Time-derivative adaptive silicon photoreceptor array

Tobi Delbrück and Carver A. Mead

Computation and Neural Systems Program, 139-74
California Institute of Technology
Pasadena CA 91125
Internet email: tdelbruck@caltech.edu

ABSTRACT

We designed and tested a two-dimensional silicon receptor array constructed from pixels that temporally high-pass filter the incident image. There are no surround interactions in the array; all pixels operate independently except for their correlation due to the input image. The high-pass output signal is computed by sampling the output of an adaptive, high-gain, logarithmic photoreceptor during the scanout of the array. After a pixel is sampled, the output of the pixel is reset to a fixed value. An interesting capacitive coupling mechanism results in a controllable high-pass filtering operation. The resulting array has very low offsets. The computation that the array performs may be useful for time-domain image processing, for example, motion computation.

1. TIME-DOMAIN IMAGE PROCESSING

Real-time image processing is expensive. Much of the computational load involved in computing motion parallax, optical flow, and object tracking lies in the image preprocessing, before any sophisticated global vision algorithms are applied.

Specialized parallel digital processors like the PIPE machine have been used to do real-time, time-domain, image processing. These machines are reprogrammable and flexible in their applications, and have been used to implement algorithms developed by the machine vision community,[2] and also to model biological visual function.[11] Since they use a large number of high speed, high power, digital chips, they are of limited usefulness in terms of teaching us how to *build* vision systems that map naturally onto silicon.

One of the fundamental preprocessing operations is high-pass filtering, or temporal differentiation. This operation is useful for all types of image motion computation. In this paper, we describe a circuit that computes a pure high-pass, temporally filtered version of the incident image. This time-derivative operation is accomplished with a pixel measuring 86 microns on a side, and the entire core of the chip, consisting of 68 by 43 pixels, consumes about 4 mW of power (1.4μW per pixel). The pixels have zero DC response, and the offsets between pixels are very small.

Many models of motion discrimination rely on inputs with strictly high-pass, or time-derivative, characteristics. Much of the previous work on silicon retinas has focused on the spatial aspects of image processing, and in those models, time is treated largely as a dimension over which to adapt.[5, 8, 9] In more recent work, the time-domain has been treated more centrally. Mahowald[4] reported a design which elegantly integrates both time and space into the same pixel. Her pixels have responses with both high-pass and low-pass characteristics, and they incorporate feedback from

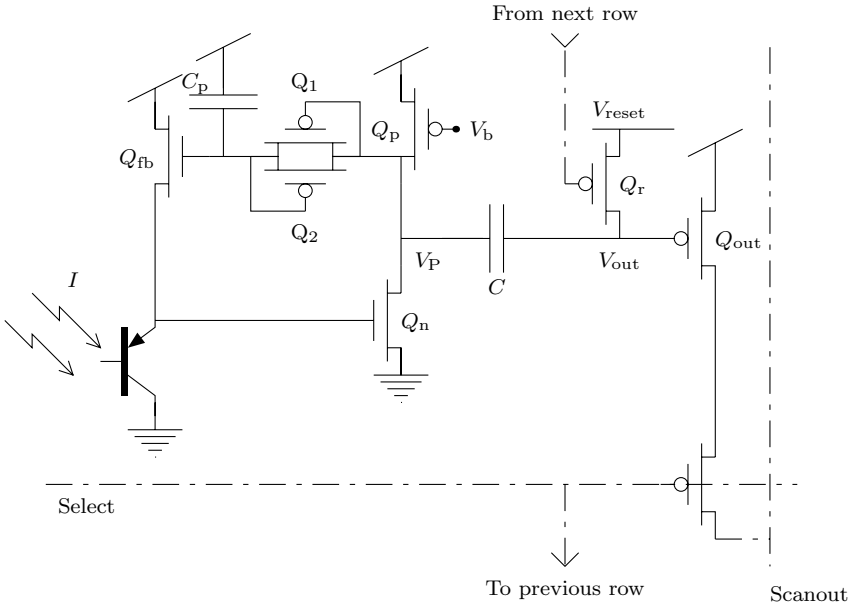

Figure 1: Time-derivative photoreceptor. V_P is output from the high-gain, adaptive photoreceptor. V_{out} is the sampled high-pass output.

