



University of Pennsylvania  
ScholarlyCommons

---

Departmental Papers (BE)

Department of Bioengineering

---

April 2004

# Optic Nerve Signals in a Neuromorphic Chip I: Outer and Inner Retina Models

Kareem A. Zaghloul  
*University of Pennsylvania*

Kwabena A. Boahen  
*University of Pennsylvania*, [boahen@seas.upenn.edu](mailto:boahen@seas.upenn.edu)

Follow this and additional works at: [http://repository.upenn.edu/be\\_papers](http://repository.upenn.edu/be_papers)

---

## Recommended Citation

Zaghloul, K. A., & Boahen, K. A. (2004). Optic Nerve Signals in a Neuromorphic Chip I: Outer and Inner Retina Models. Retrieved from [http://repository.upenn.edu/be\\_papers/14](http://repository.upenn.edu/be_papers/14)

Copyright 2004 IEEE. Reprinted from *IEEE Transactions on Biomedical Engineering*, Volume 51, Issue 4, April 2004, pages 657-666.  
Publisher URL: <http://ieeexplore.ieee.org/xpl/tocresult.jsp?isNumber=28546&puNumber=10>

This material is posted here with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of the University of Pennsylvania's products or services. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by writing to [pubs-permissions@ieee.org](mailto:pubs-permissions@ieee.org). By choosing to view this document, you agree to all provisions of the copyright laws protecting it.

This paper is posted at ScholarlyCommons. [http://repository.upenn.edu/be\\_papers/14](http://repository.upenn.edu/be_papers/14)  
For more information, please contact [libraryrepository@pobox.upenn.edu](mailto:libraryrepository@pobox.upenn.edu).

---

# Optic Nerve Signals in a Neuromorphic Chip I: Outer and Inner Retina Models

## Abstract

We present a novel model for the mammalian retina and analyze its behavior. Our outer retina model performs bandpass spatiotemporal filtering. It is comprised of two reciprocally connected resistive grids that model the cone and horizontal cell syncytia. We show analytically that its sensitivity is proportional to the space-constant-ratio of the two grids while its half-max response is set by the local average intensity. Thus, this outer retina model realizes luminance adaptation. Our inner retina model performs high-pass temporal filtering. It features slow negative feedback whose strength is modulated by a locally computed measure of temporal contrast, modeling two kinds of amacrine cells, one narrow-field, the other wide-field. We show analytically that, when the input is spectrally pure, the corner-frequency tracks the input frequency. But when the input is broadband, the corner frequency is proportional to contrast. Thus, this inner retina model realizes temporal frequency adaptation as well as contrast gain control. We present CMOS circuit designs for our retina model in this paper as well. Experimental measurements from the fabricated chip, and validation of our analytical results, are presented in the companion paper [Zaghloul and Boahen (2004)].

## Keywords

adaptive circuits, neural systems, neuromorphic engineering, prosthetics, vision

## Comments

Copyright 2004 IEEE. Reprinted from *IEEE Transactions on Biomedical Engineering*, Volume 51, Issue 4, April 2004, pages 657-666.

Publisher URL: <http://ieeexplore.ieee.org/xpl/tocresult.jsp?isNumber=28546&puNumber=10>

This material is posted here with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of the University of Pennsylvania's products or services. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by writing to [pubs-permissions@ieee.org](mailto:pubs-permissions@ieee.org). By choosing to view this document, you agree to all provisions of the copyright laws protecting it.

# Optic Nerve Signals in a Neuromorphic Chip I: Outer and Inner Retina Models

Kareem A. Zaghoul, *Member, IEEE*, and Kwabena Boahen\*

**Abstract**—We present a novel model for the mammalian retina and analyze its behavior. Our outer retina model performs band-pass spatiotemporal filtering. It is comprised of two reciprocally connected resistive grids that model the cone and horizontal cell syncytia. We show analytically that its sensitivity is proportional to the space-constant-ratio of the two grids while its half-max response is set by the local average intensity. Thus, this outer retina model realizes luminance adaptation. Our inner retina model performs high-pass temporal filtering. It features slow negative feedback whose strength is modulated by a locally computed measure of temporal contrast, modeling two kinds of amacrine cells, one narrow-field, the other wide-field. We show analytically that, when the input is spectrally pure, the corner-frequency tracks the input frequency. But when the input is broadband, the corner frequency is proportional to contrast. Thus, this inner retina model realizes temporal frequency adaptation as well as contrast gain control. We present CMOS circuit designs for our retina model in this paper as well. Experimental measurements from the fabricated chip, and validation of our analytical results, are presented in the companion paper [Zaghoul and Boahen (2004)].

**Index Terms**—Adaptive circuits, neural systems, neuromorphic engineering, prosthetics, vision.

## I. RETINOMORPHIC SYSTEMS

THE RETINA, one of the best studied neural systems, is a complex piece of biological wetware designed to signal the onset or offset of visual stimuli in a sustained or transient fashion [25]. To encode these signals into spike patterns for transmission to higher processing centers, the retina has evolved intricate neuronal circuits that capture information contained within natural scenes efficiently [27]. This visual preprocessing, realized by the retina, occurs in two stages, the outer retina and the inner retina. Each local retinal microcircuit plays a specific role in the retina's function, and neurophysiologists have extracted a wealth of data characterizing how its constituent cell types contribute to visual processing. These physiological functions can be replicated in artificial systems by emulating their underlying synaptic interactions. In this paper, we present a novel model for how the outer and inner retina process visual

information, implement this model in silicon, and analyze its behavior. In the accompanying paper [30], we characterize how our silicon chip processes visual information and validate our analytical results.

Attempts have been made to duplicate neural function at high levels of abstraction because detailed knowledge of neural circuits was unavailable, but these neuro-inspired systems proved ill-suited to direct hardware implementation because of the mathematical operations they require. Neuromorphic systems, on the other hand, which are constructed from neuron-like elements based on physiology, connected together into local microcircuits based on anatomy, can now duplicate certain brain computations in silicon [9], [19].

The advantage these neuromorphic chips offer over software models is their ability to replicate neural computations in real-time at low power. This implementation efficiency translates to a greater ability to explore model parameters to further understand the underlying biological system and, by communicating with other neuromorphic chips, a greater ability to replicate more complicated neural systems. Furthermore, with the continued miniaturization of silicon technology and with a greater understanding of neural computations, attention has recently been shifted to developing prosthetic devices capable of restoring neural function [6], [32]. While these efforts have been relatively successful in the cochlea, fully implantable solutions remain elusive [21]. Neuroprostheses, such as an intraocular retinal implant, may benefit from a fully integrated approach made possible by the greater autonomy and functionality of neuromorphic systems.

The most recent effort to model retinal processing in silicon incorporated outer retina circuitry as well as bipolar cell (BC) and amacrine cell interactions in the inner retina [3]. This outer retina circuit took the difference between the photoreceptor signal and its spatiotemporal average, computed by a network of coupled lateral elements [horizontal cells (HCs)], through negative (inhibitory) feedback. Cone-coupling in this model attenuated high-frequency noise to realize a spatial bandpass filter and dynamic range was extended by implementing local automatic gain control. Furthermore, this model used HC activity to boost cone to HC excitation [3], which eliminated receptive-field expansion and temporal instability. This gain boosting mechanism has some physiological basis since glutamate release from cones is modulated by HC hemichannels [13], and may be enhanced further by HC autofeedback [14]. Our outer retina design is based on this approach [3].

The BC and amacrine cell interactions introduced by this earlier design represented a model for inner retina processing that, much like the circuit discussed here, attempted to capture

Manuscript received September 6, 2002; revised June 25, 2003 This work was supported in part by the National Institutes of Health (NIH) under Vision Training Grant T32-EY07035 and in part by the Whitaker Foundation under Grant 37005-00-00. The work of K. A. Zaghoul was completed while he was with the Department of Neuroscience, University of Pennsylvania, Philadelphia, PA 19104 USA. *Asterisk indicates corresponding author.*

K. A. Zaghoul is with the Department of Neurosurgery, University of Pennsylvania, Philadelphia, PA 19104 USA (e-mail: kazaghoul@uphs.upenn.edu).

\*K. Boahen is with the Department of Bioengineering, 120 Hayden Hall, 3320 Smith Walk, University of Pennsylvania, Philadelphia, PA 19104 USA (e-mail: boahen@seas.upenn.edu).

