi_i v_i(x) = 0. \]

Using \( \xi(x) = [K(1 + \epsilon f(\nu(x)))]^{-1} \) we expand this equation in order to directly include the activity of the boundary system and get
\[ k_i(v(x) - c_i(x)) - \rho \xi_i(x) \cdot \left( \frac{d}{\nu(x)} - \frac{d}{\nu_i(x)} \right) = 0. \]  

(20)

We observe that this condition does not in general coincide with the condition denoted by Eq. (17). However, correspondence is achieved:

1. when the contribution of the boundary system is switched off by setting \( \epsilon = 0 \), then we get \( \xi(x) = 1/K \) and Eq. (20) directly reduces to \( k_i(v(x) - c_i(x)) - (\rho/K) v_i(x) = 0 \), or

2. when a brightness compartment is covered with a constant distribution of boundary activation, i.e. \( \xi(x) = k \); in those cases the negative term in brackets vanishes due to the zero gradient in the boundary activity distribution and Eq. (20) reduces to \( k_i(v(x) - c_i(x)) - kpv_i(x) = 0 \).

Case 1 is only of minor interest since the elimination of any BCS contribution contradicts the key functionality of the architecture, namely the complementary action of two autonomous sub-systems to generate coherent percepts. We will now consider the requirements for Case 2 conditions.

The ability of the model to deal with abrupt as well as gradual transitions in the input luminance distribution has been demonstrated in this paper. So far, model simulations e.g. by [1, 63, 14] have only dealt with step contrast edges. The contour system in the extended model now automatically switches between modes to generate sharply localized boundary signals at luminance edges and gradual distributed responses at fuzzy boundaries [33]. For luminance edges sharply localized maximum amplitude boundary activations are generated via non-linear contrast processing and subsequent sharpening in a recurrent competitive field. Such boundary signals virtually suppress the lateral spreading of activity between neighboring sites in the diffusion system. In turn, these contour signals generate a compartmentation of the diffusion system into segregated filling-in domains. Thus, for example for a Mondrian, the image domain (denoted by \( R \)) is segregated into \( r \) regions \( R_r \), \( i = 1, \ldots, r \) with boundaries ideally represented by Dirac impulse functions. For an individual region \( R \), such as the central patch in the stimulus sketched in Fig. 18, we get \( \xi(x) = \Sigma \delta(x - x_i) \), with \( x_i \) to denote the locations of contrast edges. The equilibrium filling-in activity in the homogeneous region \( R \) is \( k_i(v(x) - c_i(x)) - kpv_i(x) = 0 \), i.e. we find the condition derived for Case 2. For gradual transitions in which the boundary system generates ‘fuzzy’ print signals, the filling-in activity cannot freely diffuse. In this case, the distributed activity yields a smooth function \( \xi(x) \) of non-localized boundaries. Therefore, the negative component in the smoothness term of Eq. (20) does not vanish. Thus, for the processing of gradual luminance transitions and discontinuous luminance gradients, the logic of the processing strategy does not follow an optimization scheme in the framework of regularization theory.

The data compatibility term in Eq. (17) contains a confidence measure. The purpose of this factor is to selectively evaluate closeness of the reconstruction at sparse locations where input data is available. The input sources \( c(x) \) to the individual diffusion layers can be approximately specified as:
\[ c^-(x) = \max \left[ \pm \frac{L_{\text{on}}(x)}{L(x)} \right] \]

with \( L(\cdot) \) the input luminance distribution. Input generators for the filling-in of compartments for homogeneous luminance regions are thus sampled only at places near boundaries. Under those conditions, the steady-state filling-in activity does not correspond to the regularized solution of a reconstructed brightness surface. In the filling-in case, we have \( k_i = 1 \) for each spatial location irrespective of the presence of any generator. Therefore, the similarity between the reconstruction and the initial measurement is not used as a penalty term in the sense of a regularization functional. However, previous approaches have combined contrast estimates with the low-pass filtered representation of the input luminance prior to the filling-in stage [14]. In the model proposed by Grossberg and Todorović [1] the raw activation distribution \( x^+ \) generated by the initial ON-center/OFF-surround shunting interaction has been used. In both approaches a dense input field is fed to the diffusion stage. The problem of brightness reconstruction now changes to the problem of proper smoothing the input activation generated by the initial processing stages. Hence, the models of brightness-from-luminance processing described in Ref. [1] and [12, 14] meet the requirements of filling-in to be compatible with \( v(x) - c(x) - \lambda \Delta v(x) = 0 \) denoted in Eq. (17) (\( k_i = 1 \)).

