

SILICON RETINA FOR AUTOFOCUS

Tobi Delbrück

Institute for Neuroinformatics (INI)

ETH/Univ. Zürich

Winterthurerstr. 190

8057 Zürich, Switzerland

tobi@ini.phys.ethz.ch, <http://www.ini.unizh.ch/~tobi>

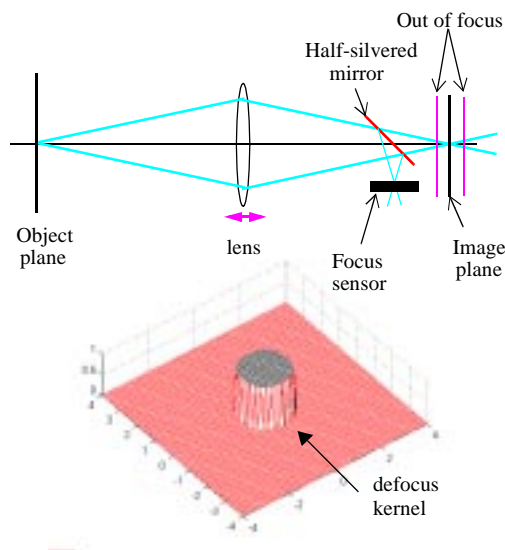
ABSTRACT

This paper describes a silicon retina that measures image sharpness. The idea is to use this sensor in the image plane of an autofocus camera system [3]. The chip has 25×26 pixels, each $(60 \mu\text{m})^2$, is fabricated on a $(2.2\text{mm})^2$ $1.2 \mu\text{m}$ CMOS process, and consumes $100 \mu\text{A}$ with 5 V supply.

1. ALGORITHM

Image defocus corresponds to a filtering operation on the image with a circular cookie-cutter kernel whose diameter is proportional to the distance the image is away from the plane of focus and to the lens f number—the ratio of focal length to lens aperture. This geometry and the kernel are shown in Figure 1. This figure also shows how a focus sensor can sit in the optical path of an imaging system by using a half silvered mirror.

FIGURE 1 Defocus geometry and blur kernel.



In 1968, Berthold Horn [4] studied the problem of measuring image sharpness for the purpose of autofocusing. This study was in the context of Project MAC, a vision system based on a Vidector (a kind of scanning photometer), and an early LISP machine. He used information from the discrete Fourier transform of the image to measure the spatial frequency energy. As the image is defocused, the high frequencies are cut off. The energy peaks at the point of best focus. In the chip reported here, we use the simplest energy measure

(and one that is natural for us to build) by simply summing up the energy in the spatial gradient:

$$E = \int_{\text{image}} (\nabla I)^2$$

where E is the energy measure and ∇I is the gradient of the image. This functional of the image results in a peaked function of focus setting. It is important to note that the functional of ∇I must be expansive, so that higher contrast edges are weighted more.

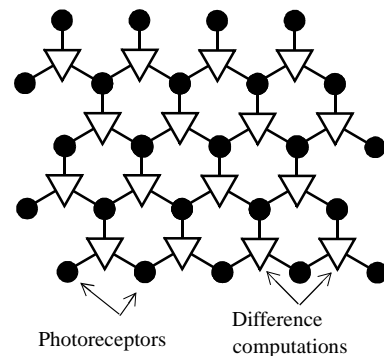
In practice, the pixel spacing has an important effect on this measure because we can only estimate the derivatives of spatial frequencies up to a limit set by the spacing. Any spatial wavelengths in the image that are shorter than twice the pixel spacing will be aliased. So the sampling grid will place a limit on the precision by which we can locate the position of maximum sharpness. In the case of a single perfectly sharp edge, the pixel spacing prevents us from focusing the edge better than the pixel spacing. But an edge that is extended over an array of pixels and that is not aligned with the array, or a band-limited image, will be sampled at multiple points along their gradients, and we can do better than the case of a single pair of pixels.

2. CHIP ARCHITECTURE

I start by describing the architecture of the chip in this section. In Section 3, I describe the elements used on the chip.

Figure 2 shows how the pixels are arranged in a hexagonal array. Each pixel has an adaptive photoreceptor and a 3-input antibump circuit. The antibump circuit computes an *expansive* measure of the absolute differences between three neighboring pixels. The sharper the image, the larger these differences. The expansive measure means that a sharp image (where pixel differences are greater) result in a larger sharpness measure. The focus measure is the sum of the antibump output currents over all the pixels.

