INTEGRATED VISION SENSOR FOR DETECTING BOUNDARY CROSSINGS

S. Zahnd, P. Lichtsteiner, T. Delbrück

Institute of Neuroinformatics, University and ETH Zürich
Winterthurerstr 190, Zürich 8057, Switzerland
sam@ini.phys.ethz.ch

ABSTRACT

What is the simplest solution for counting objects that cross a field of view? We reduce the problem to detecting localized illumination change on two spatially separated borders. Each circular border is formed from pixels that measure the magnitude of local changes in illumination relative to the global average. These local change signals are summed over each border and then amplified. Off-chip processing will determine the movement direction by cross-correlating these signals with appropriate delays. We have implemented the crossing-detection functionality with low complexity and power consumption using 5.3mm^2 on a standard 5V 0.8µm double-poly double-metal CMOS process. Crossing objects that cover 1/8 of the border with contrasts down to a few percent can be detected.

1. INTRODUCTION

Standard occupancy detectors that use discrete pyroelectric passive infrared (PIR) sensors equipped with plastic multi-lens arrays give a reliable estimate of the presence of moving IR sources, but can say nothing about movement direction or location. (That is why presence detectors can forget about you and switch off the lights in your office even though you are still sitting there.) Our chip is a vision-based person counter designed to improve the reliability of standard PIR-based occupancy detectors by providing an estimate of the present number of people occupying a room. Standard image acquisition and processing hardware is not economically justified for this application. We developed a dedicated image sensor chip with very low off-chip processing and resource requirements. The key to keep the computation simple is to observe only the boundary of the selected area. By classifying movement events as move-in or move-out, we can estimate the present number of people in the room. We make no attempt to distinguish multiple simultaneous crossings. We intend to combine our sensor with a PIR sensor. When the PIR sensor no longer detects movement, our sensor will provide additional information on the likely number of people still present. At the same time, the PIR sensor signal can be used to reset an erroneous nonzero count to zero when sufficient time passes without any PIR movement signal. The occupancy count obtained using our sensor will simply improve the reliability of the occupancy estimation.

Since our sensor should work in a relatively uncontrolled environment, with varying illumination and object speed or size, it has to provide the following functionality:

- ***Illumination Change*** information must be extracted from the scene, independent of the absolute illumination and contrast polarity. The detection needs to be highly sensitive in order to sense contrasts of down to few percent. The sensor must adapt to the absolute illumination level, which can change several orders of magnitude; and
- ***Global Illumination Change Rejection*** must be provided in order to be immune against illumination changes that affect the whole field of view.

Our approach uses two spatially separated illumination change detectors that are provided by building two concentric rings of photosensors, each of which used to try to detect crossing objects. The scalar signals obtained from these two rings are very low bandwidth and can be correlated in real time by the cheapest microcontroller. In future work we plan to correlate these signals off-chip with a variety of temporal delays (in contrast with [1;2]) to interpret them as move-in or move-out events. (We believe this signal detection task is quite difficult to preprogram in simple analog circuitry, so we will process the required heuristics and thresholding in reconfigurable firmware.)

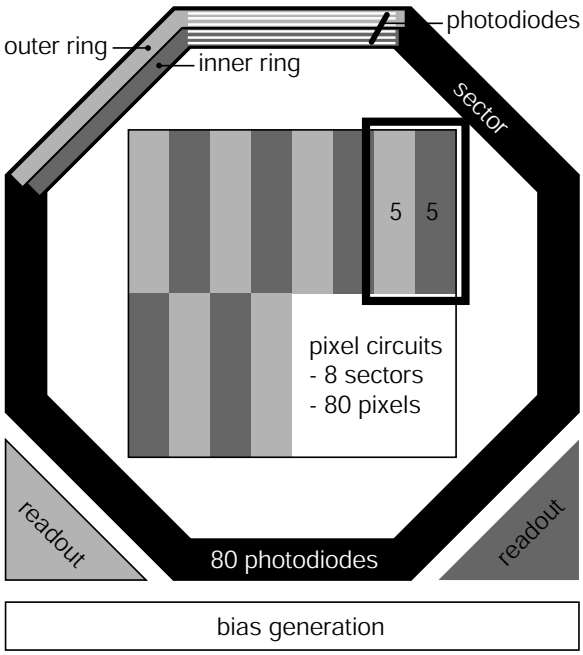

Figure 1. Floor plan of the specialized sensor chip, with octagonal rings of photodiodes. Localized illumination change detection is done in the pixel circuits, located in the center. Gray level represents units dedicated to the same ring. The readout blocks amplify the summed pixel outputs.