Digital Object Identifier 10.1109/TBME.2003.821039

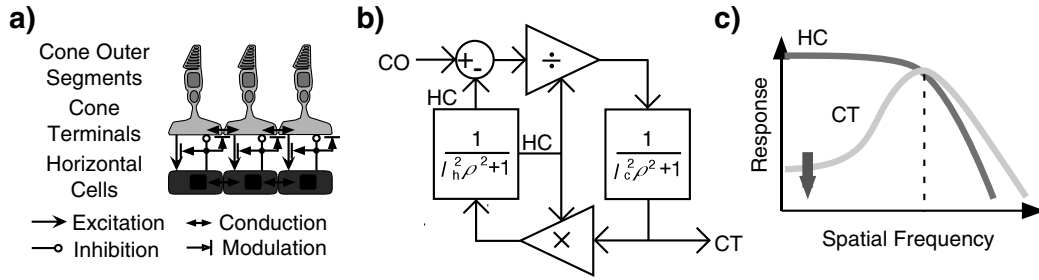


Fig. 1. Outer retina model. (a) Neural circuit: CTs receive a signal that is proportional to incident light intensity from the cone outer segment (CO) and excite HCs). HCs spread their input laterally through gap junctions, provide shunting inhibition onto CT, and modulate cone coupling and cone excitation. (b) System diagram: Signals travel from CO to CT and on to HC, which provides negative feedback. Excitation of HC by CT is modulated by HC, which also modulates the attenuation from CO to CT. These interactions realize local automatic gain control in CT and keep receptive field size invariant. Both CT and HC form networks, connected through gap junctions, that are governed by their respective space constants,  $l_c$  and  $l_h$ . (c) Frequency responses: Both HC and CT low-pass filter input signals, but because of HCs larger space constant,  $l_h$ , HC inhibition eliminates low-frequency signals, yielding a spatially bandpass response in CT.

temporal adaptation through adjustment of amacrine cell feedback inhibition. However, this earlier design did not include the retina's complementary push-pull architecture [25]. Hence, at low-frequency stimulation, dc current levels tended to rise and, thus, modulation of the feedback loop gain was not realized as intended [3]. Furthermore, this previous implementation did not replicate synaptic interactions at the ganglion cell (GC) level and, thus, did not faithfully capture the behavior of the retina's major output pathways [3].

In this paper, we present a novel model for processing in the mammalian retina that addresses the design flaws in [3] and extends that model to include GC level synaptic interactions. Our model is based on the functional architecture of the mammalian retina, and on physiological studies that have characterized the computations performed by the retina. Additionally, we build upon these earlier studies to present new circuit components that may account for some of the computations found in the retina. While the ultimate goal of our modeling effort is to match the mammalian retina, a detailed comparison between our model's outputs and the retina's outputs will be presented in a forthcoming paper.

The present paper presents a detailed analysis of our mathematical model and describes its silicon implementation. The remainder of this paper is organized as follows. In Section II, we present our outer retina circuit model, briefly revisiting a circuit design described earlier [3], and our model of BC circuitry that divides signals into ON and OFF channels. In Section III, we introduce a novel model for the inner retina that adjusts its loop gain, allowing it to adapt to different input power spectra. Finally, in Section IV, we discuss some of the assumptions underlying our analysis and justifications for our model. Section V concludes the paper. An accompanying paper presents experimental results that characterize our model's behavior.

## II. OUTER RETINA MODEL

Our model for the outer retina's synaptic interactions, which realize spatiotemporal bandpass filtering and local gain control, was described previously [3]. The synaptic interactions underlying our model are repeated in the Fig. 1(a) for reference. Briefly, photons incident on the cone (CO) reduce current flow to its cone terminal (CT), causing a hyperpolarization and a decrease in neurotransmitter release. CTs excite HCs which apply

shunting (or attenuating) inhibition on to CT [25]. This reciprocal interaction between cones and HCs, and electrical coupling of both cell types to their neighbors through gap junctions [25], creates a spatiotemporally bandpassed signal at CT that adapts to light intensity [26].

From the synaptic interactions, we derived the block diagram shown in Fig. 1(b) by modeling both the cone and HC networks as spatial low-pass filters. We can derive the system level equations that describe how CO determines HC and CT activity, represented by  $i_{hc}$  and  $i_{ct}$  respectively, from this block diagram. These equations, reproduced from [3], are

$$i_{hc}(\rho) = \left( \frac{A}{(l_c^2 \rho^2 + 1)(l_h^2 \rho^2 + 1) + \frac{A}{B}} \right) \frac{i_{co}}{B} \quad (1)$$

$$i_{ct}(\rho) = \left( \frac{(l_h^2 \rho^2 + 1)}{(l_c^2 \rho^2 + 1)(l_h^2 \rho^2 + 1) + \frac{A}{B}} \right) \frac{i_{co}}{B} \quad (2)$$

where  $l_c$  and  $l_h$  are the cone- and horizontal-network space-constants, respectively,  $B$  is the attenuation from CO to CT, and  $A$  is the amplification from CT to HC. HCs have stronger coupling in our model (i.e.,  $l_h$  is larger than  $l_c$ ), causing their spatial low-pass filter to attenuate lower spatial frequencies. Thus, HC low-pass filters the signal while CT bandpass filters it, as shown in Fig. 1(c), with the same corner frequency. We can determine this peak spatial frequency,  $\rho_A$ , by taking the derivative of CT's system equation and equating it to zero. We find

$$\rho_A = \left( \sqrt{\frac{A}{B}} - \frac{l_c}{l_h} \right)^{1/2} \frac{1}{\sqrt{l_c l_h}}. \quad (3)$$

To realize local automatic gain control in our model, we set HC activity proportional to intensity and use this activity to modulate CO to CT attenuation,  $B$ , by changing cone-to-cone conductance. This modulation adapts cone sensitivity to different light intensities [5]. Furthermore, to overcome the expansion in receptive-field size that this modulation caused in earlier designs, we complemented HC modulation of cone gap-junctions with HC modulation of cone leakage conductance, through shunting inhibition, making  $l_c$  independent of luminance. We also compensated for the change in loop-gain with HC modulation of cone excitation, by keeping  $A$  proportional to  $B$ , thus fixing  $\rho_A$ .

We can determine how the CT activity produced by an edge depends on these parameters by assuming that the dominant contribution is from the peak and, hence, we insert the value of  $\rho_A$  into CT's system equation

$$i_{ct}(\rho_A) = \frac{\frac{l_h}{l_c}}{\frac{l_h}{l_c} + 2 - \frac{l_c}{l_h}} \frac{i_{co}}{B}$$

where we have set  $A = B$ , assuming a proportionality constant of unity for simplicity.  $i_{co}$  represents the frequency component of  $i_{co}$  at  $\rho_A$ . Hence, CT activity will be proportional to contrast if  $B$ , which is determined by HC activity, is proportional to local intensity. We can determine the local HC activity and, hence,  $B$ , by summing responses to the dc component (i.e.,  $\rho = 0$ ) and the peak spatial frequency (i.e.,  $\rho = \rho_A$ )

$$B = \frac{i_{hc}(0) + i_{hc}(\rho_A)}{I_m} = \left( \frac{\langle i_{co} \rangle}{2} + \frac{i_{co}}{\frac{l_h}{l_c} + 2 - \frac{l_c}{l_h}} \right) \frac{1}{I_m}$$

where  $I_m$  is the intensity level at which the CT attenuation,  $B$ , is unity, and where  $\langle i_{co} \rangle$  represents the local average of input  $i_{co}$  and  $A = B$  as before. Therefore

$$i_{ct}(\rho_A) = \frac{\left( \frac{l_h}{l_c} \right) \tilde{i}_{co} I_m}{\tilde{i}_{co} + \left( \frac{l_h}{l_c} + 2 - \frac{l_c}{l_h} \right) \frac{\langle i_{co} \rangle}{2}}. \quad (4)$$

From (4), we see that CT activity is primarily determined by spatial contrast,  $\tilde{i}_{co}/\langle i_{co} \rangle$ , saturates when increasing signal fluctuation causes this ratio to become large, and is entirely independent of absolute luminance. In the case when signal fluctuation,  $\tilde{i}_{co}$ , is large, increasing the ratio  $l_h/l_c$  increases the level of cone saturation. In the limit where  $l_h \gg l_c$ ,  $l_h$  and  $l_c$  have negligible effect on the system's peak sensitivity at low contrasts (i.e.,  $\langle i_{co} \rangle \gg \tilde{i}_{co}$ ).