FIGURE 2 Focus chip architecture



3. THE CIRCUIT ELEMENTS

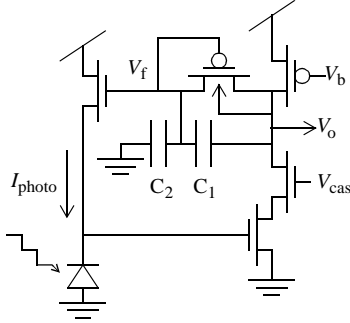
The pixel uses only two circuit elements: an 5-transistor adaptive photoreceptor [1] and a 7-transistor antibump circuit [2]. In addition, the chip includes a bias generator that generates all circuit bias currents and an output circuit that produces a pulse-frequency modulated output from the summed focus signal.

3.1 Adaptive photoreceptor [1]

Over a range of 6-7 decades of illumination, the photoreceptor shown in Figure 3 outputs a signal V_o whose variations in response to scenes with varying reflectivity are nearly invariant to illumination level. The output of the receptor V_o is referenced around the past history of the illumination, and has a transient gain determined by C_2/C_1 of about 1 volt/decade. The DC gain is much lower, only 100 mV/decade. The lower limit of illumination is determined by dark current and desired response speed; it is presently useful down to bright moonlight conditions.

The photoreceptor adaptation means that in response to a static image, the focus signal eventually disappears. In practice, the adaptation is much slower than the dynamics of a focusing operation.

FIGURE 3 Adaptive photoreceptor.



3.2 Antibump energy measurement circuit

The spatial gradient energy measurement is done by the antibump circuit with 3 inputs shown in Figure 4. The antibump output current grows large only when at least two of the inputs V_i differ sufficiently. The circuit used here adds another leg to the one I described in the original paper [2]. The output current I of this circuit can be derived from the fundamental subthreshold transistor equation [7]

$$I_{ds} = I_0 e^{\kappa V_g - V_s} (1 - e^{-V_{ds}}).$$

The result of this computation is

$$\frac{I}{I_b} = \frac{1}{1 + \frac{S}{\sum e^{\kappa V_i} \sum e^{-\kappa V_i}}},$$

where I_b is the bias current, $\kappa \approx 0.8$ is the back gate coefficient, and S is the effective strength ratio (effective W/L ratio) between the stacked transistors and the output transistors. Voltages are in units of kT/q . This expression only depends on voltage differences, and because the three voltage differences must sum to zero, it really only depends on two voltage differences—call them V_{12} and V_{13} . Two-

dimensional plots of the shape of this function are shown in Figure 5.

FIGURE 4 3-input antibump circuit. Inputs $V_{1,2,3}$, output I .

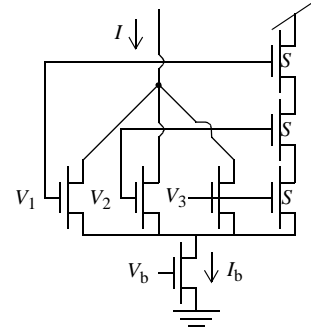
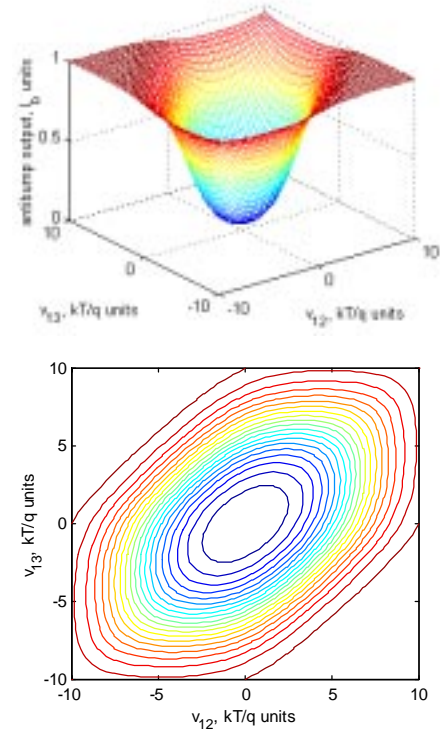


FIGURE 5 3-input antibump circuit theoretical output, mesh and contour plots. $S=100$, $\kappa=0.7$.



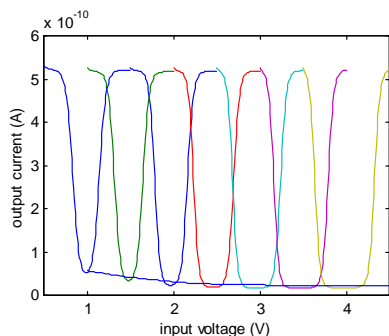
The parameter S is crucial—it determines the shape of the function near zero and the dynamic range of the computation. We only care about $S \gg 9$. In this case, when $V_{12} = V_{13} = 0$,

$$\frac{I}{I_b} \approx \frac{9}{S}$$

The dynamic range, which we define as the ratio between maximum and minimum output current, is $S/9$. If S is too small, then the dynamic range is limited but the function has a nice parabolic shape around zero. When S is too large, the function has a large flat region around zero, and low contrast images produce no detectable output, at least on a linear scale. A happy balance exists somewhere in between.