To summarize the discussion, we can conclude that the filling-in mechanism used as an elementary processing stage exhibits some similarities with the regularized solution for the problem of brightness reconstruction utilizing quadratic functionals. We have identified certain conditions to achieve such correspondence. This provides a guideline for further investigation into the development of regularization schemes for biological vision. In practical terms, the results of these findings are useful e.g. for an approximation of filling-in using neural network learning. The formal equivalence between regularization theory and radial basis

\footnote{For excitatory and inhibitory saturation levels of different magnitudes, it has been proved [12, 14] that the output of this processing stage contains a non-zero DC level that corresponds to a compressed and low-pass filtered representation of the input luminance distribution.}
function networks has been proven in Refs [64,65] (see also Ref. [66] for the incorporation of a classifier based on Gaussian basis functions into an ART-architecture). Training of such networks can be utilized to find approximate solutions for cases that correspond to the regularized solution for smoothed input, similar to the approximation of anisotropic diffusion described in Refs [67,68].

4.4. Mechanisms for brightness reconstruction complementary to regularization approaches

We have identified conditions in which the filling-in process, that has been suggested on the basis of empirical evidence, can generate a regularized solution of an inverse problem of brightness reconstruction from initial estimates (filling-in signals). The mechanism discussed in the above formulation realizes a physical membrane regularization for dense activities injected in regions that have been formed by the compartments generated in the BCS. In order to contrast our approach to the solution of an inverse problem of brightness reconstruction and more traditional approaches, we consider the following example of an Ehrenstein figure (see Fig. 8). Under normal viewing conditions (approx. 30 cm distance) the arrangement of radially oriented black bars on a light background appears together with an illusory circular disk in the center that is induced by the line ends and that appears brighter than the same luminance surround. The strength of the illusory contour that delineates the circle from the background depends on the orientation of the inducing lines (bars) relative to the local tangents of the illusory circle [7]. Now consider what is intended to be an inverse process of brightness reconstruction. The most direct approach for the solution of this problem would be to reconstruct a luminance distribution that minimizes the difference between the reconstruction and the original input in a pixel-by-pixel fashion thereby also achieving local smoothness in the reconstruction. In standard regularization terms [55] as well as in recent schemes for anisotropic diffusion (e.g. [36,38]) this would result in the reconstruction of the dark regions—representing the individual lines—and the reconstruction of the light homogeneous background region of constant luminance. This solution could be denoted as 'precise' as it treats the reconstruction problem in terms of measurement of light quantities.

Contrary to this approach, we acknowledge that the "... purpose of vision is to see things, not to judge light." (J. Lettvin, quoted from [69], p. 184). The optical input is segmented by the visual system into regions that are separated by perceptual boundaries. The mechanisms that are responsible for robust boundary formation must free the initial visual input data from its inherent uncertainties. The appearance of certain illusory percepts is thus an expression of adaptive mechanisms designed for the generation of informative visual representations of the world and can be used as probes of the machinery instantiated for this adaptive behavior of the brain. Along these lines, an alternative solution to the brightness reconstruction problem relies on the generation of a form-sensitive boundary representation generated by the separate BCS. The activity distribution generated by the BCS in turn defines the form-sensitive distribution of spatially variant regularization parameters, denoted by the function $f(x)$. In this framework, brightness filling-in is based on region compartments that are generated by contrast edges, the grouping of fragmented edge elements into dense boundary contours, and illusory outlines generated by long-range grouping.5

Fig. 9 shows how the BCS/FCS framework can provide the perceptually correct solution when processing an Ehrenstein stimulus, as simulated in [46]. Plot (a) shows the stimulus input configuration consisting of dark inducing lines on a bright background. The gray level activities of plot (b) code the resulting filtering activities of the ON (bright) and OFF (dark) channels. These provide inputs to separate ON and OFF filling-in domains, respectively. The boundary computations of the BCS (plot (c)) make use of the long-range co-operation processes of the full-blown two-dimensional implementation and are thus able to generate boundary signals corresponding to the circular illusory contour. Finally, plot (d) shows the final model output where gray level codes brightness intensifies. Note that the center (illusory) disk is brighter than the (same luminance) background, as in the illusion.