the surrounding context of pixels. Mann[6] has characterized photoreceptor circuits that have a logarithmic response and a nicely controllable band-pass characteristic, and that use UV-programmable static offset compensation,[3] as in Mead's adaptive silicon retina.[8]

The primary contrast of the present work with Mahowald's and Mann's work is that the present work is completely specialized for computing a pure high-pass filtering operation, and the computation itself is intimately linked to the sampling process used to access the pixel outputs. This specialization results in an efficient implementation. The pixel reported here is 1.77 times smaller than Mahowald's pixel, and 3.5 times smaller than Mann's pixel.

2. PIXEL OPERATION

The time-derivative circuit is a high gain, self-adapting, sampled photoreceptor. Figure 1 shows the circuitry inside a single pixel. In broad outline, the transistors on the left of the figure act as a high-gain adaptive photoreceptor.[4, 1] The output from this receptor, V_P, is coupled into the circuitry on the right of the figure, which is a high-pass sampling structure that is linked to the scanning process by which data is read serially from the chip. The high-pass output from the circuit is V_{out}, and is converted a current by Q_{out} to be scanned out of the pixel. The pixel sits in a two-dimensional array. Successive rows of the array are scanned onto a monitor for display.[10]

In the rest of this section, we describe first the operation of the adaptive photoreceptor, and then the high-pass filtering operation.

2.1. Adaptive photoreceptor operation

Light is absorbed by the phototransistor, producing a current that is supplied by a feedback transistor Q_{fb}. The phototransistor base is formed by a well, and the emitter is formed by a piece of active diffusion sitting in the well. The collector is the substrate. The current gain of this type of phototransistor is quite high, typically several hundred, and the quantum efficiency near one at visible wavelengths.[7] Even though only 3% of the area of the actual pixel is not covered by second metal, the chip is sensitive enough to work in a dimly lit room while looking through an $f/4$

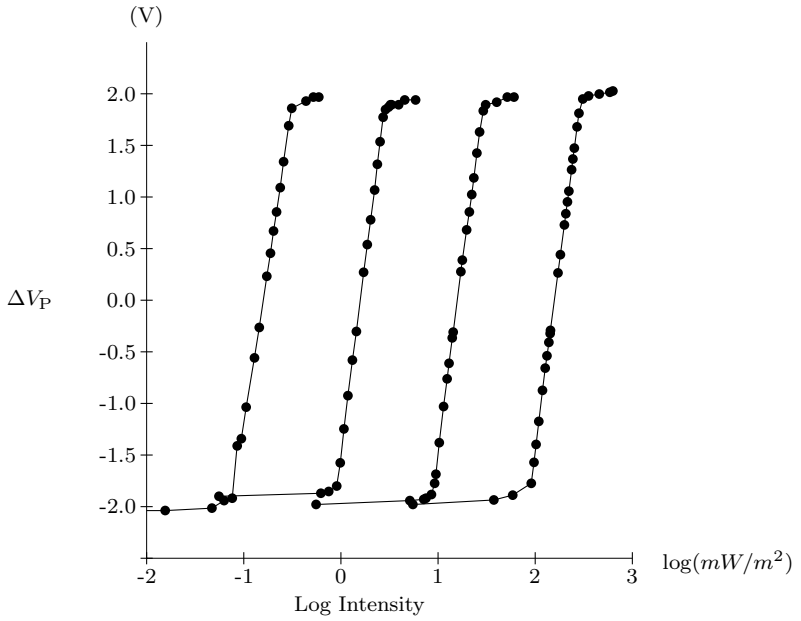

Figure 2: Response amplitudes to step changes in incident intensity, at different adaptation levels. For each curve, we plot the amplitude of the step response on the ordinate, versus the intensity to which we step on the abscissa. The curves were taken at different adaptation intensities. For example, the second curve from the right was taken at an adaptation intensity of 1.15 log units. A decrement of intensity to 1 log unit produced a response of -1.6 volts amplitude. The highest intensity on the scale corresponds to the illuminance of a scene under typical fluorescent lighting conditions.

lens.

