

PROCEEDINGS OF SPIE

[SPIDigitalLibrary.org/conference-proceedings-of-spie](https://spiedigitallibrary.org/conference-proceedings-of-spie)

Silicon retina with adaptive photoreceptors

Misha A. Mahowald

Misha A. Mahowald, "Silicon retina with adaptive photoreceptors," Proc. SPIE 1473, Visual Information Processing: From Neurons to Chips, (9 July 1991); doi: 10.1117/12.45540

SPIE.

Event: Orlando '91, 1991, Orlando, FL, United States

Silicon retina with adaptive photoreceptors

M.A. Mahowald

California Institute of Technology
Pasadena, California, 91125

Abstract

The central problem faced by the retina is to encode reliably small local differences in image intensity over a several decade range of background illumination. The distal layers of the retina adjust the transducing elements to make this encoding possible. Several generations of silicon retinae that integrate phototransducers and CMOS processing elements in the focal plane are modeled after the distal layers of the vertebrate retina.^{1,2,4} We describe a silicon retina with an adaptive photoreceptor that responds with high gain to small spatial and temporal variations in light intensity. Comparison with a spatial and temporal average of receptor response extends the dynamic range of the receptor. Continuous, slow adaptation centers the operating point of the photoreceptor around its time-average intensity and compensates for static transistor mismatch.

1 Introduction

The silicon retina is a two-dimensional array of phototransducers and processing elements integrated on the surface of a CMOS chip. The computations it performs are inspired by the computations performed in the distal layers of the vertebrate retina, which include the cones, the horizontal cells and the bipolar cells. The cones are arrayed two-dimensionally; they transduce the incoming light image into neural signals. Horizontal cells extend their processes parallel to the photoreceptor array and make contact with their neighbors via gap junctions. The horizontal cells spatially and temporally average photoreceptor signals. Hyperpolarizing and depolarizing bipolar cells partially rectify the receptor response and transmit a differential signal to the retinal ganglion cells, the output cells of the retina. Bipolar cell response is based on the difference between the average computed by horizontal cells and the response of individual photoreceptors.

The most crucial function of these first three layers is to adapt to prevailing light conditions. The photoreceptors, horizontal cells and bipolar cells take widely-varying amounts of incoming light and produce a signal with much narrower dynamic range that nonetheless captures the important information in an image. Adaptation is necessary if the system is to respond sensitively to small local changes in image intensity when the background intensities may vary by a factor of one million.

The silicon retina of Mead and Mahowald^{1,4} modeled the bipolar cell response of the mudpuppy⁷. The computation in the silicon retina was completely feed forward. It used a local average of intensity computed by a resistive network as a reference for a high gain "bipolar cell" response. By reporting only differential signals, the output of the silicon retina remained invariant over several orders of magnitude in background light intensity. However, this silicon retina did not employ adaptation within the photoreceptor itself. The receptor was very simple; a diode load to convert the photocurrent generated by a bipolar transistor into a voltage that was logarithmic in light intensity over four to five orders of magnitude. This arrangement gave the system a good dynamic range, but the gain was so low that realistic image contrasts produced signals of about the same magnitude as the subthreshold transistor mismatches. The output amplifier was just as happy to amplify transistor mismatches as it was to amplify real intensity variations in the image.

Biological retinae achieve high-gain responses from their transducers while avoiding saturation using a combination of mechanisms that operate over different time scales. Tiger salamander cones rapidly shift

their response range using the horizontal cell network as a reference point.⁶ The cones, therefore, maintain a high-gain response to small changes in illumination and do not saturate when presented with semi-global background intensity variation. In addition to this network-mediated shift in operating point, cones show intrinsic adaptation over longer time scales. Normann and Werblin⁵ showed that mudpuppy cones adapt to center their output around a long-time average intensity. The combination of rapid feedback from the horizontal cells and a slow adaptive process allows the cone to report small changes in illumination over six to seven orders of magnitude in background intensity.