5 Classical approaches for the regularization of brightness

---

1 The complete specification of the BCS is not described in this paper, but only a simplified version for the one-dimensional case. For a full description of the machinery involved as well as the presentation of simulation results we refer the reader to, for example, Refs [6,7,42,71].
reconstruction use local energy measures of the reconstructed function. If these measures exceed a certain threshold, the reconstructed brightness function is allowed to cease and fracture at the expense of a penalty introduced by the discontinuities, finally minimizing global energy (see Refs [60,70]). The resulting functional often becomes non-convex so that additional mechanisms have to be incorporated in order to achieve globally optimum results (e.g. [72,73]). In our approach, the generation of the \( f(x) \) activities by an 'external' mechanism allows the energy function in the brightness system to be convex such that the system is guaranteed to converge to a unique optimum.

In the present approach, the reconstruction is performed separately in two parallel domains, namely ON and OFF. This design has been motivated by several psychophysical findings, as well as the result by Schiller and colleagues [74] that physiological ON and OFF channels remain segregated until visual area V1. The combination of ON and OFF filling-in solutions necessary for the determination of brightness can only, however, determine relative contrasts in the stimulus. Since the visual system is in fact sensitive to absolute luminance levels ([75–77]; see also [78]), the reconstructed brightness must be assigned to a precise, or absolute, magnitude. This is achieved in the current model.
by incorporating a segregated pathway that represents the low-pass filtered and non-linearly transformed input luminance distribution. This activity profile provides an ‘anchor’ of absolute brightness level.

5. Simulation results

In the present section, the behavior of the model described in Section 3 is illustrated through a set of computer simulations. Eqs. (8)–(10) are solved using a fourth-order Runge–Kutta numerical integration scheme until activities equilibrate. All other model stages are assumed to equilibrate rapidly and therefore the algebraic equations of the steady-state solutions are used directly. Note that steady-state solutions are employed only to optimize computation time given that we are interested in accounting here for ‘equilibrium’ perceptual states. In applications where one is interested in modeling temporally evolving events the dynamic form of the equations specifying the model can be applied directly.

We first present some simulation results that attempt to account for classic brightness effects, such as Mach bands. We then apply the model to some recent data by Georgeson [2] and discuss the implications of our approach to the question regarding the nature of the early visual representation. Finally, we present some simulations that illustrate how the present approach can be applied to real images.

5.1. Classical brightness effects

The first set of stimuli has been selected to demonstrate the capacity of the model to predict various challenging data. In particular, we investigate (a) low- and high-contrast missing fundamental stimuli, (b) the generation of Mach bands, and (c) the appearance of brightness staircases in comparison to sequences of Craik–O’Brien–Cornsweet ‘cusps’. The stimuli used are all one-dimensional slices through two-dimensional patterns.

The stimulus of the missing fundamental wave (studied first by Campbell et al. [79]) is generated by having the fundamental harmonic component removed from a square wave. This stimulus is particularly interesting as its appearance changes qualitatively depending on the amount of contrast. For high contrasts it has a vertical appearance, i.e. the modulation of the cusps is seen, whereas for low contrast, subjects report perceiving a square wave. Therefore, a one-parameter variation of the input stimulus (contrast magnitude) causes a non-linearity in the qualitative appearance of the final brightness percept. Our simulations correctly predict the effects of contrast on the percept, as shown in Fig. 10, with the high contrast missing fundamental wave shown in (a) and the low contrast missing fundamental wave shown in (b).

In order to provide a more detailed view of the behavior of the model, Fig. 11 shows the processing results at several model stages for the high contrast missing fundamental. Each plot displays the activity levels of a given model stage at equilibrium as a function of spatial position. Plots displaying ON and OFF activities use solid and dotted lines, respectively. The graphs that appear in the bottom row are those already shown in Fig. 10. Row 4 displays the activity profiles generated for the ON and OFF center-surround shunting networks and the segregated channels of ON and OFF contrast, as well as the luminance-driven pathway. The simulation of brightness processing employs three spatial scales. For rows 1–3, each row displays, from left to right: (i) the activities generated at the complex cell level; (ii) boundary processing in the recurrent field; (iii) contrast-driven diffusive filling-in in separate filling-in domains (ON and OFF activities are shown simultaneously; display of negative OFF activities for illustration purposes only); and (iv) the final scale-sensitive brightness appearances. Different scales appear in rows 1–3 (coarse to fine).