$Q_{\rm fb}$ is configured as a source follower. For typical intensities, $Q_{\rm fb}$ operates in subthreshold, so the voltage at the emitter of the phototransistor changes logarithmically with changes in the intensity. The gate voltage of $Q_{\rm fb}$ is constant for short times, so when the intensity increases by some fraction $I_{\rm new}/I_{\rm old}$, the emitter voltage decreases by $\frac{kT}{q}\ln(I_{\rm new}/I_{\rm old})$.

Transistors $Q_{\rm n}$ and $Q_{\rm p}$ amplify and invert the voltage changes at the emitter of the phototransistor, producing the output $V_{\rm P}$ of the photoreceptor. A bias voltage, $V_{\rm b}$, determines the cutoff frequency for the receptor by setting the bias current in the inverting amplifier. The voltage gain of the inverting amplifier is typically -50 for transistors that are 6 microns long.

The receptor output voltage, $V_{\rm P}$, is fed back through $\rm Q_1$ and $\rm Q_2$ to capacitor $C_{\rm p}$, which stores a temporally low-pass filtered version of the receptor output $V_{\rm P}$. In quiescence, the voltage at the emitter will settle to whatever voltage it takes to make $Q_{\rm n}$ sink as much current as $Q_{\rm p}$ is supplying. Similarly, the voltage on $C_{\rm p}$ will settle to whatever voltage it takes to make $Q_{\rm fb}$ supply the current sunk by the phototransistor.

Transistors $\rm Q_1$ and $\rm Q_2$ act as a very high resistance element for small voltages, and as a low resistance element for large voltages. The current through this element increases exponentially with the voltage in both directions, but is extremely small for small voltages. For small changes in intensity, over moderate time scales, the receptor does not adapt, since no current flows onto or off $C_{\rm p}$. Hence, for small changes in intensity, the feedback loop is open, and the receptor is maximally sensitive. For large intensity changes, or over long time scales, the receptor adapts to the new intensity. This behavior is sensible for a system that must be sensitive, yet also have large dynamic range and an ability to rapidly adjust to new conditions. This receptor is a simplified version of one described in Delbrück and Mead,[1] and is very similar to one described in Mahowald.[4]

Figure 2 shows response characteristics of the receptor for different adaptation levels, plotted in a style similar to that used by biologists in showing the transfer characteristics of retinal receptor cells. We see that the receptor can adapt

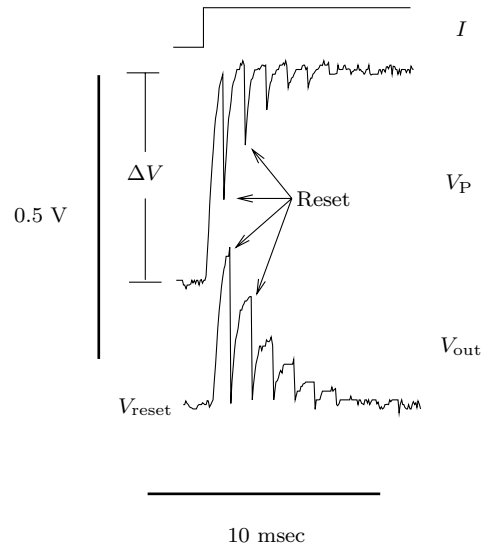

Figure 3: Response of photoreceptor ($V_{\rm P}$) and output ($V_{\rm out}$) to a step increase of intensity I.

over more than four decades of background intensity. The gain of any adaptive, logarithmic sensor is conveniently measured in units of output change per change in log input; measured on this scale, the receptor discussed here has a gain of about 3 volts per e-fold intensity change.