A silicon analogue of these adaptive processes was first incorporated in a silicon retina by Mead³ using an ultraviolet-programmable floating gate. As is done in the tiger salamander retina, the operating point of a high-gain receptor was modulated by feedback from the resistive network.⁶ Although the transduction processes in cones and in silicon are unrelated, slow adaptation plays an important role in silicon circuits. Slow adaptation was incorporated in the Mead retina in order to keep transistor mismatches from being amplified by the feedback from the resistive network. After adaptation this retina responded quite well to low contrast images without offsets. However, the adaptation needed to be repeated if the background light level changed significantly. Because ultraviolet could not be used continuously to adapt the chip while it was in operation, this adaptive retina had practical limitations.

We report the development of a photoreceptor that adapts continuously while the retina is in operation. We show data from a single receptor that demonstrates the performance of the receptor in isolation. We compare the response of the adaptive retina to a low-contrast edge to that of the Mead and Mahowald retina.

2 The Adaptive Retina

The architecture of the adaptive retina is similar to those of previously described silicon retinae.^{1,2,4} Each pixel in the network is linked to its six neighbors with circuits that approximate resistive elements³, to form the hexagonal array shown in Figure 1. Each node of the array has a single bias circuit to control the strength of the six associated resistive connections. The photoreceptors, which are described in detail in the next section, act as voltage inputs that drive the resistive network through saturating conductances.

The resistive network, which models the horizontal cells of the vertebrate retina, computes a spatially and temporally weighted average of photoreceptor inputs. The spatial scale of the weighting function is determined by the product of the lateral resistance and the conductance coupling the photoreceptors into the network. Varying the conductance of the transconductance amplifier or the strength of the resistors changes the space constant of the network, and thus changes the effective area over which signals are averaged. At a particular space constant, the temporal response of the network is determined by the product of the fixed capacitance of each node, and the absolute value of the conductance coupling the photoreceptors into the network. The network represents a spatially and temporally low-pass filtered version of the receptor inputs.

Because the resistive network feeds back to the receptor, the final receptor output is an amplified version of the difference between the average computed by the resistive network, and the photocurrent of the receptor. The averaging distance of the network must be set at a relatively high value for this difference to be significant. The output amplifier produces a bidirectional current output based on the difference between the receptor output and the average computed by the resistive network.

2.1 The Adaptive Receptor

Our adaptive photoreceptor, which is analogous to the vertebrate cone, includes a transducing element embedded in several parallel feedback loops. A schematic of the receptor is shown in Figure 2. For the phototransducer, we used a bipolar transistor, which produces a current proportional to the number of photons it absorbs. The photocurrent is supplied from the power supply through the action of the transistor Q1. The gate of transistor Q1 is modulated by the action of three parallel feedback loops, which act a three different time scales. Because Q1 has an exponential current-voltage relation in subthreshold, the voltage response of the receptor is proportional to the logarithm of the light intensity. The current through Q1

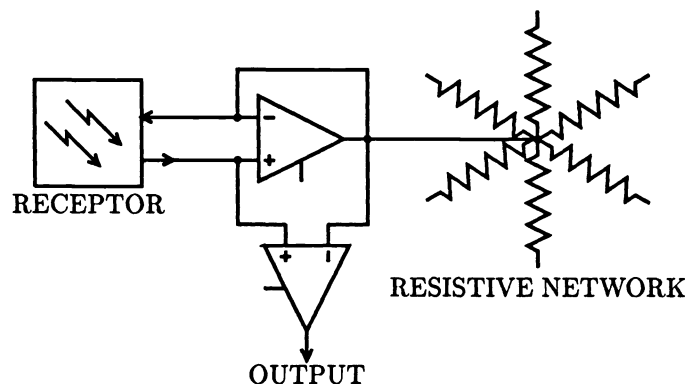


Figure 1: Schematic of a pixel. The output is the difference between the potential of the local receptor and that of the resistive network. The network computes a weighted average over neighboring pixels. The network feeds back to the receptor to suppress its response in uniform areas of the image.

clamps the emitter voltage, V_E , to be roughly equal to the absolute value of the gate-source voltage on the bias transistor. Small variations in V_E are amplified by the inverting stage comprising the bias transistor and Q2. The output of the receptor, V_{out} , is the voltage of the inverting amplifier.