The combined behavior of the complex cell circuit and the recurrent field is such that sharp boundary signals are generated at locations of localized step contrasts, largely irrespective of the local contrast amplitude. Thus for both low- and high-contrast missing fundamental stimuli, the edge-like transitions of the stimuli produce sharp boundary signals. At the same time, shallow luminance variations produce weak complex cell responses that are not contrast-enhanced. Thus, the ridges and troughs in the high-contrast missing fundamental are able to survive as non-zero boundary signals. Overall, for the high-contrast stimulus, the recurrent field for boundary generation sharpens the large activities associated with the ‘edges’.

\footnote{All other stimuli have also been processed in three different spatial scales. For all simulations, a fixed set of parameters has been used. For a list of the parameter values as well as a discussion of the parameter choices the reader is referred to Ref. [33].}

\footnote{For the low-contrast missing fundamental, these boundary signals are not able to survive because the complex cell signals do not exceed the firing threshold.}
but also allows preservation of activities outside the domain of feedback (regulated by the spatial couplings utilized in the circuit). The preserved boundary activations are of sufficient amplitude that they are able to block, or trap, the diffusive filling-in process. Note that this only occurs in the two finest scales (scales 1 and 2). The coarsest scale (scale 3) does not contain surviving boundary signals and thus its associated filling-in activities more closely resemble a square-wave modulation. The final brightness prediction is taken as the average of the brightness activity profiles generated for individual scales.

The Mach band phenomenon illustrates another brightness perception effect that directly depends upon a one-parameter variation of the input luminance distribution (steepness of the ramp-like luminance transition). It further demonstrates the influence of scale in the current model. The strength of appearance of light and dark bands (Mach bands) at shallow transitions follows an inverted-U function such that for moderate widths the bands appear more vividly, becoming weaker as the ramp width is increased or decreased [80]. Moreover, the bands disappear altogether for very narrow ramp widths (less than 4–7.5 arcmin).
[81,80]. This observation rules out simple explanations that are based on center-surround processing as the responses in the ON and OFF contrast channels are maximal for step profiles (see Refs [29,30] for discussion).

In terms of the present model, depending on the steepness of the transition, the processing of ramps at small scales will not respond with juxtaposed ON and OFF responses after initial filtering and segregation of contrast channels. The simple cell response at the stage of oriented contrast detection will therefore not receive extra input activity that in turn would cause the recurrent circuit to sharpen the activity profile. As a consequence, the distribution of contrast activity is kept unsharpened so that the spatially extended boundary signals block the diffusion of the overshoots and undershoots produced by the initial center-surround filtering, thereby producing brightness overshoots and undershoots, or Mach bands. Fig. 12(a) shows the input luminance distribution and the final appearance of the brightness profile with overshoots and undershoots at the ‘knee’ points. In Fig. 12(b) a more detailed view of the processing is given. The plot in the upper row on the left shows an overlay of dense boundaries (dotted lines) and the contrast-driven feature signals (solid lines) generated for the full trapezoidal wave. A zoomed-in plot of activities generated for one ramp transition shows how boundaries can block diffusive filling-in over several network units (marked by dots). The bottom row shows final equilibrated activities for ON and OFF filling-in and the activation of the luminance-driven pathway. Fig. 12 further illustrates that, in terms of our proposal, the processing of spatial luminance profiles is not simply dependent on bases functions that account for brightness transitions with their final superposition.

The next example compares the results of processing sequences of luminance Craik–O’Brien–Cornsweet cusps with a luminance staircase containing plateau regions of corresponding width (see Fig. 13). The processing of both luminance profiles produces almost identical activity distributions in the ON and OFF contrast channels (not shown). Therefore, the distinct brightness appearance of both input profiles cannot be accounted for by the behavior of the contrast-driven channels alone. In the model, the behavior of the luminance-driven channel is responsible for the difference between the percepts. By definition, the luminance-driven pathway faithfully follows the modulation present in the input. As it provides a ‘baseline’ of activity that is combined with ON and OFF filling-in information to produce brightness, the model correctly predicts the perceived sequence of brightness steps for the luminance staircase11 (see Fig. 13(a)). On the other hand, a sequence of luminance cusps does not appear as a sequence of brightness steps, since the luminance-driven channel does not support such percept (see Fig. 13(b)). For this stimulus only the outermost regions differ in brightness; the internal stimulus segments end up having almost the same brightnesses.