The adaptive element consisting of $\rm Q_1$ and $\rm Q_2$ has the undesirable property of introducing large random offsets in the DC output of the receptor. Because the incremental resistance of the adaptive element is extremely high for small differential voltages, any tiny leakage current or transistor mismatch causes a large output voltage variation. If the adaptive element is constructed using two well-type transistors, as shown in Figure 1, the predominant offset is caused by diode leakage current in $\rm Q_1$ and $\rm Q_2$. The effect of this leakage is to cause a large negative offset between the output of the receptor, $V_{\rm P}$, and the voltage on $C_{\rm p}$. This leakage-current effect is exacerbated by light scattering into the well in which $\rm Q_1$ and $\rm Q_2$ sit, and the effect is quite pronounced even if the only opening in the metal covering the circuit is directly over the phototransistor emitter. Other variants of this adaptive element are possible using native-type transistors, or a combination of native- and well-type transistors, but we have found none that have ideally symmetric operation. For these reasons, a receptor array constructed solely from these adaptive receptors is not very useful. Despite these large DC offsets, the response of the photoreceptor to *changes* in intensity is very predictable.

2.2. High-pass filtering operation

The output $V_{\rm P}$ from the adaptive photoreceptor is capacitively coupled to the temporal-differentiation part of the circuit. The computation of the derivative is intimately linked to the sampling of the pixel output. After the pixel output $V_{\rm out}$ is sampled, by driving the select line low, the pixel output node $V_{\rm out}$ is reset to the voltage $V_{\rm reset}$ by driving the gate of $Q_{\rm r}$ low. This sampling and reset sequence happens naturally, during the scanning of the array. The rest of the time, $V_{\rm reset}$ is high. This scheme results in a high-precision, low-offset output, since the pixel will output nothing at all if nothing changes in the photoreceptor, and the offset will be limited to the offset in the single transistor $Q_{\rm out}$. This single transistor may be made large with little area penalty.

At first glance, the time-derivative operation appears to be obvious: the pixel simply outputs any change in the photoreceptor output since the last sample was taken. Indeed, the operation does happens this way under certain conditions. Under other conditions, however, the operation is more sophisticated.

Figure 3 shows how $V_{\rm P}$ and $V_{\rm out}$ behave during the sampling of a single step-increase of intensity. The receptor

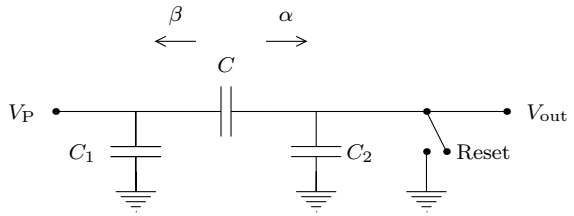

Figure 4: Abstraction of sampling process. C_1 and C_2 are the parasitic capacitances not shown in Figure 1.

output V_P initially increases by ΔV in response to the intensity increase. The increase in V_P is coupled to the output node V_{out} and is sampled. The sampling occurs just before the reset. When the output node V_{out} is driven back down to V_{reset} by the reset signal, the output of the receptor itself, V_P, is driven down also, since it is capacitively coupled to the output node. This action will not leave the receptor output where it wants to be; instead it will rebound towards its desired output voltage ΔV. This rebound will capacitively couple to the output node, in effect creating a new sample. This cycle repeats itself ad infinitum.

Each time a sample of the receptor output is taken, the sample will be smaller than before. The reason for this behavior is that the coupling between the receptor and the pixel output is not perfect – it is limited by the parasitic capacitance of the output node. Similarly, the coupling back from the pixel output to the receptor is limited by the receptor parasitic capacitance. Hence, each time the receptor drives the output node, and each time the output node drives the receptor, some *fraction* of the original signal will be lost. After N samples, the original output signal will decay to γ^N of its starting value, where γ is the inefficiency of the coupling. This kind of decay is simply a discretely sampled exponential decay. Hence, the output signal will decay exponentially in time. This operation forms the high-pass, time-derivative signal.

Figure 4 shows an abstraction of the circuitry involved in the differentiation process. We define the capacitive division ratios

$$\begin{aligned} \alpha &= \frac{C}{C + C_2}, \\ \beta &= \frac{C}{C + C_1}. \end{aligned} \tag{1}$$

α is the fraction by which a change at the receptor node gets reduced at the output node. Similarly, β is the fraction by which a change at the output node gets reduced at the receptor node.