CT activity governed by (4) in our model exhibits behavior remarkably similar to the mammalian retina. Physiologists have found that cone responses as a function of intensity of light stimulation against a background, or contrast, are described by a simple equation

$$V = \frac{IV_m}{I + \sigma}$$

where  $V$  is the peak amplitude of the cone response produced by a given level of stimulating light intensity,  $I$ .  $V_m$  is the maximum response and  $\sigma$  is the background intensity. The response reaches half of the maximum when  $I = \sigma$  [20]. This adaptive behavior is preserved across five decades of background intensities. In our model, cone responses take on the same S-shaped response with contrast when  $l_h \approx l_c$ .

The complete outer retina CMOS circuit that implements local gain control and spatiotemporal bandpass filtering, while using HC modulation to maintain invariant spatial filtering and temporal stability, is shown in Fig. 2(a) (adapted from [3]) for two adjacent nodes. Photocurrents discharge  $V_c$ , increasing CT activity, and excite the HC network through an nMOS transistor followed by a pMOS current mirror. HC activity, represented by  $V_h$ , modulates this CT excitation and inhibits CT activity—by dumping this same current on to  $V_c$ . Cone signals,  $V_c$ , are electrically coupled to the six nearest neighbors

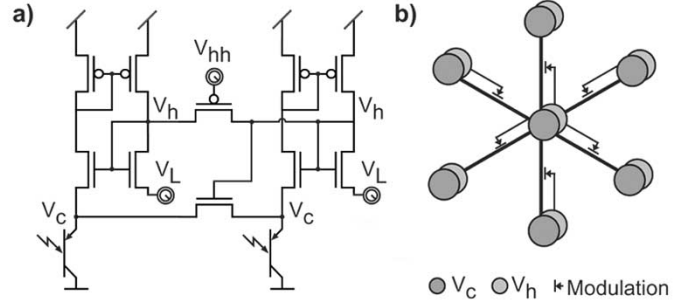


Fig. 2. Outer retina circuit. (a) Circuit: A phototransistor draws current through an nMOS transistor whose source is tied to  $V_c$  and whose gate is tied to  $V_h$ . This transistor passes a current proportional to the product of CT and HC activity, thus modeling HC shunting inhibition. In addition, this current, mirrored through pMOS transistors, dumps charge on the HC node,  $V_{hh}$ , modeling CT excitation of HC, and HC autofeedback.  $V_L$  sets the mean level of  $V_c$ , governing CT activity. (b) CT coupling is modulated by HC activity. A HC node [ $V_h$  in (a)] gates three of the six transistors coupling its CT node [ $V_c$  in (a)] to its nearest neighbors.

through nMOS transistors whose gates are controlled locally by  $V_h$  [Fig. 2(b)]. These cone signals gate transistors that feed currents into the BC circuit, such that increases in  $V_c$ , which tracks the level we set  $V_L$  to, increase the BC activity. HC signals also communicate with one another, through pMOS transistors, but this coupling is modulated globally by  $V_{hh}$ , since interplexiform cells that adjust HC coupling are not present in our chip [15].

We rectify the cone signal using the circuit shown in Fig. 3 (described in [29]) to model complementary signaling by BCs. Briefly, CT activity is represented by a current,  $I_c$ , that is inversely related to  $V_c$  (from our description above). And we compare  $I_c$  to a reference current,  $I_r$ , that is in turn inversely related to a reference bias,  $V_{ref}$ .<sup>1</sup> Therefore, we define two new currents,  $I'_c \propto 1/I_c$  and  $I'_r \propto 1/I_r$ , to simplify our description of the bipolar circuit. Mirroring these currents on to one another preserves their differential signal,  $I'_r - I'_c$ , which equals  $I_{ON} - I_{OFF}$  [29].  $V_{bq}$  in our circuit sets quiescent current levels—decreasing  $V_{bq}$  causes more current to flow through both channels—making rectification in our circuit incomplete. We set  $I'_r$  equal to the mean value of  $I'_c$  such that the difference is positive when light is brighter ( $I'_c$  decreases) and negative when light is dimmer ( $I'_c$  increases). We can see that as cone activity increases ( $V_c$  falls, translating to a rise in  $I_c$ , and a corresponding fall in  $I'_c$ ), current is diverted through the ON channel, but this current level quickly saturates. On the other hand, as cone activity decreases ( $V_c$  rises, translating to a fall in  $I_c$ , and a corresponding rise in  $I'_c$ ), current flows through the OFF channel and increases as the reciprocal of cone activity [29]. Thus, our bipolar circuitry divides signals into ON and OFF channels, as expected, but the division is not symmetric because one side of our circuit is fixed to the reference bias,  $V_{ref}$ .

### III. INNER RETINA MODEL

Our model of inner retina neurocircuitry, which realizes low-pass and high-pass temporal filtering, adapts temporal

<sup>1</sup>In practice, we cannot simply tie  $V_{ref}$  to  $V_L$  because mean CT activity,  $V_c$ , is slightly higher than  $V_L$ , due to drain-voltage mismatch between nMOS and pMOS transistor-pairs in the outer retina circuit.

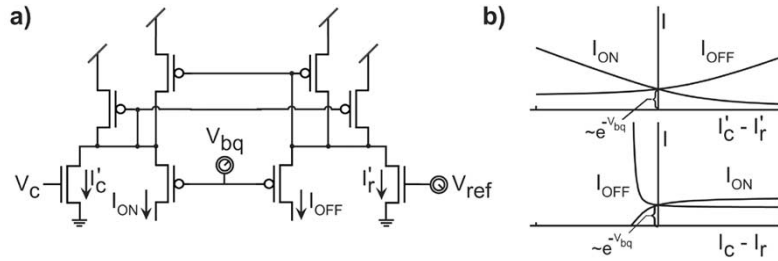


Fig. 3. BC circuit. (a) CT activity,  $V_c$ , drives a current,  $I'_c$ , that is compared to a reference current,  $I'_r$ , driven by a reference bias,  $V_{ref}$ . By mirroring these currents on to one another, the difference between them is encoded by either  $I_{ON}$  or  $I_{OFF}$ .  $V_{bq}$  determines the level of residual dc signal present in  $I_{ON}$  and  $I_{OFF}$ . (b) The difference between  $I'_c$  and  $I'_r$  determines differential signalling in  $I_{ON}$  and  $I_{OFF}$  (top). When  $V_c = V_{ref}$  (i.e.,  $I'_c = I'_r$ ), residual dc currents are proportional to  $e^{-V_{bq}}$ . Directly plotting the difference between cone activity,  $I_c$ , and  $I_r$  yields the curves on bottom. Increases in cone activity cause ON currents to saturate while decreases in cone activity cause OFF currents to increase reciprocally.

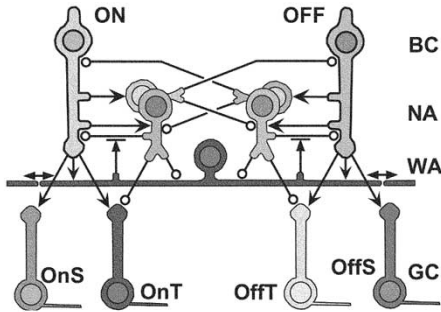


Fig. 4. Inner retina synaptic interactions. ON and OFF BCs relay cone signals to GCs, and excite NAs and WAs. NAs inhibit BCs, WAs, and transient GCs; their inhibition onto WAs is shunting. WAs modulate NA presynaptic inhibition and spread their signals laterally through gap junctions. BCs also excite local interneurons that inhibit complementary BCs and NAs.

dynamics to input frequency and to contrast, and drives GC responses, is shown in Fig. 4. Briefly, BC inputs to the inner retina excite GCs, an electrically coupled network of wide-field amacrine cells (WAs), and narrow-field amacrine cells (NAs) that provide feedback inhibition on to the bipolar terminals (BTs) [16]. WA, which receives excitation from both ON and OFF BT and inhibition from both ON and OFF NA, modulates presynaptic feedback inhibition from NA to BT. We use a large membrane capacitance to model the NA's slow, sustained, response, which leads to a less sustained response at the BT through presynaptic inhibition [18]. Push-pull inhibition is realized by a third set of amacrine cells in our model and is implemented by the ON-OFF BC circuit. These BT signals excite both sustained and transient GCs, but transient cells receive feedforward inhibition from NAs as well [25].