5.2. Brightness perception and the nature of early visual representation

Luminance variations present in retinal images lead to perceptual brightness variations. In the past decade several research groups have proposed that the outputs of multiple channels tuned to orientation and spatial frequency are combined in order to produce a feature code for symbolic image description [83,84]. According to this view, human perception is based on the encoding of orientation, contrast and blur of features such as bars and edges. The central assumption of feature-encoding theories is that one of the major tasks of the visual system is to quickly extract the most salient information from an image. In the process, detail, such as the gradual variation of luminance, is lost (see, e.g. [39,85]).

While feature-encoding schemes have proved successful in predicting several psychophysical results in pattern discrimination (e.g. [86–88]) and may be useful to probe underlying even- and odd-symmetric neural mechanisms, we do not consider them appropriate formalisms to describe brightness perception. A vocabulary consisting of ‘edges’, ‘bars’, and ‘plateaus’ is simply not rich enough to describe

---

11 Note that the brightness profile is not perfectly uniform along the luminance plateaus. Smaller regions would produce even more uneven brightnesses, as predicted by the Chevreul illusion (see e.g. Ref. [82]).
real world scenes in general. Brightness gradients are ubiquitous in natural scenes and need to be accounted for by theories that model brightness data. They cannot be simply discarded, or ignored. For example, studies of the perception of lightness in three-dimensional ellipsoids showed an improvement in lightness constancy for such three-dimensional objects when compared to those obtained with flat rectangular shapes under similar illumination conditions [89]. Moreover, ample evidence indicates that smooth luminance gradients are required for the proper perception of three-dimensional shape and are employed by 'shape-from-shading' algorithms (e.g. [90–92]). We propose that filling-in models better account for brightness perception. Furthermore, unlike feature-encoding schemes they are capable of addressing the growing body of data on dynamic aspects of brightness perception (see Section 6).

To illustrate the above issues in a more direct form, we address the recent results of Georgeson [2] with $f + 3f$ compound gratings. Georgeson has recently investigated the perception of one-dimensional patterns composed of two spatial frequencies, $f$ and $3f$. The experiments were performed in order to test two alternative feature-based models, that proposed in Ref. [93] and the local energy model [84]. Fig. 14 shows three of the waveforms studied, with $0$, $\pi$ and $\pi/2$ phase relationship. The model proposed by Georgeson is able to indicate the position of lines and edges and predicts, for example, that 12 features (six lines, or bars, and six edges) are present in each cycle of the waveform with 0 phase. At the same time, for this pattern, the energy model predicts only two features per cycle (corresponding to the mean-luminance cross-over points). Georgeson [2] reports that in a psychophysical task with $f + 3f$ patterns where subjects were asked to indicate the center

Fig. 13. Sequence of local contrast transitions: (a) luminance staircase distribution with the input distribution, the activity profile in the luminance-driven channel, and the final brightness prediction; and (b) sequence of luminance cusps (plots from left to right according to the luminance staircase in (a)).

Fig. 14. Compound gratings ($f + 3f$) studied by Georgeson [2]. The stimuli are generated by summing the two component waveforms with a given phase relationship.
position of the perceived bar and edges, for a 0 phase stimulus, subjects marked all 12 features predicted by his model.

Although subjects were able to indicate the positions of bars and edges more or less reliably, the question arises: do they employ the outputs of edge and line detectors of the type proposed by Georgeson? As discussed, the present filling-in proposal suggests that boundary-regulated diffusion is responsible for the perception of brightness. Instead of a representation in terms of a discrete sequence of events (or symbols), a representational medium that is spatially organized—through the use of spatially organized fields of activity accounts for the geometric structure of percepts. Fig. 15 shows the results of computer simulations of the filling-in model for $f + 3f$ compound gratings as shown in Fig. 14. The model describes the visual appearance well in all three cases. Moreover, the model correctly accounts for the striking mismatch between luminance and brightness relationships. For example, for the 0 phase compound, although the inner regions of the high-luminance and low-luminance portions of the waveform have greatly different luminances (see arrows), the corresponding brightnesses are very similar (see arrows), and for some stimulus configurations, an actual brightness reversal occurs (see Fig. 14). The 0 phase compound illustrates a luminance–brightness mismatch that is perhaps even more extreme than the standard example of simultaneous contrast where the same mid-gray appears differently bright on white and black surrounds. Compare also the brightness relationships between the positions indicated by arrows for the $\pi$ and $\pi/2$ compounds.