Suppose that the initial step increase in intensity results in a change ΔV at the receptor node, V_P. This change results in a change $\alpha\Delta V$ at the output node, V_{out}. After the reset, the receptor gets driven back down by $\alpha\beta\Delta V$. The receptor rises back up by this amount, driving the output up by $\alpha(\alpha\beta)\Delta V$, completing one sampling cycle. After N samples, the output will be reduced to $\alpha(\alpha\beta)^N\Delta V$.

If the sampling frequency is $f = 1/T$, then we may rewrite the preceding expression for the output voltage as

$$\begin{aligned} V(t) &= \alpha(\alpha\beta)^{t/T}\Delta V \\ &= \alpha e^{(t/T)\ln(\alpha\beta)}\Delta V \\ &= \alpha e^{-tf\ln(\frac{1}{\alpha\beta})}\Delta V. \end{aligned}$$

We can identify the effective time constant of the exponential decay of the response:

$$\tau = \frac{1}{f\ln(\frac{1}{\alpha\beta})}. \tag{2}$$

Hence, the higher the sampling frequency, the lower the time constant. Both α and β are smaller than one; the closer they are to one, the longer the time constant.

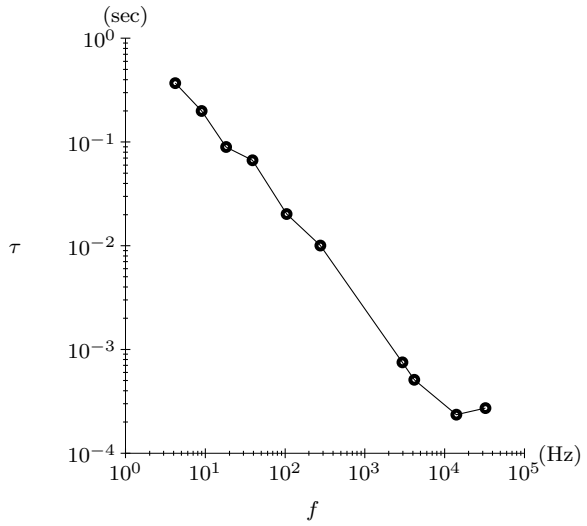

Figure 5: Measured time constant τ as function of sampling frequency f. Stimulus was a step increase of intensity. Reset pulse was always 8.9 μs long, and output node was reset to 2.75 V. Time constant was obtained my measuring $\alpha\beta$ ratio, and then applying Equation 2.

If we assume that both C_1 and C_2 are equal and are much smaller than C,

$$C' = C_1 = C_2 << C,$$

then we can easily compute an approximation to the above formula for τ:

$$\tau \approx \frac{1}{f}\frac{C}{2C'}.$$

All of this analysis, however, can obscure the essential operation of the circuit. If we look at Figure 3, we see that the pixel output V_{out} is reduced by the fraction $\alpha\beta \approx 23/33$ by each sample. A simple computation based on Equation 1 shows that the ratio $\frac{C}{C'}$ that would result in this ratio is about 5. This ratio is consistent with the layout and sizing of the capacitors and transistor gates in the test pixel, demonstrating that, to first order, the preceding analysis is correct.

For this analysis to be correct, it is important that two conditions be met. First, the pixel output must not change *during* the reset period. Otherwise, the response will decay more quickly than the above analysis would indicate. Hence, when we actually use this pixel, we set the bias current determined by V_{b} low enough that the output does not move much during the reset phase. This setting has the additional benefit that it filters out the ubiquitous 60 cycle intensity variations due to incandescent filament heating and fluorescent bulb discharges. Second, the sampling frequency must be high enough that the pixel itself does not adapt during the sampling period. Otherwise, we will sample the pixel adaptation process, and not intensity contrast change. As long as the sampling frequency is larger than a few Hz, this second condition will be met.

Figure 5 shows measurements of the time constant τ as a function of the sampling frequency f. We see that the time constant is proportional to $1/f$ over most of the range, as we expect. For high sampling frequencies f, the time constant becomes constant, since it is limited by the bias current determined by V_{b}. For low sampling frequencies, τ is limited by the adaptation of the receptor itself.