Our model of the inner retina is described by the system diagram in Fig. 5(a). We derived this diagram from the synaptic interactions of Fig. 4 by modeling the NA as a low-pass filter. Responses of the different inner retina cell types in this model to a step input are shown in Fig. 5(b). BC activity is a low-pass filtered version of light input to the outer retina. Increase in BC causes an increase in BT and a much slower increase in NA. The difference between BT activity and gain-modulated NA activity determines WA activity, which in turn sets the gain of NA feedback inhibition on to BT and WA. Thus, after a unit step input, BT activity initially rises but NA inhibition, setting in later, attenuates this rise until BT activity is equaled by gain-modulated NA activity. WA represents our local measure of contrast and

receives full-wave rectified input from ON and OFF BT and NA and, thus, rises above its baseline value of unity at both onset and offset. BT drives the sustained GC, GCs, which responds for the duration of the step, while the difference between BT and NA activity drives the transient GC, GCt, which decays to zero.

From the block diagram in Fig. 5(a), we can derive the system level equations for NA and BT with the help of the Laplace transform

$$i_{na} = \frac{g\epsilon}{\tau_A s + 1} i_{bc}, \quad i_{bt} = \frac{\tau_A s + \epsilon}{\tau_A s + 1} i_{bc} \quad (5)$$

where  $g$  is the gain of the excitation from BT to NA and where

$$\tau_A \equiv \epsilon \tau_{na}, \quad \epsilon \equiv \frac{1}{1 + wg} \quad (6)$$

where  $\tau_{na}$  is the time constant of NA and  $w$  represents WA activity, which determines feedback strength. From (5) and (6), we can see that BT high-pass filters and NA low-pass filters the BC signal; they have the same corner frequency,  $1/\tau_A$ . This closed-loop time-constant,  $\tau_A$ , depends on the loop gain  $wg$  and, therefore, on WA activity. For example, stimulating the inner retina with a high frequency would provide more BT excitation (high-pass response) than NA inhibition (low-pass response) to WA. WA activity and, hence,  $w$ , would subsequently rise, reducing the closed-loop time-constant  $\tau_A$ , until the corner frequency  $1/\tau_A$  reaches a point where BT excitation equaled NA inhibition. This drop in  $\tau_A$ , accompanied by a similar drop in  $\epsilon$ , will also reduce overall sensitivity and advance the phase of the response.

The system behavior governed by (5) and (6) is remarkably similar to the contrast gain control model proposed by Victor [28], which accounts for response compression in amplitude and in time with increasing contrast. Victor proposed a model for the inner retina whose high-pass filter's time-constant,  $T_S$ , is determined by a "neural measure of contrast,"  $c$ . The governing equation is

$$T_S = \frac{T_0}{1 + \frac{c}{c_{1/2}}}.$$

This model's time-constant depends on contrast in much the same way that our model's time constant depends on WA activity (6), where Victor's  $T_0$  is similar to our  $\tau_{na}$  and where Victor's ratio  $c/(c_{1/2})$  is represented by how much WA activity

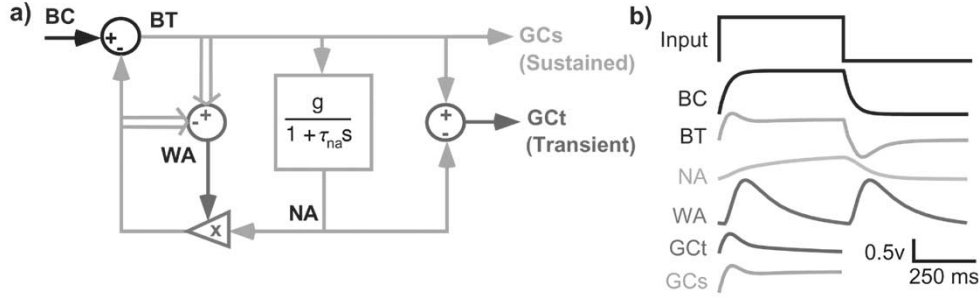


Fig. 5. Inner retina model. (a) System diagram: NA signals represent a low-pass filtered version of BT signals and provide negative feedback on to the BC. The WA network modulates the gain of NA feedback (X). WA receives full-wave rectified excitation from BT and full-wave rectified inhibition from modulated NA (double arrows). BT drives sustained GCs directly while the difference between BT and NA drives transient ganglion cells (GCt's). (b) Numerical solution to inner retina model with a unit step input of 1 V. Traces show 1 s of ON cell responses for BC, BT, NA, GCt, and WA. Outer retina time constant,  $\tau_o$ , is 96 ms;  $\tau_{na}$  is 1 s.

increases above its value at dc in our model. As this activity is sensitive to temporal contrast (see below), we propose that our WA cells are the anatomical substrate that computes Victor's neural measure of contrast.

In our model, the loop gain  $wg$  is set by the local *temporal contrast*. Because WA cells are coupled together through gap junctions, WA activity reflects inputs from BT and NA weighted across spatial locations. These pooled excitatory and inhibitory inputs should balance when the system is properly adapted

$$w \uparrow i_{na} \downarrow = \uparrow i_{bt} \downarrow + i_{surr} \quad (7)$$

$$\Leftrightarrow w = \frac{\uparrow i_{bt} \downarrow}{\downarrow i_{na} \downarrow} + \frac{i_{surr}}{\downarrow i_{na} \downarrow} \quad (8)$$

where we define  $i_{surr}$  as the gap-junction current resulting from spatial differences in WA activity  $w$ .  $\uparrow i_{bt} \downarrow$  and  $\downarrow i_{na} \downarrow$  are full-wave rectified versions of  $i_{bt}$  and  $i_{na}$ , computed by summing ON and OFF signals. If all different phases are pooled spatially, these full-wave rectified signals will not fluctuate and will be proportional to amplitude. And if we ignore  $i_{surr}$ ,  $w$  will simply be  $|i_{bt}|/|i_{na}|$ , yielding a temporal measure of contrast, since it is the ratio of a temporal difference (high-pass signal,  $i_{bt}$ ) and a temporal average (low-pass signal,  $i_{na}$ ). From (5), we see that at dc, this ratio is equal to  $1/g$ ; hence, the loop gain is unity and the dc gain  $\epsilon = 1/2$ .

GC responses in our inner retina model are derived from the BT and NA signals of (5). Specifically, BT signals directly excite both sustained and transient types of GCs, but transient cells receive feedforward NA inhibition as well. The system equations determining GCs's and GCt's responses, derived from (5), as a function of the input to the inner retina,  $i_{bc}$ , are

$$i_{GCs} = \frac{j\tau_A\omega + \epsilon}{j\tau_A\omega + 1} i_{bc} \quad (9)$$

$$i_{GCt} = \frac{j\tau_A\omega + \epsilon(1-g)}{j\tau_A\omega + 1} i_{bc}. \quad (10)$$

When BT-to-NA excitation has unity gain ( $g = 1$ ), feedforward inhibition causes a purely high-pass (transient) response in GCt while GCs retains a low-pass (sustained) component. With a small loop-gain,  $wg$ , the dc gain,  $\epsilon$ , approaches  $1/2$  and BT's and GCs's responses become all-pass. However, as the loop gain increases,  $\epsilon$  decreases and these responses become high-pass. The change in  $\epsilon$  with loop gain is matched in both BT and NA,

and so taking the difference between these two signals always cancels the sustained component in GCt. Thus, GCt produces a purely high-pass response irrespective of the system's loop gain.

#### A. Adapting to Temporal Frequency

Most physiological studies have characterized the retina's response when stimulated with only one temporal frequency. We can adopt a similar approach to characterize our model's GC responses, but to do so, we must determine how our system adapts to a single temporal frequency by deriving a mathematical expression for  $w$ 's dependence on contrast,  $c$ , as well as input frequency,  $\omega$ . As we stimulate our model with the same temporal frequency and contrast at all spatial locations, such that there is no difference between surround and center loop gain, we can set  $i_{surr} = 0$  in (8), provided the spatial frequency is not too high. Then we introduce the frequency dependencies of  $i_{bt}$  and  $i_{na}$  in order to determine the behavior of the system's closed-loop gain,  $wg$ .

We assume sinusoidal inputs,  $c \sin(\omega t)$ , with amplitude  $c$ . We also assume that the frequencies we are interested in are less than the outer retina's temporal cutoff, so that we can ignore outer retina filtering for the moment, and that  $i_{bt}$  and  $i_{na}$  only have energy at these input frequencies, plus a dc component. Hence, we obtain the spectra

$$i_{bt} = b_0\delta(\omega) + c \frac{j\tau_A\omega_i + \epsilon}{j\tau_A\omega_i + 1} \delta(\omega - \omega_i) \quad (11)$$

$$i_{na} = n_0\delta(\omega) + c \frac{g\epsilon}{j\tau_A\omega_i + 1} \delta(\omega - \omega_i) \quad (12)$$

where  $\epsilon$ ,  $\tau_A$ , and  $g$  are defined as above and  $\omega_i$  is the input frequency.  $b_0$  and  $n_0$  are the residual activity (dc and offsets) in  $i_{bt}$  and  $i_{na}$ , respectively. From our bipolar circuitry, we find that  $b_0$  is determined by  $V_{bq}$ . The source of NA residual activity,  $n_0$ , is explained below.

