A Contrast Sensitive Silicon Retina with Reciprocal Synapses

Kwabena A. Boahen
Computation and Neural Systems
California Institute of Technology
Pasadena, CA 91125

Andreas G. Andreou
Electrical and Computer Engineering
Johns Hopkins University
Baltimore, MD 21218

Abstract

The goal of perception is to extract invariant properties of the underlying world. By computing contrast at edges, the retina reduces incident light intensities spanning twelve decades to a twentyfold variation. In one stroke, it solves the dynamic range problem and extracts relative reflectivity, bringing us a step closer to the goal. We have built a contrast-sensitive silicon retina that models all major synaptic interactions in the outer–plexiform layer of the vertebrate retina using current–mode CMOS circuits: namely, reciprocal synapses between cones and horizontal cells, which produce the antagonistic center/surround receptive field, and cone and horizontal cell gap junctions, which determine its size. The chip has 90×92 pixels on a 6.8×6.9mm die in 2μm n–well technology and is fully functional.

1   INTRODUCTION

Retinal cones use both intracellular and extracellular mechanisms to adapt their gain to the input intensity level and hence remain sensitive over a large dynamic range. For example, photochemical processes within the cone modulate the photocurrents while shunting inhibitory feedback from the network adjusts its membrane conductance. Adaptation makes the light sensitivity inversely proportional to the recent input level and the membrane conductance proportional to the background intensity. As a result, the cone's membrane potential is proportional to the ratio between the input and its spatial or temporal average, i.e. contrast. We have

764

A Contrast Sensitive Silicon Retina with Reciprocal Synapses   765

developed a contrast–sensitive silicon retina using shunting inhibition.

This silicon retina is the first to include variable inter–receptor coupling, allowing one to trade–off resolution for enhanced signal–to–noise ratio, thereby revealing low–contrast stimuli in the presence of large transistor mismatch. In the vertebrate retina, gap junctions between photoreceptors perform this function [5]. At these specialized synapses, pores in the cell membranes are juxtaposed, allowing ions to diffuse directly from one cell to another [6]. Thus, each receptor's response is a weighted average over a local region. The signal–to–noise ratio increases for features larger than this region—in direct proportion to the space constant [5].

Our chip achieves a four–fold improvement in density over previous designs [2]. We use innovative current–mode circuits [7] that provide very high functionality while faithfully modeling the neurocircuitry. A bipolar phototransistor models the photocurrents supplied by the outer–segment of the cone. We use a novel single–transistor implementation of gap junctions that exploits the physics of MOS transistors. Chemical synapses are also modeled very efficiently with a single device.

Mahowald and Mead's pioneering silicon retina [2] coded the logarithm of contrast. However, a logarithmic encoding degrades the signal–to–noise ratio because large signals are compressed more than smaller ones. Mead et. al. have subsequently improved this design by including network–level adaptation [4] and adaptive photoreceptors [3, 4] but do not implement shunting inhibition. Our silicon retina was designed to encode contrast directly using shunting inhibition.