2. OPTICS AND FLOORPLANING

Figure 1 shows the octagonally structured floor plan. The light sensitive area is composed of photodiodes with a very large aspect ratio that are aligned along the boundary of the octagonal structure. The chip is sensitive to crossing of this border region. Each ring has 5 photodiodes per sector and 8 sectors per ring, which results in a total of 80 pixels. The pixel processing circuits are placed in the center of the octagon.

Since the sensitive region is only the border of the image, we could use very low cost fisheye optics that have decent optical properties only on the periphery of the image.

Why not build the sensor using only two giant circular photodiodes? We split each ring into sectors because the size of the moving object is expected to be a small fraction of the ring circumference. Each of the pixels independently measures the magnitude of the local illumination change, relative to an adapted level and to the average over the ring. These independent measurements are summed to result in crossing signals—one for each ring. If we had made a single giant circular photodiode, a crossing object would only change the photocurrent by a tiny fraction—which would be difficult to detect—and moreover, we would have no way of distinguishing between localized movement signal and global illumination change.

3. CHIP IMPLEMENTATION

The chip consists mainly of the circuits shown in **Figure 2**. Each pixel has an adaptive photoreceptor and a change detector. The photoreceptors have high gain only for transient signals. The photoreceptor outputs are averaged over the ring. The change detector measures the magnitude of the difference between the average and the individual photoreceptor output. These metrics are summed over the ring and amplified by the readout circuit to result in a scalar change signal from each ring.

From a neuromorphic perspective the global average acts like a giant retinal horizontal cell, while the change detectors are like summing activity from rectifying ON and OFF retinal bipolar cells. In effect we have built a very specialized silicon retina with only two outputs.

3.1 Adaptive Photodetection

An adaptive photoreceptor transduces the photocurrent to voltage[3]. **Figure 2** shows the photoreceptor circuit along with the rest of the circuits involved in the detection of border crossings. The steady state response V_{rcp} of the photoreceptor is logarithmic in the background illumination I_{bg}. The feedback loop is closed through a resistor-like adaptive element Adap that has an exponential voltage dependent conductance. It provides long time scale adaptation to absolute illumination. For small changes i in I_{ph}, the DC response is approximately linear in the contrast i/I_{bg}:

$$v_{rcp}\left(i / I_{\mathrm{bg}}\right) = \frac{U_{\mathrm{T}}}{\kappa} \frac{i}{I_{\mathrm{bg}}}. \qquad (1.1)$$

U_{T} is the thermal voltage, and κ is the back-gate coefficient. Mfb operates in subthreshold to provide the logarithmic illumination encoding. Mn and Mp form a high-gain inverting amplifier that closes the feedback loop through the DC path provided by the adaptive element Adap.