These currents only have energy at  $\omega = 0$  and  $\omega = \omega_i$ , and so, using Parseval's relationship, the ratio of the magnitudes of these two currents is simply the square-root of the ratio of the energy contained in  $i_{bt}$ 's two impulses over that in  $i_{na}$ 's. Setting  $n_0 = gb_0$  to simplify the computation, we find

$$wg = \sqrt{1 + \frac{\tau_{na}^2\omega^2}{1 + (\frac{b_0}{c})^2 \frac{1}{\epsilon^2} + (\frac{b_0}{c})^2 \tau_{na}^2\omega^2}} \quad (13)$$

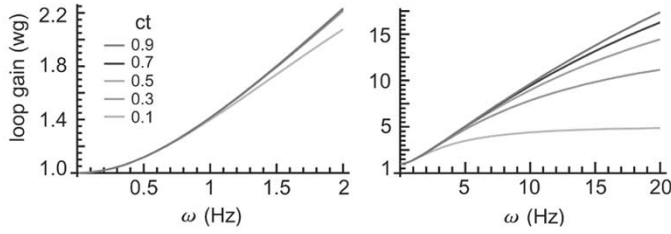


Fig. 6. Change in loop gain with frequency. The system loop gain,  $wg$ , depends both on temporal frequency and on contrast. Plots of this relationship are shown on both a small (left) and large (right) frequency scale. Loop gain rises with temporal frequency,  $\omega$ , and saturates at a point determined by the contrast level. For a given temporal frequency, higher contrasts generate a larger loop gain.

gives the system loop-gain as a function of  $c$ ,  $b_0$ , and  $\omega$ , when  $\epsilon\tau_{na}$  is substituted for  $\tau_A$ . Recall that  $\epsilon \equiv 1/(1 + wg)$ , which means that the loop gain  $wg$  determines  $\epsilon$ . Numerical solutions to this relationship are plotted in Fig. 6 for five different contrast levels, with a  $\tau_{na}$  of 1 s.

Finding approximations to (13) over different temporal frequency ranges, we obtain

$$wg \approx \begin{cases} 1, & \omega < \frac{1}{\tau_{na}} \\ \sqrt{1 + \tau_{na}^2 \omega^2}, & \frac{1}{\tau_{na}} < \omega < \frac{c}{b_0 \tau_{na}} \\ \sqrt{1 + \left(\frac{c}{b_0}\right)^2}, & \omega > \frac{c}{b_0 \tau_{na}}. \end{cases}$$

In the dc case, when  $\omega = 0$ , the system's loop gain is 1, as expected. As  $\omega$  exceeds  $1/\tau_{na}$ ,  $wg$ , which sets the closed-loop time-constant,  $\tau_A$ , tracks  $\omega$  over the range  $1/\tau_{na} < \omega < (c/b_0)(1/\tau_{na})$ . In this region, the inner retina effectively adapts to temporal frequency by matching its corner frequency to the input. The adaptation, however, only takes place over these intermediate frequencies—below  $1/\tau_{na}$ , the system's corner frequency is floored at  $2/\tau_{na}$ , and above  $(c/b_0)(1/\tau_{na})$ , it saturates at  $(c/b_0)(1/\tau_{na})$ .

We can use these loop gain expressions to determine the frequency response of our GC outputs. For temporal frequencies below  $c/b_0\tau_{na}$ , we can make the approximation  $wg \approx 1 + \tau_{na}\omega$ . Ignoring the dc component, and invoking the relationship  $\tau_A \equiv \tau_{na}/(1 + wg)$  and  $\epsilon \equiv 1/(1 + wg)$ , the GC frequency responses simplify to

$$i_{GCs} = \frac{j\tau_{na}\omega + 1}{(j + 1)\tau_{na}\omega + 2} i_{bc} \quad (14)$$

$$i_{GCt} = \frac{j\tau_{na}\omega + (1 - g)}{(j + 1)\tau_{na}\omega + 2} i_{bc}. \quad (15)$$

We can see from these equations that for input  $i_{bc} = c\sin(\omega t)$ , GCs' response rises from  $c/2$  at dc to  $c/\sqrt{2}$  at  $\omega \gg 1/\tau_{na}$ . Actually, it rises to  $c$  at frequencies above  $c/b_0\tau_{na}$ . Hence, we expect that GCs' response will be hardly affected by changes in  $\tau_{na}$  since GCs' response is practically flat across all  $\omega$ . Finally, we expect GC's response to remain independent of  $g$ , since it does not appear in the equations.

On the other hand, the GCt response rises from  $c(1 - g)/2$  at dc to  $c/\sqrt{2}$  at  $\omega \gg 1/\tau_{na}$ , following GCs' behavior. Because  $g \approx 1$ , GCt's response is close to zero at dc. Reducing  $\tau_{na}$  will shift the onset of its saturation range to higher frequencies and,

thus, reduce the slope of its rise. Furthermore, although  $g$  does not affect the temporal dynamics of the GCt's response, we can see that mismatch in  $g$  will introduce low-frequency responses in GCt. In summary, this analysis makes specific predictions about how the open-loop time-constant  $\tau_{na}$ , temporal frequency  $\omega$ , and contrast  $c$  interact to determine the frequency response, when we stimulate with a single temporal frequency.

### B. Adapting to Contrast

In addition to tracking temporal frequency, the inner retina also adapts to input contrast. We can see from Fig. 6 that loop-gain increases as we increase stimulus contrast and, hence, the system will adjust its corner frequency,  $\omega_A$ , so as to increase it, causing a speed up in the GC response and rejecting a larger range of low frequencies. This contrast-dependence arises because the point where loop gain,  $wg$ , saturates with temporal frequency occurs at higher temporal frequencies for higher contrasts. This happens because NA's response drops below its dc offset  $gb_0$  at frequencies above  $c/b_0\tau_{na}$ , making it impossible to track frequencies beyond this point.

However, to explore how our model responds to contrast in natural scenes, which stimulate the retina with several temporal frequencies simultaneously, we need to understand how much each spectral component contributes to setting the loop gain,  $wg$ . For any individual frequency, WA activity will reflect what the adapted activity for that individual frequency ought to be only to a limited extent since the WA activity will reflect the weighted contributions of *all* the frequencies to loop gain. Thus, although sensitivity to all frequencies drops when stimulus contrast increases, low-frequency gains are attenuated more [24] since the system's loop gain is pushed to a higher level by the higher frequencies than would normally be the case if low input frequencies solely determined system loop gain.

We use the same approach as above to formalize this intuition, but we now integrate the energy over the entire spectrum of temporal frequencies present to determine how  $i_{bt}$  and  $i_{na}$  activity affect loop gain,  $wg$ . We assume a flat spectrum, with amplitude proportional to contrast, that is assumed to be low-pass filtered by the outer retina. In this case, we define the contrast spectral density,  $d$ , as contrast, or amplitude, per unit frequency. The relation between  $d$  and the input signal contrast,  $c$ , is given as  $d^2 \equiv c^2/\omega_o$ , since we divide the input spectrum's energy, or contrast squared, by the bandwidth we are interested in. In this case,  $i_{na}$  and  $i_{bt}$ 's spectra are given by

$$i_{bt} = \left( b_0\delta(\omega) + d \frac{j\tau_A\omega + \epsilon}{j\tau_A\omega + 1} \right) \left( \frac{1}{j\tau_o\omega + 1} \right)^2 \quad (16)$$

$$i_{na} = \left( n_0\delta(\omega) + d \frac{g\epsilon}{j\tau_A\omega + 1} \right) \left( \frac{1}{j\tau_o\omega + 1} \right)^2 \quad (17)$$

where  $\epsilon$ ,  $\tau_A$ ,  $g$ ,  $b_0$ , and  $n_0$  are defined as above. To facilitate our analysis, we approximate the outer retina as two low-pass filters, with an identical fixed time constant of  $\tau_o$ , that sharply attenuate frequencies greater than  $\omega_o = 1/\tau_o$ . Sketches of  $|i_{bt}|$  and  $|i_{na}|$ 's spectra are shown in Fig. 7(a).