The feedback loops that determine the receptor output operate on three different time scales. The fastest time-scale feedback signal is transmitted through the capacitor C_R , which couples V_{out} to the gate of Q2. The ratio of C_R to the sum $C_F + C_H + C_R$, sets the instantaneous gain of the receptor. Feedback from the resistive network to the receptor occurs at a slower time scale due to the temporal low-pass characteristics of the resistive network. If illumination is uniform so that the resistive network follows the receptor response, the gain of the receptor is reduced to be proportional to $C_R + C_H / C_R + C_H + C_F$. Feedback from the resistive network prevents saturation of the receptor when the background illumination level changes. A slow feedback loop through the diode connected transistors, Q3 and Q4, reduces the gain of the receptor for long times to $kT/q\kappa$ volts per e-fold increase in photocurrent from the transducer. The time scale of this adaptation is set by the reverse leakage current through Q3 or Q4. These transistors share a common gate that is tied to the well in which the transistors are sitting. No matter which way the light changes, one of the diodes will be reversed biased. This slow adaptation insures that transistor mismatches will not be amplified by feedback from the resistive network. Continuous adaptation of this kind provides a “single point correction” at every operating point.

The response of an isolated photoreceptor is shown in Figure 3. The action of the resistive network was simulated by connecting the feedback capacitor, C_H , directly to the the output of the photoreceptor. This condition is analogous to a fullfield illumination change with the time constant of the resistive network set very fast. The circuit performed largely as expected. However, the time constant of adaptation is faster for decrements in light level than it is for increments in light level. We believe that the difference can be accounted for by a photon-generated leakage current that pulls the well containing the diode-connected transistors towards ground. At rest, the leakage current is supplied by the inverter via Q4. When the output of the inverter goes low, Q4 no longer counteracts the leakage current, which then flows onto the gate of Q1.

2.2 Results

The output of the adaptive silicon retina is the difference between the spatio-temporal average computed by the resistive grid and the receptor output. Therefore, the retina shows an spatially antagonistic center-surround response. If we compare the spatial response of the adaptive retina to that of the Mead and

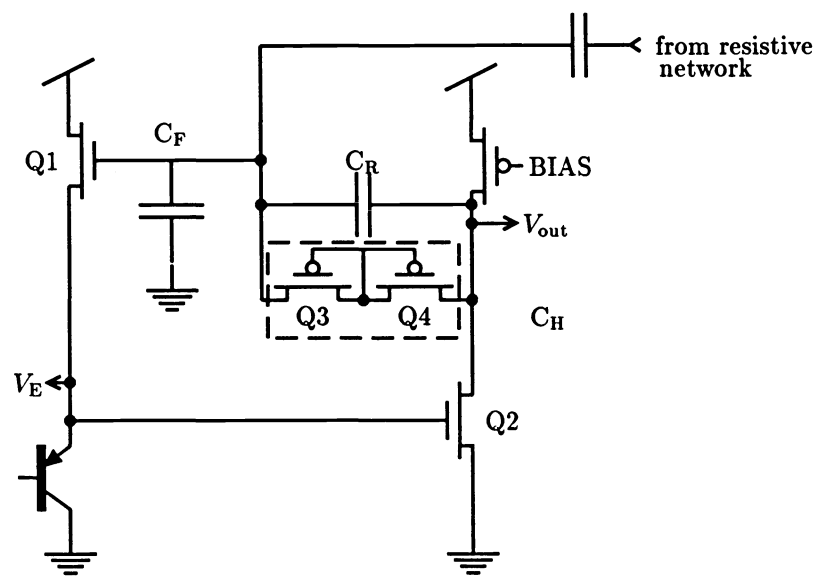


Figure 2: Schematic of a photoreceptor. The number of collected photons is amplified by the β of the bipolar transistor. The current of the bipolar transistor supplied by the current through the source follower transistor, Q1. The gate of Q1 is modulated by three parallel feedback elements, C_R , C_H , and a pair of back-to-back diodes, Q3 and Q4, whose gates are tied to the well.