The remainder of this paper is organized as follows. The neurocircuitry of the distal retina is described in Section 2. Diffusors and the contrast–sensitive silicon retina circuit are featured in Section 3. We show that a linearized version of this circuit computes the regularized solution for edge detection. Responses from a one–dimensional retina showing receptive field organization and contrast sensitivity, and images from the two–dimensional chip showing spatial averaging and edge enhancement are presented in Section 4. Section 5 concludes the paper.
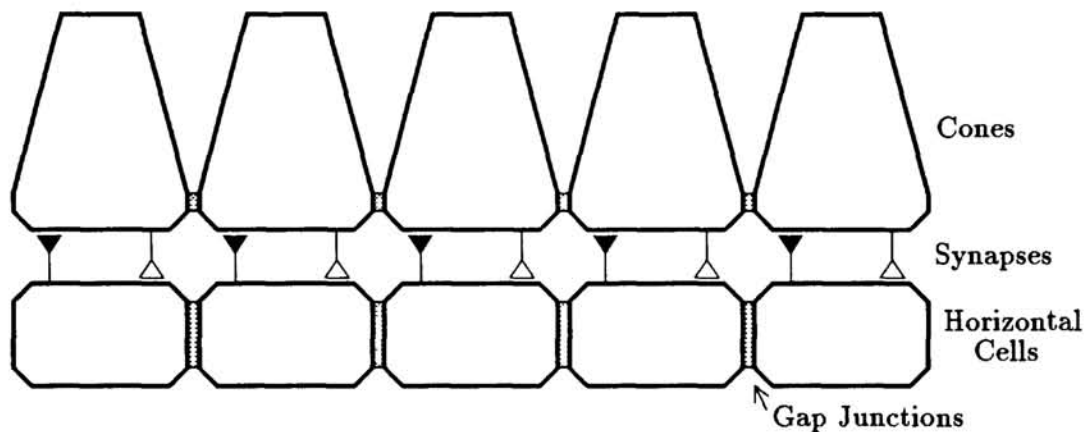
Figure 1: Neurocircuitry of the outer–plexiform layer. The white and black triangles are excitatory and inhibitory chemical synapses, respectively. The grey regions between adjacent cells are electrical gap junctions.

766   Boahen and Andreou

2   THE RETINA

The outer plexiform layer of the retina produces the well–known antagonistic center/surround receptive field organization first described in detail by Kuffler in the cat [11]. The functional neurocircuitry, based on the red cone system in the turtle [10, 8, 6], is shown in Figure 1. Cones and horizontal cells are coupled by gap junctions, forming two syncytia within which signals diffuse freely. The gap junctions between horizontal cells are larger in area (larger number of elementary pores), so signals diffuse relatively far in the horizontal cell syncytium. On the other hand, signals diffuse poorly in the cone syncytium and therefore remain relatively strong locally. When light falls on a cone, its activity increases and it excites adjacent horizontal cells which reciprocate with inhibition. Due to the way signals spread, the excitation received by nearby cones is stronger than the inhibition from horizontal cells, producing net excitation in the center. Beyond a certain distance, however, the reverse is true and so there is net inhibition in the surround.

The inhibition from horizontal cells is of the shunting kind and this gives rise to to contrast sensitivity. Horizontal cells depolarize the cones by closing chloride channels while light hyperpolarizes them by closing sodium channels [9, 1]. The cone's membrane potential is given by

$$V = \frac{g_{Na}E_{Na} + g_D V_{net}}{g_{Na} + g_{Cl} + g_D} \tag{1}$$

where the conductances are proportional to the number of channels that are open and voltages are referred to the reversal potential for chloride. g_D and V_{net} describe the effect of gap junctions to neighboring cones. Since the horizontal cells pool signals over a relatively large area, g_{Cl} will depend on the background intensity. Therefore, the membrane voltage will be proportional to the ratio between the input, which determines g_{Na}, and the background.
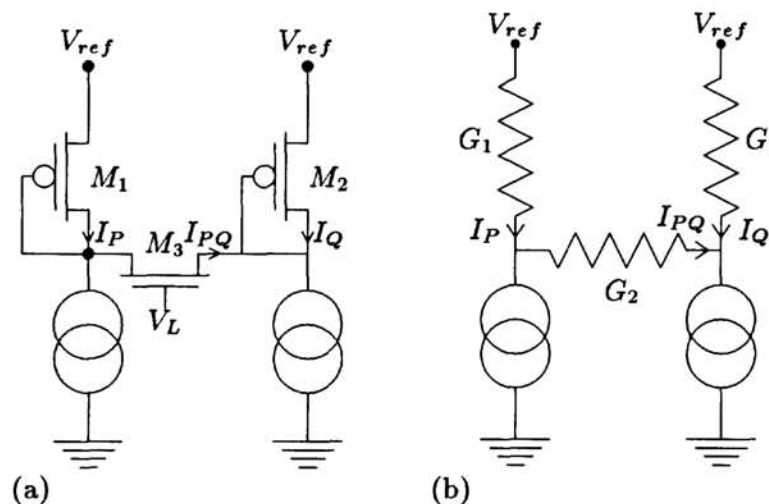
Figure 2: (a) Diffusor circuit. (b) Resistor circuit. The diffusor circuit simulates the currents in this linear resistive network.