To find the loop gain, we take the ratio of the magnitudes of  $i_{bt}$  and  $i_{na}$ , computed using Parseval's relationship. Simplifying



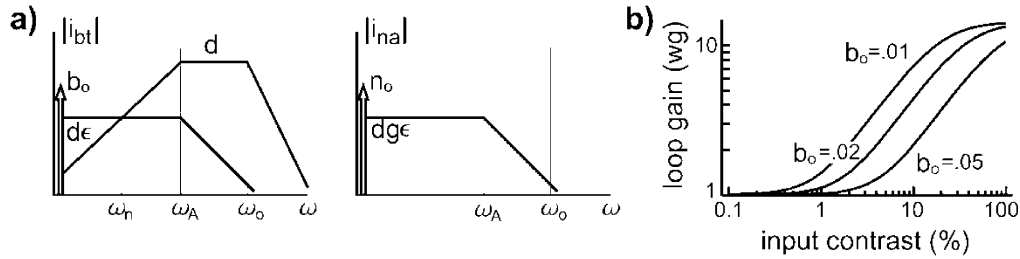


Fig. 7. Effect of contrast on loop gain. (a) BT activity,  $i_{bt}$ , is the sum of three components: a dc component that depends on residual BT activity,  $b_0$ , a low-pass component that equals  $d\epsilon$  and cuts off at  $\omega_A$ , and a high-pass component that rises as  $d\tau_A\omega$ , exceeds the low-pass component at  $\omega_n = 1/\tau_{na}$ , and saturates at  $\omega_A$ . The outer retina provides an absolute cutoff at  $\omega_o$ . NA activity,  $i_{na}$ , is the sum of a dc component that depends on NA residual activity,  $n_0$ , and a low-pass component that equals  $d g \epsilon$  and that cuts off at  $\omega_A$ . Loop gain,  $wg$ , is determined by the ratio between the energy in  $i_{bt}$  and the energy in  $i_{na}$ . (b) A numerical solution for loop gain as a function of contrast ( $c = d\sqrt{\omega_o}$ ) for three different levels of residual activity,  $b_0$ . As  $b_0$  increases, the curves shift to the right, implying that the contrast signal is not as strong.  $\tau_{na}$  is 1.038 s and  $\tau_o$  is 77 ms for this simulation.

our analysis by setting  $n_0 = g b_0$  and by treating the temporal cutoffs at  $\omega_A$  and  $\omega_o$  as infinitely steep, we find

$$wg = \left( \frac{b_0^2 + \int_0^{\omega_A} (d^2\epsilon^2 + d^2\tau_A^2\omega^2)d\omega + \int_{\omega_A}^{\omega_o} d^2d\omega}{b_0^2 + \int_0^{\omega_A} (d\epsilon)^2d\omega} \right)^{1/2} \quad (18)$$

$$= \sqrt{1 + \frac{c^2 - \frac{2}{3} \frac{\tau_o}{\epsilon\tau_{na}} c^2}{b_0^2 + \frac{\epsilon\tau_{na}}{\tau_{na}} c^2}}. \quad (19)$$

Recalling from (6) that  $\epsilon$  is a function of loop gain  $wg$ , we can find a numerical solution to how  $wg$  depends on input contrast. This relationship is shown in Fig. 7(b) for three different values of  $b_0$ . Loop gain  $wg$  approaches 1 as contrast per unit frequency,  $d$ , approaches 0.  $wg$  rises sublinearly with contrast over a range and saturates at a value determined by the ratio  $\tau_{na}/\tau_o$ . The contrast that realizes this saturation is determined by the amount of residual BC activity  $b_0$ . As  $b_0$  increases, the linear regime shifts to higher contrasts. This implies that the  $b_0$  determines the system's contrast threshold—the higher  $b_0$ , the more contrast we need to produce a given loop gain. This property, whereby the amount of residual activity determines the strength of the input contrast signal, is analogous to Victor's  $c_{1/2}$  term from (7).

The gross behavior of this analytical solution is captured by the approximation

$$wg \approx \begin{cases} \sqrt{1 + \left(\frac{c}{b_0}\right)^2} & c < c_S \\ \frac{3}{5} \frac{\tau_{na}}{\tau_o} & c > c_S. \end{cases} \quad (20)$$

Loop gain rises sublinearly with contrast and saturates when contrast exceeds some threshold,  $c_S$ , that depends on  $b_0$ . We can derive approximately what this threshold,  $c_S$ , is by recognizing that  $wg$  saturates when  $\epsilon\tau_{na}c^2/\tau_{na} \gg b_0^2$  in the denominator of the second term of (19). Assuming that  $\epsilon \approx 1/wg$  in this case, the saturation value for  $wg$  given above as  $3\tau_{na}/5\tau_o$  yields a limiting contrast value of  $c_S \approx b_0\tau_{na}\sqrt{3/5}/\tau_o$ .

To better illustrate how varying levels of contrast affect the range of temporal frequencies over which adaptation takes place, we can recognize that, below  $c_S$ , loop gain  $wg$  is approximated by  $1 + c/b_0$  quite well. Hence, our expressions for GC outputs (9) and (10) can be rewritten as

$$i_{GCs} = \frac{j\tau_{na}\omega + 1}{j\tau_{na}\omega + \frac{c}{b_0} + 1} i_{bc} \quad (21)$$

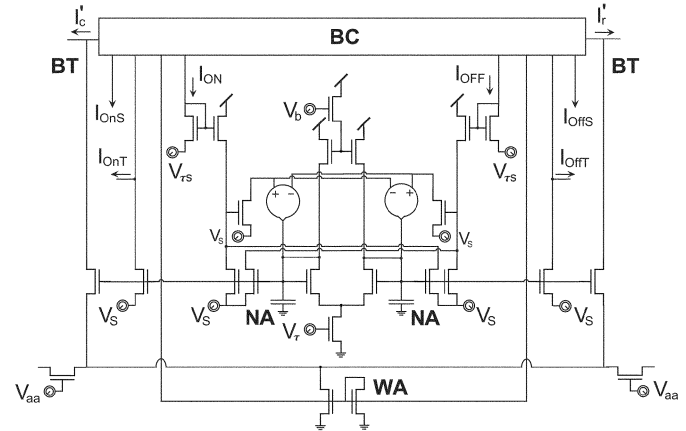


Fig. 8. Inner retina circuit. The inputs to the circuit,  $I'_c$  and  $I'_r$ , represent complementary inputs to the rectifying BC circuit. We make five copies at the output of this circuit (only three are shown in the figure). One copy drives the ON-OFF low-pass filter to excite the NA on one side of the circuit while inhibiting NA on the other side. A second copy is used to excite WAs, which modulates NA feedback inhibition on to BT. The remaining three copies feed ON and OFF versions of transient and sustained GCs (OnT, OnS, OffT, and OffS).

$$i_{GCt} = \frac{j\tau_{na}\omega}{j\tau_{na}\omega + \frac{c}{b_0} + 1} i_{bc}. \quad (22)$$

Frequencies below  $(c/b_0 + 1)/\tau_{na}$  are attenuated by  $c/b_0 + 1$  while frequencies above are not. Thus, the amount of attenuation, as well as the range of frequencies over which this attenuation takes place, gets larger with increasing contrast.

From our analysis, we see that both temporal frequency and contrast are interrelated in their effect on the dynamics governing our inner retina model. Furthermore, because of the relationship between temporal frequency and contrast, changes in one of these parameters cause changes in the model's adaptation to the other. Increasing either one of these parameters causes the time constants of our model both to speed up, which affects how our inner retina processes natural images. This adaptation to contrast and temporal frequency makes retinal coding more efficient, from an information theoretic perspective (see Discussion).

### C. Circuit Implementation

The complete inner retina circuit that realizes the synaptic interactions of Fig. 4 is shown in Fig. 8 (described in detail in [29]). Briefly, differential signals from BC circuitry,  $I_{ON}$  and

$I_{\text{OFF}}$ , drive either half of the circuit. Excitatory BT-to-NA current excite the NA node on one side of our circuit while also inhibiting the NA node on the complementary side, thus realizing push-pull inhibition. These inputs are divided by the sum of ON and OFF NA activity to compensate for the exponential voltage-current relationship in the subthreshold regime. NA provides feedback inhibition on to BT, modulated by WA activity. Convergence of ON and OFF signals on to WA implements full-wave rectification for both BT excitation and NA inhibition of WA. WA nodes are coupled to one another through an nMOS diffusion network gated by  $V_{aa}$ . A copy of NA current inhibits the *transient* GC. In fact, our circuit generates two copies of transient signals so that we can pool transient GC inputs over larger areas [30].