S is the “effective strength ratio.” It is *not* the drawn W/L ratio, because the short and narrow channel effects have a very pronounced effect. Short transistors act much shorter than their drawn length, and narrow transistors act much narrower. It is easy to obtain $S > 10^5$ by using minimum length in the stacked transistors and minimum width in the output transistors, but this large S results in a very wide flat spot around zero. Spice (BSIM 3v3) is very poor at predicting the measured value of S . But there is a saving grace: the effective S depends on the common mode input voltage. Figure 6 shows measured antibump output currents for different common mode input voltages for a 2-input antibump circuit. We can see from these results that the effective S becomes larger as common mode increases. Both short and narrow channel effects become more pronounced with increased gate-substrate voltage [8]. By moving the photoreceptor common mode output voltage, we can, in principle, optimize the dynamic range and sensitivity of the focus measurement.

FIGURE 6 Measured antibump output current for different common mode input levels. Each curve shows output with one input fixed, the other sweeping around it. Bottom curve shows output for zero differential input while sweeping common mode.



3.3 Output circuit

In order to interface to inexpensive microcontrollers, the focus signal current computed by the chip core is input to a pulse frequency generator. A simple ADC may be formed by connecting the resulting pulse train to the external counter clock input on a microcontroller.

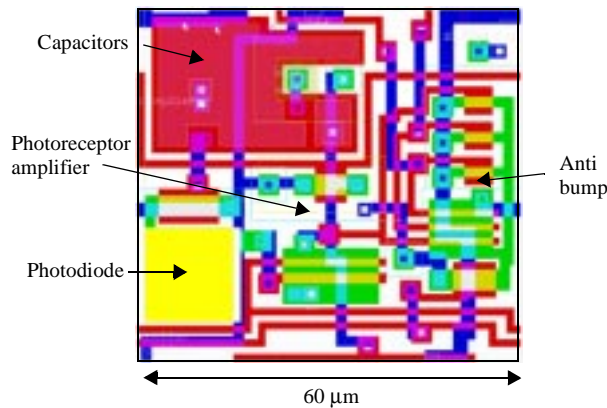
3.4 Bias generator

The focus chip has a bias generator that generates all internal bias currents, so that they are nearly independent of threshold voltage, supply voltage and temperature. The bias circuit is based on a beta-multiplier “Vittoz loop” that generates a known master reference current [5] using a single external carbon resistor. A combination of on-chip diffusion resistance and off-chip carbon resistance – with opposite temperature coefficients – provides some degree of temperature compensation. A Vittoz pseudo-resistive divider [6] derives the other bias currents from this master current.

3.5 Layout

Each pixel in the 25 by 26 array measures $(60\mu\text{m})^2$; the chip is built in a $1.2\mu\text{m}$ MOSIS [9] process. The layout of one pixel is shown in Figure 7.

FIGURE 7 Focus sensor pixel



4. MEASUREMENTS

In each of the following figures, the summed output current from all the pixels is converted into a voltage using a $10\text{ M}\Omega$ resistor. Figure 8 shows the response of the focus chip to a square wave grating pattern in response to focus changes, for different f numbers. As we decrease the lens aperture, the depth of field increases, so the focus function widens. I don't know why it also shifts to peak at closer distance. Figure 9 shows the response to the same grating pattern versus grating contrast, at the optimum focus setting. The shape of the response follows the shape of the antibump characteristic, modified by the display screen and photoreceptor response properties. Figure 10 shows the response versus grating spatial frequency. The response is linear with spatial frequency because more and more edges are visible to the chip, until spatial aliasing and display screen limitations set in.

FIGURE 8 Measured response of focus chip to defocus.