A Contrast Sensitive Silicon Retina with Reciprocal Synapses   767

3   SILICON MODELS

In the subthreshold region of operation, a MOS transistor mimics the behavior of a gap junction. Current flows by *diffusion:* the current through the channel is linearly proportional to the difference in carrier concentrations across it [2]. Therefore, the channel is directly analogous to a porous membrane and carrier concentration is the analog of ionic species concentration. In conformity with the underlying physics, we call transistors in this novel mode of operation *diffusors*. The gate modulates the carrier concentrations at the drain and the source multiplicatively and therefore sets the *diffusivity*. In addition to offering a compact gap junction with electronically adjustable 'area,' the diffusor has a large dynamic range—at least five decades.

A current–mode diffusor circuit is shown in Figure 2a. The currents through the diode–connected well devices M_1 and M_2 are proportional to the carrier concentrations at either end of the diffusor M_3. Consequently, the diffusor current is proportional to the current difference between M_1 and M_2. Starting with the equation describing subthreshold conduction [2, p. 36], we obtain an expression for the current I_{PQ} in terms of the currents I_P and I_Q, the reference voltage V_{ref} and the bias voltage V_L :

$$I_{PQ} = e^{\kappa V_L - V_{ref}}(I_Q^{1/\kappa} - I_P^{1/\kappa}) \tag{2}$$

For simplicity, voltages and currents are in units of $V_T = kT/q$, and I_0, the zero bias current, respectively; all devices are assumed to have the same κ and I_0. The ineffectiveness of the gate in controlling the channel potential, measured by $\kappa \approx 0.75$, introduces a small nonideality. There is a direct analogy between this circuit and the resistive circuit shown in Figure 2b for which $I_{PQ} = (G_2/G_1)(I_Q - I_P)$. The currents in these circuits are identical if $G_2/G_1 = \exp(\kappa V_L - V_{ref})$ and $\kappa = 1$. Increasing V_L or reducing V_{ref} has the same effect as increasing G_2 or reducing G_1.

Chemical synapses are also modeled using a single MOS transistor. Synaptic inputs to the turtle cone have a much higher resistance, typically $0.6G\Omega$ or more [1], than the input conductance of a cone in the network which is $50M\Omega$ or less [8]. Thus the synaptic inputs are essentially current sources. This also holds true for horizontal cells which are even more tightly coupled. Accordingly, chemical synapses are modeled by a MOS transistor in *saturation*. In this regime, it behaves like a current source driving the postsynapse controlled by a voltage in the presynapse. The same applies to the light–sensitive input supplied by the cone outer–segment; its peak conductance is about $0.4G\Omega$ in the tiger salamander [9]. Therefore, the cone outer–segment is modeled by a bipolar phototransistor, also in saturation, which produces a current proportional to incident light intensity.

Shunting inhibition is not readily realized in silicon because the 'synapses' are current sources. However, to first order, we achieve the same effect by modulating the gap junction diffusitivity g_D (see Equation 1). In the silicon retina circuit, we set V_L globally for a given diffusitivity and control V_{ref} locally to implement shunting inhibition.

A one–dimensional version of the current–mode silicon retina circuit is shown in Figure 3. This is a direct mapping of the neurocircuitry of the outer–plexiform layer (shown in Figure 1) onto silicon using one transistor per chemical synapse/gap junction. Devices M_1 and M_2 model the reciprocal synapses. M_4 and M_5 model