Because we have control over both WA coupling ( $V_{aa}$ ) and BT-to-NA gain ( $V_{\tau s}$ ), we can explore how changing these parameters changes GC responses. As WA modulation determines the dynamics of GC responses, we expect the extent of spatial coupling in the WA network, controlled by  $V_{aa}$ , to affect circuit dynamics. In our model, WA activity reflects inputs weighted across spatial locations, and is influenced by differences in center and surround WA activity. Relative contributions from different loop gains at different spatial locations are determined by the resistance of the WA network, with  $i_{\text{sur}} = \Delta w/R$  in (8). Here,  $\Delta w$  is the difference between WA activity in a surround location,  $w_s$ , and a center location,  $w_c$ , and  $R$  is the resistance coupling these two locations. Thus, depending on the strength of WA coupling, if WA activity in the surround is larger than that in the center, we expect the loop gain in the center to increase, whereas if the opposite is true, we expect loop gain in the center to decrease.

BT-to-NA gain is determined by the relationship between  $V_{\tau s}$ ,  $V_S$ , and  $V_\tau$  [29].  $V_S$  sets the source voltages for the NA subcircuit such that WA activity can be represented by voltage deviations below  $V_S$ . A bias voltage,  $V_\tau$ , determines leakage current,  $I_\tau$ , from NA's large membrane capacitor and, hence, the time constant for both common-mode and differential signals. And the gain of BT-to-NA excitation is determined by a bias voltage,  $V_{\tau s}$ , applied to the source of the current mirror's input transistor. Ideally,  $V_{\tau s}$  should be set equal to  $V_S + V_\tau$  for a gain of one [29]. If  $V_{\tau s} > V_S + V_\tau$ , then the gain,  $g$ , is greater than one, thus, WA activity should be lower and GCt responses should be inhibited. However, if  $V_{\tau s} < V_S + V_\tau$ , then  $g$  is less than one, causing the opposite effects [29]. Finally, another bias voltage,  $V_b$ , sets quiescent NA activity (i.e.,  $n_o$  in the model's equation) by exciting NA on both sides of the circuit with a current that is divided by ON and OFF NA activity (not shown). This ensures that ON and OFF NA activity are limited by a geometric mean constraint (just like  $V_{bq}$  does in the BT circuit).

#### IV. DISCUSSION

In this paper, we have presented our model and analyzed its behavior. In order to construct our model, and in order to facilitate its analysis, we made several simplifications. We address these assumptions and present our justifications here.

While our analysis of the outer retina demonstrates how our model filters signals in space, capacitances associated with the

cone and HC networks create time constants that produce temporal filtering that is inseparable from spatial filtering. Details of our outer retina's temporal properties and their effect on spatial filtering are discussed elsewhere [3]–[5]. Briefly, CT responses in our model change from bandpass to low-pass in space as temporal frequency increases. They also change from bandpass to low-pass in time as spatial frequency increases. This interplay between spatial and temporal filtering creates a CT response that peaks at a single temporal frequency for low spatial frequencies and at a single spatial frequency for low temporal frequencies. HCs modulate synaptic interactions to realize luminance adaptation in our outer retina model, but in the mammalian retina this role is most likely played by dopaminergic amacrine cells [12] and light sensing GCs [2].

Although temporal filtering in the outer retina is not addressed in this paper, we do use a simplified version of the outer retina's temporal filter for our analysis of contrast adaptation in the inner retina. This analysis, presented in Section III, models the outer retina as two low-pass filters with the same time constant,  $\tau_o$ . We felt this simplification was valid because our inner retina analysis is limited to high spatial frequencies where WA averaging across space occurs. At these high spatial frequencies, spatiotemporal filtering in the outer retina is indeed low-pass in time [3]–[5]. We examined responses in this simpler condition to determine how the inner retina adapts to temporal frequencies and contrast. The same adaptation process still takes place when a more complex spatiotemporally filtered signal is presented by the outer retina and, hence, the simple picture we have painted may provide useful intuition.

Our inner retina model is based upon the functional architecture and physiological behavior of the mammalian retina. All the synaptic interactions we included have been identified anatomically, as well as the complementary ON and OFF channels. Complementary signaling starts at cone-to-bipolar synapses, by using a sign-reversing synapse in one channel and half-wave rectification in both [8]. Push-pull interactions maintain complementary signaling in the inner retina through reciprocal inhibition between ON and OFF channels, via vertically diffuse amacrine cells that span the ON and OFF laminae [23]. Serial inhibition [10] may also play a vital role in these interactions.

While the outer retina adapts its sensitivity to light intensity, the inner retina adapts its low-pass and high-pass temporal filters to contrast and frequency. Encoding signals found in natural scenes optimally requires the retina's bandpass filters to peak at the spatial and temporal frequency where input signal power drops to the noise floor [1], [27]. Therefore, as stimuli with different spectra are presented to the retina, optimal coding requires this filter's peak frequency to move accordingly. Thus, the retina must adapt to spatial and temporal frequency and to luminance and contrast to continue to convey information efficiently to higher cortical structures. As stimulus power increases—as in the case of increased contrast—optimal filtering demands that the peak of this bandpass filter moves to higher frequencies, with a corresponding drop in low-frequency sensitivity.

Our inner retina model's temporal filter realizes this adaptation to contrast—GC responses compress in time and amplitude when driven by steps of increasing contrast [28]—by adjusting

its time constant [24], [28]. Our model proposes that this adaptation is implemented through WA modulation of NA feedback (pre-synaptic inhibition). Thus, we offer an anatomical substrate through which earlier dynamic models can be realized. A likely biological candidate for this WA is the A19 amacrine cell [16], which lies at the border between ON and OFF laminae, has thick dendrites, a large axodendritic field, and is electrically coupled to other A19 amacrine cells through gap junctions. Our inner retina circuit design also corrects flaws in the design in [3], which failed to produce temporal adaptation, and adds GC level synaptic interactions and convergence.

Our retina model's output is encoded as spike trains in four types of GC outputs. We sought to model parasol (also called Y in cat) and midget (also called X in cat) GCs, which respond in a transient and in a sustained manner, respectively, at stimulus onset or offset. Both types of GCs receive synaptic input from BCs and amacrine cells, although parasol cells receive more amacrine synapses, and, presumably, more feedforward inhibition [11], [17]. Hence, we included feedforward inhibition in our transient cells, but not in sustained cells. Furthermore, parasol cells also sample the visual scene nine times more sparsely than midget cells, and have proportionately larger receptive fields [7]. Thus, we tiled our transient cells less densely, and included convergence of bipolar signals on to these cells (see [30]). Ninety percent of the total primate GC population is made up of ON and OFF midget and parasol cells [22] and so we concentrated our modeling efforts on these four cell types.

## V. CONCLUSION

By morphing the retina's functional architecture into a VLSI circuit, we were able to create a circuit that expresses some of the defining features of visual processing in the mammalian retina. In our outer retina model, interaction between an excitatory cone network, which has relatively small space and time constants, and an inhibitory HC network, which has larger space and time constants, creates a bandpass spatiotemporal response. This model adapts to input luminance through HC modulation of cone-coupling and cone excitation, and through aut feedback, to produce a contrast signal at the CT. From our analysis, we see that CT activity is determined by input contrast, and independent of mean luminance. Hence, our model's CT's S-shaped behavior conveys a contrast signal to subsequent retina circuitry that does not change with light intensity, similar to the behavior exhibited by the mammalian retina. BCs in our model rectify these signals into ON and OFF channels to replicate the retina's complementary signaling. Our design was based on push-pull signaling in the mammalian retina that divides signals into complementary pathways to ensure efficient coding.

Furthermore, by morphing the inner retina's functional architecture into a VLSI circuit, we were able to extend previous models of visual processing in the mammalian retina. In our inner retina model, modulation of NA presynaptic inhibition by WAs realizes contrast gain control and time-constant adaptation. WAs serve as an anatomical substrate for computing a neural measure of contrast in our model. We use this measure of contrast in the same way that the retina is hypothesized to—modulation of feedback inhibition changes the system's closed-loop

time constant. This adaptation allows our model to optimally encode signals by adjusting its temporal filter based on input frequency and contrast. Thus, in response to natural scenes, our model measures all input frequencies present and computes how the system should best filter these signals.