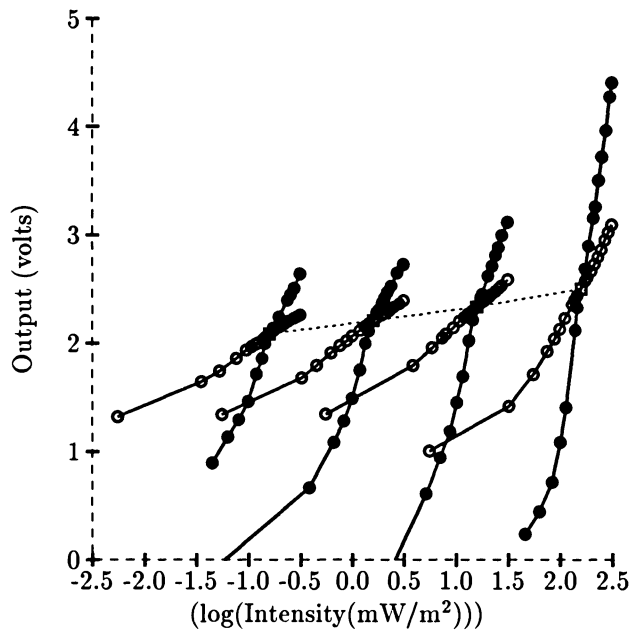


Figure 3: Response of the photoreceptor to changes in light intensity of a light emitting diode (l.e.d). Each set of curves was taken within the same range of currents through the l.e.d. Neutral density filters were used to shift the light intensity over four orders of magnitude. The dotted line indicates the receptor's DC response to illumination measured after fully adapting to that illumination level. The gain of the adapted response of the receptor is low. Solid circles indicate the instantaneous response when C_H is tied to a fixed reference voltage. Light level was briefly displaced and peak response was measured. The gain is slightly different in the upward and downward direction due to differences in parasitic capacitances in the diode feedback element for upward and downward steps. Open circles show the receptor response when capacitive coupling from the output of the inverter to the gate of Q1 is increased by connecting the inverter output to C_H . The gain of the receptor is reduced. This condition models the effect of uniform illumination on a retina with feedback from the horizontal cells to the receptors. The gain of the receptor to changes in intensity increases at high light levels because Q1 is going above threshold.

Mahowald retina shown in Figure 4, we see that the adaptive retina response is more regular. Pixel to pixel variations are due to mismatches in the output amplifier and the small variations in the receptor output that remain after adaptation. The position of the edge is much more apparent in the response of the adaptive retina than it is in the output of the Mead and Mahowald retina. In addition to DC variation, the pixels have a distribution of different gains for changes in intensity. The variation in gain is not corrected by adaptation. We believe that the variation can be reduced by increasing the size of the C_R feedback capacitor.