2.3. Complete system operation

The pixel is incorporated in a two dimensional array. The array is scanned out onto a monitor using a fully monolithic video scanner.[10] During the scanning, successive rows of the array are scanned onto the monitor. While each row is

Figure 6: Photographs of the monitor display, showing chip output in response to a moving hand, and in response to a set of dark diamond-shaped squares moving to the right. Vertical lines are due to clock feedthrough in the MOS switches used to multiplex the output.[10] At the scanning frequency of $f \approx 60$ Hz, the time constant τ is long enough that the persistence of a response extends over multiple frames (also see Figure 5). The chip is fabricated in 2μ double poly, double metal CMOS; there are 68 rows and 43 columns of pixels. The die size is 4.6 by 6.8 mm.

being scanned, the previous row is reset. Hence, the reset phase of the scanning process only occupies approximately $1/M$ of the total array scan time, where M is the number of rows. This is a small fraction, so the condition of the previous paragraph is easily met.

To be properly appreciated, the output from this circuit must be viewed in real time. A dynamic image loses a great deal of saliency when statically viewed. Figure 6 makes an attempt to show how the chip responds to some typical moving input patterns.

3. CONCLUSION

We have designed and tested a silicon receptor array that temporally high-pass filters the incident image. This imager could be useful as a preprocessing stage for later computation of image motion, optical flow, motion parallax, and other derived quantities that rely on dynamic image features.

4. ACKNOWLEDGMENTS

This work was supported by the Office of Naval Research (grant NAV N00014-89-J-1675), 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] T. Delbrück and C.A. Mead, "An Electronic Photoreceptor Sensitive to Small Changes in Intensity," D.S. Touretsky (ed.) *Advances in Neural Information Processing Systems,* vol. 1, pp. 720–727, Morgan Kaufmann, San Mateo, 1988.

[2] D.A. Fay and A.M. Waxman, "Real-time Early Vision Neurocomputing," *1991 IEEE INNS Intl. Joint Conf. on Neural Networks – Seattle*, vol. 1, pp. 621–626, IEEE, Piscataway, 1991.

[3] L. Glasser, "A UV Write-Enabled PROM," H. Fuchs (ed.), *1985 Chapel Hill Conference on VLSI,* pp. 61–65, Computer Science Press, Rockville, 1985.

[4] M.A. Mahowald, "Silicon Retina with Adaptive Photoreceptor," *Proc. SPIE/SPSE Symposium on Electronic Science and Technology: from Neurons to Chips,* vol. 1473, April 1991.

[5] M.A. Mahowald and C.A. Mead, "Silicon Retina," C.A. Mead, *Analog VLSI and Neural Systems*, pp. 257–278, Addison-Wesley, Reading, 1989.

[6] J. Mann, "Implementing Early Visual Processing in Analog VLSI: Light Adaptation," *Proc. SPIE/SPSE Symposium on Electronic Science and Technology: from Neurons to Chips,* vol. 1473, April 1991.

[7] C.A. Mead, "A Sensitive Electronic Photoreceptor," H. Fuchs (ed.), *1985 Chapel Hill Conference on VLSI,* pp. 463–471, Computer Science Press, Rockville, 1985.

[8] C.A. Mead, "Adaptive Retina," C. Mead and M. Ismail (eds.), *Analog VLSI Implementation of Neural Systems*, pp. 239–246, Kluwer Academic Publishers, Boston, 1989.

[9] C.A. Mead and M.A. Mahowald, "A Silicon Model of Early Visual Processing," *Neural Networks,* vol. 1, pp. 91–97, 1988.

[10] C.A Mead and T. Delbrück, "Scanners for Use in Visualizing Analog VLSI Circuitry", *Analog VLSI for Signal Processing,* in press, to appear sometime 1991.

[11] F. Werblin, "Synaptic Connections, Receptive Fields, and Patterns of Activity in the Tiger Salamander Retina," *Investigative Ophthalmology and Visual Science*, vol. 32, no. 3, pp. 459-483, 1991.