While Georgeson’s model [2,93] can perhaps explain the hills and valleys of brightness present in the stimuli—due to the triggering of combinations of edges and bars—its linear mechanisms cannot easily predict the brightness relationships within a stimulus, such as the brightness reversal of the 0 phase stimulus. For such stimulus, the low contrast ‘ripples’ (see arrows for the 0 phase stimulus) do not produce bar and edge features of enough contrast to overcome the large brightness steps associated with the ‘edges’ at the mean-luminance cross-over points.

The computations of the filling-in model detailed in Fig. 16 show some of the key computational steps in the generation of the final brightness for the 0 phase compound. Note that an initial non-linearity such as encoded in our ON-channel is not sufficient to produce the observed brightness reversal—although it helps. In the model the reversals originate from the interactions of ON and OFF contrast-driven signals that undergo diffusion in separate filling-in domains. Similar analyses apply to the other stimuli shown in Fig. 15.

5.3. Natural images and automated inspection

Although the modeling efforts presented herein are largely motivated by psychophysical and physiological data, the model presented can also be used with natural images and applied to image processing applications of
automated inspection. In a full two-dimensional implementation, Grossberg and colleagues [71] have studied the performance of several BCS/FCS variants applied to synthetic aperture radar (SAR) imagery. One of the schemes investigated employed a full-blown two-dimensional BCS together with the present proposal of treatment of feature signals, namely, the use of contrast-driven and luminance-driven channels. The behavior of the system is illustrated in Fig. 17. The upper row (panels (a), (b)) displays the raw input image and the result of a signal transformation that uses a logarithmic mapping function. The sensor signal is typically of the range of five orders of magnitude and is corrupted by speckle noise of high amplitude. The resulting image functions can therefore be characterized as having low signal-to-noise ratio. The result of processing in the BCS is shown in (c) and the final filled-in brightness distribution in (d). Again, brightness is derived from a 'baseline' of activity (in the luminance-driven pathway) that is combined with the equilibrated ON and OFF filling-in information. The results qualitatively illustrate that the present approach can be used in conjunction with pattern classification methods for successful image recognition given that the model is able to highlight important image regions, such as the trees and the bridge overpass. For

Fig. 17. Processing result for an application of the BCS/FCS model in the technical domain of automated inspection: (a,b) input synthetic aperture radar (SAR) image of upstate New York scene consisting of highway with bridge overpass, and its logarithmic transformed representation (note that the raw sensor signal is typically of the range of five orders of magnitude and disturbed by speckle noise of high amplitude); (c) boundaries generated in the BCS to form compartments of surface-related compartments, and (d) filled-in result generated by the FCS that forms a representation of the final brightness distribution that can be used for further inspection (reprinted with permission from [71]).
further details of the description and a comparative evaluation of the approach, see Ref. [71].

6. Temporal dynamics of brightness perception

There is an enormous literature on the spatial variables determining brightness perception. In contrast, there have been considerably fewer studies investigating temporal variables (see Ref. [94] for a summary). This gap is being bridged by the recent work of Paradiso and colleagues ([34, 95–98]; see also [99] and more recently [100], the latter for a direct demonstration of brightness induced modulation of neural firing in V1). For example, Paradiso and Nakayama investigated the role of edge information in determining the brightness of homogeneous regions by studying the temporal evolution of brightness. They reasoned that if filling-in processes are responsible for brightness perception, it would be possible to demonstrate their existence by interrupting it. Their results are consistent with the hypothesis that brightness signals are generated at the borders of their target stimuli and propagate inward at a rate of 110–150 deg s⁻¹. These experimental results have subsequently been modeled by Arrington [65] who showed that Grossberg and Todorović’s filling-in model can account for the results.

Rossi and Paradiso [98] studied the temporal dynamics of brightness induction (contrast). In displays such as that shown in Fig. 18, subjects were asked to indicate the lowest temporal frequency of luminance modulation at which the brightness of the central stripe appeared to stop modulating between light and dark; note that the luminance of the central patch was kept constant. Their main finding was that as the width of the central patch increased, the lowest temporal frequency indicated by the subjects decreased. In other words, as the width increased, the luminance modulation of the surround (such that no induced modulation was seen in the center) was slower—a relatively larger area could not be induced at a rate as fast as a small area.