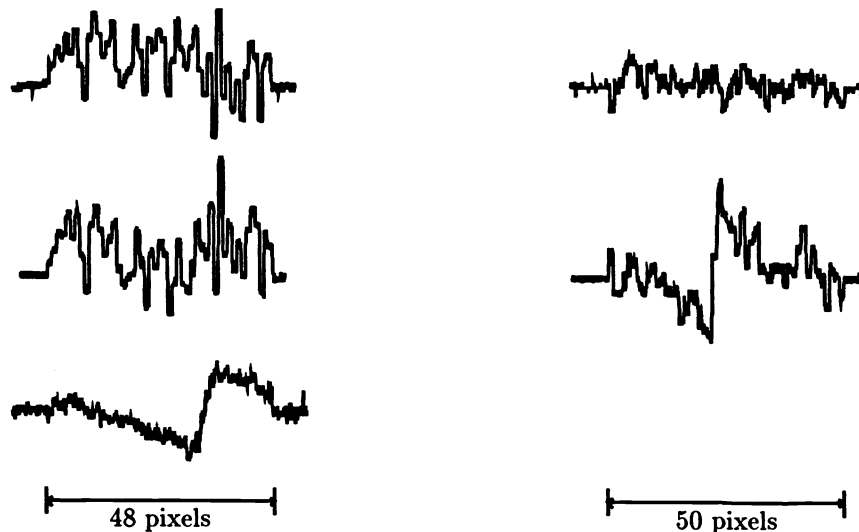


Figure 4: Comparison of the edge responses of the Mead and Mahowald retina and the adaptive retina. Stimulus was a 0.2 log unit, one-dimensional step in intensity. Data was taken by multiplexing the analog responses of a row of pixels perpendicular to the intensity edge to a digital storage scope. The current output of the chip was converted to a voltage by an off-chip sense amplifier. The gain of this amplifier, hence the voltage scale of the response, is arbitrary. The DC offsets of the photoreceptors and the output amplifiers appear as small differences in the responses of different pixels. The spatial averaging area of the resistive grids in both retinæ were large. (a) Response of the Mead and Mahowald retina. Top trace shows the response to a uniform field. Middle trace shows raw edge response. The bottom trace shows the edge response minus the uniform field response. The position of the edge is visible only after performing this differencing operation off-chip. (b) Response of the adaptive retina. Top trace shows the response to a uniform field to which the retina was adapted. Middle trace shows raw edge response. The edge was centered roughly around the intensity of the uniform field. The position of the edge is immediately apparent.

3 Conclusion

The image processing performed in the outer retinal layers, which is required to reliably signal small changes in image brightness, has important consequences for the representation of visual information. Unlike a camera, the retina reports contrasts, rather than absolute intensity. By measuring visual signals relative to a reference value, the retina increases the dynamic range of the receptors while remaining sensitive to small contrasts. In order to use a spatial average to enhance the gain of the receptor, the retina must adapt continuously to cancel out mismatches between elements, which vary with the DC operating level. Transistor

mismatch is one of the most serious problems faced by analog VLSI designers, particularly those working in the power-efficient subthreshold regime. The adaptive silicon retina, which uses the same computational strategy as its biological counterpart, is another step in the evolution of dense, low-power, analog information processors.

4 Acknowledgements

Thanks to Carver Mead and Tobi Delbruck for many fruitful discussions. This work was supported by the Office of Naval Research, the National Institute of Health, the Systems Development Foundation, and the State of California Department of Commerce Competitive Technologies Program. Computation was provided by a grant from Hewlett-Packard, and Chip Fabrication was supported by DARPA, through the MOSIS service.

5 References

1. Mahowald, M. A. and Mead, C. A., "Silicon Retina," In Mead, C. A., *Analog VLSI and Neural Systems*, pp.257-278, Addison-Wesley, Reading, MA, 1989.
2. Mead, C. "Adaptive retina." In C. Mead and M. Ismail (Eds.), *Analog VLSI Implementation of Neural Systems*, (pp. 239-246) Kluwer Academic Publishers, Boston, 1989.
3. Mead, C. A., *Analog VLSI and Neural Systems*, Addison-Wesley, Reading, MA, 1989.
4. Mead, C. A. and Mahowald, M. A., "A silicon model of early visual processing," *Neural Networks*. 1:91-97, 1988.
5. Normann, R. A., and Werblin, F. S., "Control of retinal sensitivity. I. Light and Dark Adaptation of Vertebrate Rods and Cones," *Journal of General Physiology*. 63: 37-61, 1974.
6. Skrzypek, J., "Light Sensitivity in cones is affected by the feedback from horizontal cells," *Technical report, Department of Computer Science, UCLA*, 1991.
7. Werblin, F. S., "Control of retinal sensitivity. I. Lateral interactions at the outer plexiform layer," *Journal of General Physiology*. 63: 62-87, 1974.