A number of parameters in our model and circuit determine the system's behavior. Manual control over HC coupling in our outer retina model allows us to explore how changing space constants in this network affects GC output. By adjusting the NA leakage current, we also have control over the system's open-loop time constant. Changing this time constant affects the circuits temporal response and changes the temporal frequency above which closed-loop gain saturates. And by adjusting the BT to NA gain, we have control of residual transient cell activity. Finally, by adjusting coupling strength in the WA network, we can explore the effect of spatial frequency on how loop gain is computed over space. We explore some of these parameters in a companion paper [30], where test results from a fabricated chip are presented.

By reproducing some of the synaptic interactions found in the mammalian retina, we were able to derive models for the outer and inner retina that spatiotemporally filter signals, adapt to luminance, and exhibit contrast gain control. Our procedure has been to construct a neural model based on known anatomy in order to elicit some of the known behavior found in the physiology. While our model assigns novel roles to circuit elements that can account for this behavior, we still need to validate these conjectures by comparing our model's behavior with that of the mammalian retina. This detailed comparison will be presented in a forthcoming paper.

## REFERENCES

- [1] J. Atick and N. Redlich, "What does the retina know about natural scenes," *Neural Computation*, vol. 4, no. 2, pp. 196–210, 1992.
- [2] D. M. Berson, F. A. Dunn, and M. Takao, "Phototransduction by retinal ganglion cells that set the circadian clock," *Science*, vol. 295, no. 5557, pp. 1070–3, 2002.
- [3] K. Boahen, "A retinomorph chip with parallel pathways: Encoding increasing, on, decreasing, and off visual signals," *J. Analog Integr. Circuits Signal Processing*, vol. 30, pp. 121–135, 2002.
- [4] K. A. Boahen, "The retinomorph approach: Pixel-parallel adaptive amplification, filtering, and quantization," in *Analog Integr. Circuits Signal Processing*, vol. 13, 1997, pp. 53–68.
- [5] K. A. Boahen and A. Andreou, "A contrast-sensitive retina with reciprocal synapses," in *Advances in Neural Information Processing 4*, J. E. Moody, Ed. San Mateo, CA: MorganKaufmann, 1991, vol. 4, pp. 764–772.
- [6] W. Craelius, "The bionic man: Restoring mobility," *Science*, vol. 295, no. 5557, pp. 1018–21, 2002.
- [7] L. J. Croner and E. Kaplan, "Receptive fields of p and m ganglion cells across the primate retina," *Vis. Res.*, vol. 35, no. 1, pp. 7–24, 1995.
- [8] J. B. Demb, K. A. Zaghloul, L. Haarsma, and P. Sterling, "Bipolar cells contribute to nonlinear spatial summation in the brisk-transient (*y*) ganglion cell in mammalian retina," *J. Neurosci.*, vol. 21, pp. 7447–7454, 2001.
- [9] R. Douglas, M. Mahowald, and C. Mead, "Neuromorphic analogue visi," *Annu Rev Neurosci*, vol. 18, pp. 255–281, 1995.
- [10] J. E. Dowling and B. B. Boycott, "Organization of the primate retina: Electron microscopy," in *Proc. Roy. Soc. Lond. B*, vol. 166, 1966, pp. 80–111.
- [11] M. A. Freed and P. Sterling, "The on-alpha ganglion cell of the cat retina and its presynaptic cell types," *J. Neurosci.*, vol. 8, no. 7, pp. 2303–20, 1988.
- [12] R. J. Jensen and N. W. Daw, "Effects of dopamine and its agonists on the antagonists on the receptive field properties of ganglion cells in the rabbit retina," *Neuroscience*, vol. 17, no. 3, pp. 837–855, 1986.

- [13] M. Kamermans, I. Fahrenfort, K. Schultz, U. Janssen-Bienhold, T. Sjoerdsma, and R. Weiler, "Hemichannel-mediated inhibition in the outer retina," *Science*, vol. 292, no. 5519, pp. 1178–80, 2001.
- [14] M. Kamermans and F. Werblin, "Gaba-mediated positive autofeedback loop controls horizontal cell kinetics in tiger salamander retina," *J. Neurosci.*, vol. 12, no. 7, pp. 2451–63, 1992.
- [15] H. Kolb, N. Cuenca, H. H. Wang, and L. Dekorver, "The synaptic organization of the dopaminergic amacrine cell in the cat retina," *J. Neurocytol.*, vol. 19, pp. 343–366, 1990.
- [16] H. Kolb and R. Nelson, "Functional neurocircuitry of amacrine cells in the cat retina," in *Neurocircuitry of the Retina: A Cajal Memorial*, A. Gallego and P. Gouras, Eds. New York: Elsevier, 1985.
- [17] —, "Off-alpha and off-beta ganglion cells in cat retina: II. neural circuitry as revealed by electron microscopy of hrp stains," *J. Comp. Neurol.*, vol. 329, no. 1, pp. 85–110, 1993.
- [18] G. Maguire and P. Lukasiewicz, "Amacrine cell interactions underlying the response to change in tiger salamander retina," *J. Neurosci.*, vol. 9, pp. 726–35, 1989.
- [19] C. Mead, *Analog VLSI and Neural Systems*. Reading, MA: Addison-Wesley, 1989.
- [20] R. A. Normann and I. Perlman, "The effect of background illumination on the photoresponses of rod and green cones," *J. Physiol.*, vol. 286, pp. 491–507, 1979.
- [21] J. P. Rauschecker and R. V. Shannon, "Sending sound to the brain," *Science*, vol. 295, no. 5557, pp. 1025–29, 2002.
- [22] R. W. Rodieck, "The primate retina," *Comp. Primate Biol.*, vol. 4, pp. 203–278, 1988.
- [23] B. Roska and F. Werblin, "Vertical interactions across ten parallel, stacked representations in the mammalian retina," *Nature*, vol. 410, pp. 583–587, 2001.
- [24] R. Shapley and J. D. Victor, "The contrast gain control of the cat retina," *Vis. Res.*, vol. 19, pp. 431–434, 1979.
- [25] P. Sterling, *Retina*, and G. M. Shepherd, Eds., *The Synaptic Organization of the Brain*, fourth ed. New York: Oxford Univ. Press, 1998.
- [26] J. B. Troy and C. Enroth-Cugell, "X and y ganglion cells inform the cat's brain about contrast in the retinal image," *Exp. Brain Res.*, vol. 93, pp. 383–390, 1993.
- [27] J. H. van Hateren, "A theory of maximizing sensory information," *Biological Cybern.*, vol. 68, pp. 23–29, 1992.
- [28] J. D. Victor, "The dynamics of cat retinal x cell centre," *J. Physiol.*, vol. 386, pp. 219–246, 1987.
- [29] K. A. Zaghoul and K. Boahen, "An on-off log-domain filter circuit," *IEEE Trans. Circuits Syst.*, 2004, to be published.
- [30] —, "Optic nerve signals in a neuromorphic chip II: Testing and results," *IEEE Trans Biomed. Eng.*, vol. 51, pp. 667–675, Apr. 2004.

[31] —, "A silicon model of the mammalian retina," *Neuron*, 2004, submitted for publication.

[32] E. Zrenner, "Will retinal implants restore vision," *Science*, vol. 295, no. 5557, pp. 1022–25, 2002.



**Kareem A. Zaghoul** (S'00–M'03) received the B.S. degree from the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, in the in 1995, where he made Tau Beta Kappa and Eta Kappa Nu. He recently completed a combined M.D./Ph.D. program at the University of Pennsylvania, Philadelphia. The Ph.D. degree was awarded in the Department of Neuroscience, where he worked on understanding information processing in the mammalian retina with K. A. Boahen. His work was supported by a Vision Training Grant from the National Institutes of Health and a Ben Franklin Fellowship from the University of Pennsylvania School of Medicine. He is currently a Resident Physician in the Department of Neurosurgery at the University of Pennsylvania.



**Kwabena A. Boahen** received the B.S. and M.S.E. degrees in electrical and computer engineering from the Johns Hopkins University, Baltimore MD, in the concurrent masters-bachelors program, both in 1989, where he made Tau Beta Kappa. He received the Ph.D. degree in computation and neural systems from the California Institute of Technology, Pasadena, in 1997, where he held a Sloan Fellowship for Theoretical Neurobiology.

He is an Associate Professor in the Bioengineering Department, University of Pennsylvania, Philadelphia, where he holds a secondary appointment in the electrical engineering. He was awarded a Packard Fellowship in 1999. His current research interests include mixed-mode multichip VLSI models of biological sensory and perceptual systems, and their epigenetic development, and asynchronous digital interfaces for interchip connectivity.

Dr. Boahen received a National Science Foundation (NSF) CAREER Award in 2001 and an Office of Naval Research (ONR) YIP Award in 2002.