768   Boahen and Andreou
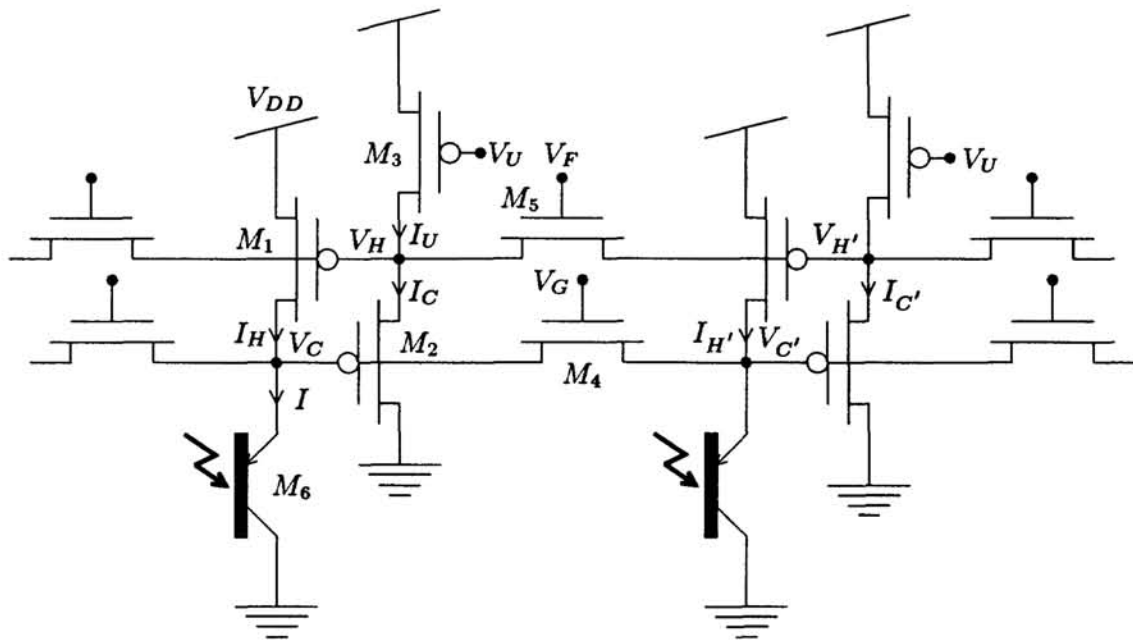
Figure 3: Current–mode Outer–Plexiform Circuit.

the gap junctions; their diffusitivities are set globally by the bias voltages V_G and V_F. The phototransistor M_6 models the light–sensitive input from the cone outer segment. The transistor M_3, with a fixed gate bias V_U, is analogous to a leak in the horizontal cell membrane that counterbalances synaptic input from the cone.

The circuit operation is as follows. The currents I_C and I_H represent the responses of the cone and the horizontal cell, respectively. These signals are actually in the post–synaptic circuit—the nodes with voltage V_C and V_H correspond to the presynaptic signals but they encode the logarithm of the response. Increasing the photocurrent will cause V_C to drop, turning on M_2 and increasing its current I_C; this is excitation. I_C pulls V_H down, turning on M_1 and increasing its current I_H; another excitatory effect. I_H, in turn, pulls V_C up, turning off M_2 and reducing its current I_C; this is inhibition.

The diffusors in this circuit behave just like those in Figure 2 although the well devices are not diode–connected. The relationship between the currents given by Equation 2 still holds because the voltages across the diffusor are determined by the currents through the well devices. However, the reference voltage for the diffusors between 'cones' (M_4) is not fixed but depends on the 'horizontal cell' response. Since $I_H = \exp(V_{DD} - \kappa V_H)$, the diffusivity in the cone network will be proportional to the horizontal cell response. This produces shunting inhibition.

3.1   RELATION TO LINEAR MODELS

Assuming the horizontal cell activities are locally very similar due to strong coupling, we can replace the cone network diffusitivity by $\hat{g} = \langle I_H \rangle g$, where $\langle I_H \rangle$ is the local average. Now we treat the diffusors between the 'cones' as if they had a fixed

A Contrast Sensitive Silicon Retina with Reciprocal Synapses   769

diffusitivity $\hat{g}$; the diffusivity in the 'horizontal cell' network is denoted by h. Then the equations describing the full two–dimensional circuit on a square grid are:

$$I_H(x_m, y_n) = I(x_m, y_n) + \hat{g} \sum_{\substack{i = m \pm 1 \\ j = n \pm 1}} \{I_C(x_i, y_j) - I_C(x_m, y_n)\} \tag{3}$$

$$I_C(x_m, y_n) = I_U + h \sum_{\substack{i = m \pm 1 \\ j = n \pm 1}} \{I_H(x_m, y_n) - I_H(x_i, y_j)\} \tag{4}$$