Rossi and Paradiso [98] compared the results of induced brightness to the perception of direct brightness changes of the central patch when its luminance was explicitly modulated up and down. Their finding was that induced brightness was much slower than direct brightness changes. They state (p. 12, italics added): ‘‘The simplest way to integrate this process [of filling-in] with the results reported here is if the brightness of a stimulus involves two processes, one based on the response to its luminance and a second based on the luminance of other areas or the contrast at the borders between areas.’’

We concur with Rossi and Paradiso that two processes are responsible for brightness perception, namely, processes related to our contrast-driven and luminance-driven pathways (see Fig. 18). In order to account for the reported differences between induced and direct brightness changes within the framework of our model, we assume that the time constants of the two pathways are different. In particular, the temporal dynamics of the luminance-driven channel is assumed to be faster than that of the contrast-driven pathway. Thus, the direct luminance changes of the central patch in Rossi and Papadiso’s settings are able to have a much

![Fig. 18. Temporal dynamics of brightness perception: (a) representation of the stimulus used in Ref. [98]. The luminance of the surrounding patches was modulated up and down with varying temporal frequencies. The luminance of the central patch was kept constant throughout a trial, while its appearance changed due to induced brightness contrast in counter phase with the surrounding modulation; (b) one-dimensional cross-section of the luminance profile of the stimulus in (a) at a time when the surround luminance was lower than the fixed center luminance; (c) schematic representation of the low-pass signal and contrast signals that undergo diffusive filling-in (only the ON signals are shown) associated with the stimulus in (b). The time course of the diffusive mechanisms is assumed to be slower than that of the luminance-driven pathway in (c) (see text).](image-url)
quicker effect on perceived brightness than the induced ones which depend on the slower filling-in process which lags behind.

7. Conclusions

In this contribution we have presented a neural architecture of brightness perception that can account for a wide variety of challenging data. The processing mechanisms are motivated by a vast amount of empirical findings. The model is in the tradition of filling-in theories with segregated sub-systems, namely BCS and FCS (as part of the FACADE theory developed by Grossberg [10]), that involve complementary processes to achieve form-sensitive segmentation of surface outlines and the determination of surface attributes. The concepts of the architecture include containment of featural diffusion activity by boundary contrast signals, the processing of sharply localized and more shallow spatial luminance transitions, and the modulation of boundary signals by a feedback circuit.

Simulations of new findings in brightness perception were used to challenge the widespread view espoused by current early vision models that important visual information is coded though a set of discrete, symbolic, visual ‘events’. We propose that our approach better captures the spatial and temporal aspects of brightness perception in particular, and of early vision in general.

A key aspect discussed in the paper focuses on the relation between steady-state activity of the FCS—which is regulated by the activation of the topographically organized BCS—and the regularized solution of ill-posed problems, in particular brightness reconstruction from sparse generators. This investigation was motivated by the observation of structural similarity between the equilibrium solution of filling-in and the Euler equation for a first-order spline (or membrane) regularization of a brightness (surface) reconstruction problem. It has been shown that the formal equivalence of the schemes only holds for particular stimulus conditions. The necessary conditions for such correspondence are: (i) the layout of homogeneous regions segregated by sharp contrast edges, such that compartments for diffusion are segregated by impulse-like BCS boundaries; and (ii) the input of a dense field of initial measurements to the filling-in domains. The BCS/FCS approach offers an alternative to the more traditional processing schemes proposed for early vision. In our framework—as part of the FACADE theory—the processes for generating a ‘web’ of boundary signals regulate the construction of a brightness representation in homogeneous regions in a context-sensitive fashion. Two types of complementary processes are handled by the BCS and the FCS, respectively. This allows for an adaptive context-sensitive reorganization process to guide the form-oriented reconstruction of feature quantities. The overall functionality offers an alternative architecture framework for geometry-driven diffusion and coherent contour processing and grouping.

Acknowledgements

We would like to express our gratitude to the three anonymous reviewers for their constructive criticism and suggestions for improvement. Heiko Neumann and Luiz Pessoa were supported in part by the German-Brasilian Academic Research Collaboration Program (DAAD-CAPES, Probral 1997, Project No. 70 712), Ennio Mingolla was supported in part by the Office of Naval Research (ONR N00014-94-1-0597).

References

[16] C. Enroth-Cugell, J. Robson, Functional characteristics and diversity
[56] R. Sprengel, Entwurf und Analyse nichtlinearer Diffusionsverfahren für die Bildverarbeitung, Infn, St Augustin, 1996.