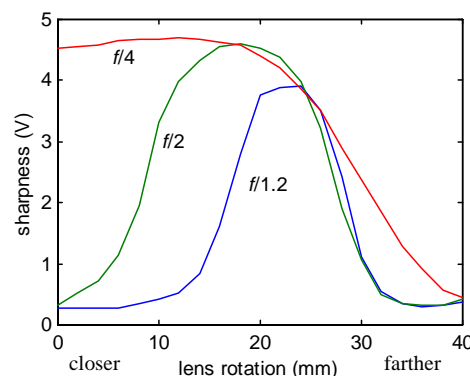


FIGURE 9 Measured response of focus chip to contrast

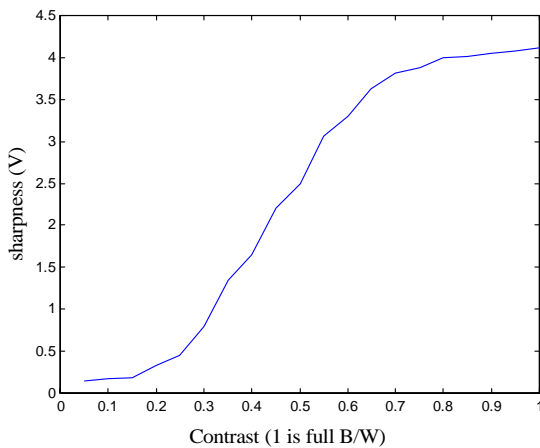
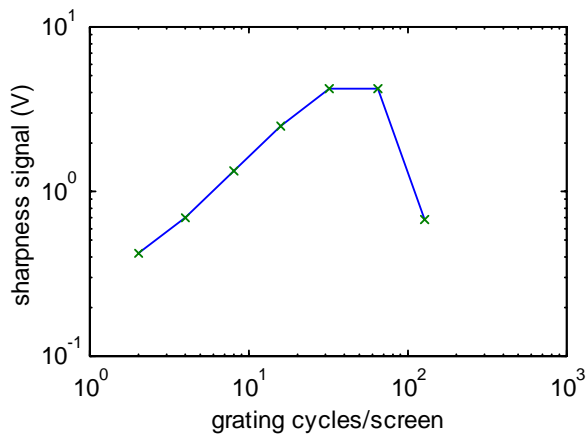


FIGURE 10 Measured response to grating spatial frequency.



5. CONCLUSIONS

My earlier work on an aVLSI focus sensor was directed at studying the control system underlying biological accommodation [3]. The focus sensor itself was a very primitive one dimensional retina that predated adaptive photoreceptors and bump circuits. The present chip is a vast improvement. A simple scaling from the obsolete 1.2 μm process used for the present chip to a mainstream 0.35 μm process would result in a pixel size of about 18 μm , allowing an array of $(110)^2$ pixels on a 5 mm^2 chip. That die would cost about 25 cents in large volume, based on today's wafer prices. It is such an obvious application of silicon retina technology that it seems worth the effort to follow this project through to the building of some optics and a microcontroller-based control system.

6. Acknowledgments

I thank Eric Vittoz and Friedrich Heitger of CSEM, Neuchatel for supporting this research for the past year. I also thank Bic Schediw of Synaptics Inc., Andre Van Shaik of the University of Sydney, and Olivier Landolt at Caltech for help with the bias generator design, and Giacomo Indiveri here at INI for the inspiration that chips with scalar or logic outputs are a lot easier to deal with than imagers. I thank Shih Chii Liu for editorial assistance and inspiration.

7. References

- [1]. T. Delbrück and C.A. Mead (1994). Analog VLSI Phototransduction by Continuous-Time, Adaptive, Logarithmic Photoreceptor Circuits. Caltech CNS Memo #30 <http://www.pcmp.caltech.edu/anaprose/tobi/recep>
- [2]. Delbrück, T. (1993) "Bump circuits for computing similarity and dissimilarity of analog voltages," Caltech CNS memo #26. <http://www.pcmp.caltech.edu/anaprose/tobi/bump>
- [3]. Delbrück, T. (1989) "A chip that focuses an image on itself," in *Analog VLSI Implementation of Neural Systems*, Kluwer Academic Publishers, pp. 171–188. <http://www.pcmp.caltech.edu/anaprose/tobi/focus>
- [4]. Horn, B. (1968) "Project MAC, Focusing," *MIT Artificial Intelligence Memo No. 160*, May 1968.
- [5]. Vittoz, E.A. and J. Fellrath (1977) "CMOS analog integrated circuits based on weak inversion operation," *IEEE Journal of Solid State Circuits*, vol. SC-12, no. 3, June 1977, pp. 224-231
- [6]. Vittoz, E.A. and X. Arreguit (1993) "Linear networks based on transistors," *Electronics Letters*, vol. 29, no. 3, 4 Feb 1993, pp. 297–298.
- [7]. Mead, C.A. (1989) *Analog VLSI and Neural Systems*. Reading, MA: Addison-Wesley.
- [8]. Tsividis, Y. P. (1987) *Operation and Modeling of the MOS Transistor*, New York, NY: McGraw Hill
- [9]. MOSIS–Metal Oxide Semiconductor Implementation Service. <http://www.mosis.org>