This system is a special case of the dual layer outer plexiform model proposed by Yagi [12]—we have the membrane admittances set to zero and the synaptic strengths set to unity. Using the second–difference approximation for the laplacian, we obtain the continuous versions of these equations

$$I_H(x, y) = I(x, y) + \hat{g}\nabla^2 I_C(x, y) \tag{5}$$

$$I_C(x, y) = I_U - h\nabla^2 I_H(x, y) \tag{6}$$

with the internode distance normalized to unity. Solving for $I_H(x, y)$, we find

$$\lambda\nabla^2\nabla^2 I_H(x, y) + I_H(x, y) = I(x, y) \tag{7}$$

This is the *biharmonic* equation used in computer vision to find an optimally smooth interpolating function '$I_H(x, y)$' for the noisy, discrete data '$I(x, y)$' [13]. The co-efficient $\lambda = \hat{g}h$ is called the regularizing parameter; it determines the trade–off between smoothing and fitting the data. In this context, the function of the horizontal cells is to compute a smoothed version of the image while the cones perform edge detection by taking the laplacian of the smoothed image as given by Equation 6. The space constant of the solutions is $\lambda^{1/4}$ [13]. This predicts that the receptive field size of our retina circuit will be weakly dependent on the input intensity since $\hat{g}$ is proportional to the horizontal cell activity.

4   CHIP PERFORMANCE

Data from the one–dimensional chip showing receptive field organization is in Figure 4. As the 'cone' coupling increases, the gain decreases and the excitatory and inhibitory subregions of the receptive field become larger. Increasing the 'horizontal cell' coupling also enlarges the receptive field but in this case the gain increases. This is because stronger diffusion results in weaker signals locally and so the inhibition decreases. Figure 5(a) shows the variation of receptive field size with intensity— roughly doubling in size for each decade. This indicates a one–third power dependence which is close to the theoretical prediction of one–fourth for the linear model. The discrepancy is due to the body effect on transistor M_2 (see Figure 3) which makes the diffusor strength increase with a power of $1/\kappa^2$.

Contrast sensitivity measurements are shown in Figure 5(b). The S–shaped curves are plots of the Michaelis–Menten equation used by physiologists to fit responses of cones [6]:

$$V = V_{max}\frac{I^n}{I^n + \sigma^n} \tag{8}$$

770   Boahen and Andreou
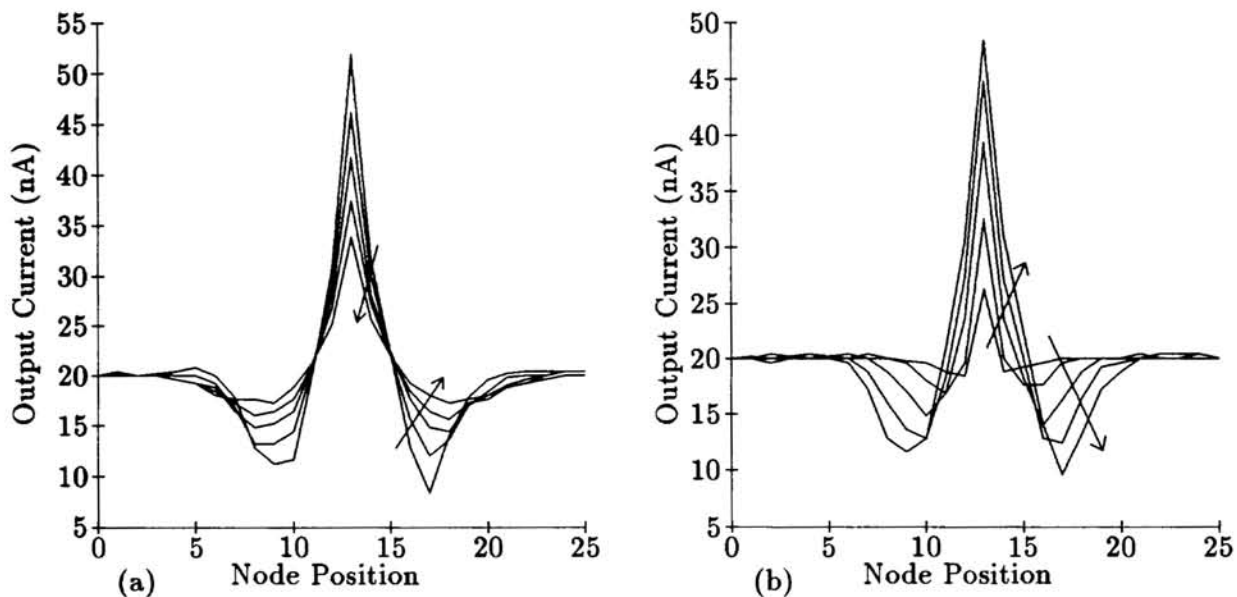
Figure 4: Receptive fields measured for 25×1 pixel chip; arrows indicate increasing diffusor gate voltages. The inputs were 50nA at the center and 10nA elsewhere, and the output current I_U was set to 20nA. (a) Increasing inter-receptor diffusor voltages in 15mV steps. (b) Increasing inter-horizontal cell diffusor voltages in 50mV steps.
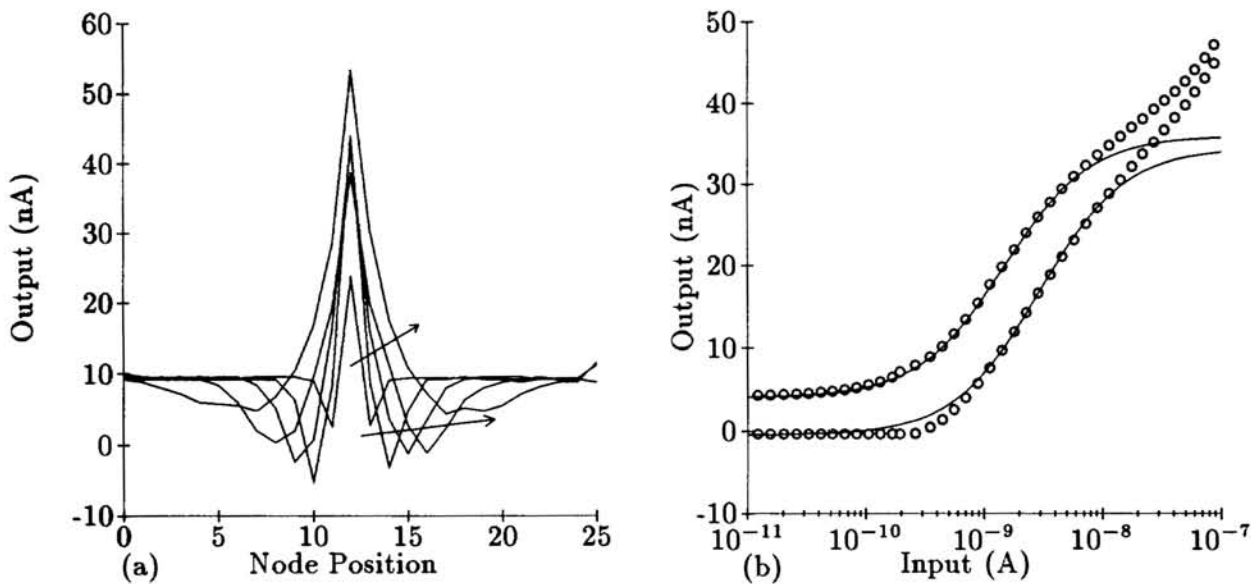
Figure 5: (a) Dependence of receptive field on intensity; arrows indicate increasing intensity. Center inputs were 500pA, 5nA, 15nA, 50nA, and 500nA. The background input was always one-fifth of the center input. (b) Contrast sensitivity measurements at two background intensity levels. Lines are fits of the Michaelis-Menten equation.