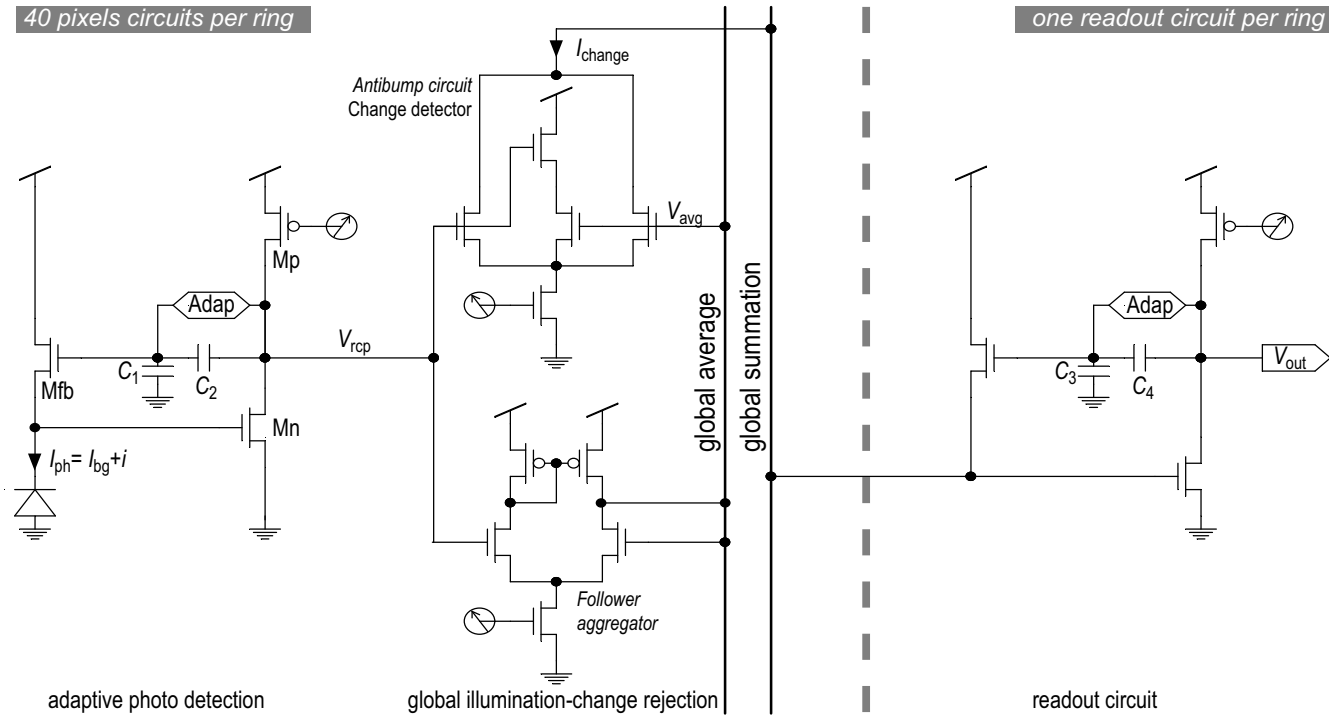

Figure 2. Schematics of keys circuits used on the chip.

The capacitive divider formed by C_1 and C_2, combined with the extremely high impedance of Adap for small voltages, results in higher gain for transient changes of illumination:

$$v_{\text{rcp}}\left(i/I_{\text{bg}}\right)=\left(\frac{C_1+C_2}{C_2}\right)\frac{U_{\text{T}}}{\kappa}\frac{i}{I_{\text{bg}}}\equiv \text{A}_1\frac{U_{\text{T}}}{\kappa}\frac{i}{I_{\text{bg}}}. \quad (1.2)$$

The normal range of contrast change for our application is 30%, which satisfies the small signal approximation. The frequency behavior of the receptor is that of a bandpass filter. Low frequency is blocked by the adaptive-element, while high frequency is limited by the capacitance at the input node and the conductance at the source of M_{fb}, which is proportional to the photocurrent.

3.2 Global Illumination Change Rejection

A global illumination change causes all reflected luminances to increase simultaneously by the same ratio. Equation (1.2) shows that this illumination change causes an identical response in each photoreceptor. Global illumination changes are rejected by measuring the magnitude of the local illumination change using the difference between photoreceptor V_{rcp} and the global average V_{avg}. **Figure 2** (middle) shows a follower aggregator circuit[4], combined with an antibump circuit[5]. The follower-aggregator computes the average photoreceptor output, and the antibump measures the magnitude of the difference with a quadratic metric. The follower-aggregator computes the average photoreceptor voltage by summing all the photoreceptor outputs through a uniform transconductance. The antibump circuit is a combination of a differential pair with a current correlator. When the differential input is small, the correlator sinks most of the bias current. As the differential input increases, more of the bias current flows to one of the outer legs. The antibump transfer function is given in subthreshold by:

$$I_{\text{change}}(\Delta V)=\frac{I_{\text{b}}}{\frac{S/4}{\cosh^2\left(\kappa\Delta V/U_{\text{T}}\right)}+1} \text{ with } \Delta V\equiv V_{\text{rcp}}-V_{\text{avg}} \quad (1.3)$$

I_{b} is the bias current provided for the antibump circuit and S is the transistor W/L ratio between the transistors in the middle and outer legs of the antibump circuit. For small ΔV the response is a constant plus a quadratic in ΔV. The second order approximation to (1.3) is

$$I_{\text{change}}(\Delta V)=\frac{1}{1+\alpha}+\frac{\alpha}{\left(1+\alpha\right)^2}\left(\frac{\kappa\Delta V}{U_{\text{T}}}\right)^2, \quad (1.4)$$

where $\alpha=S/4$. A global illumination change causes the same response in each receptor, which moves the average by the same amount and therefore has no effect on the summed antibump output current. A local change in illumination moves a single receptor output but does not change the global average by much, resulting in a detected change.

3.3 Readout Circuit

Each ring's readout circuit amplifies changes of the summed pixel change signals. An adaptive logarithmic transimpedance amplifier—as used in the photoreceptor—implements this amplification. It compresses the DC level and greatly amplifies the transient changes in the summed contrast signal. The gain of this readout stage takes the same form as for the photoreceptor (1.2). The capacitive divider formed by C_3 and C_4 determines the gain $A_2\equiv(C_3+C_4)/C_4$.

The overall system transient response must take into account that a change at a photoreceptor output affects the average computed by the follower aggregator. This means that the stimulation of a single pixel affects all other pixel outputs—albeit only slightly. The overall response of a system with m sectors of which n are stimulated by the same contrast i/I_{bg} can be written as

$$v_{\text{out}}\left(i/I_{\text{bg}}\right)=\frac{U_{\text{T}}}{\kappa}A_1^2A_2\frac{\alpha}{(1+\alpha)}\left(n-\frac{n^2}{m}\right)\left(\frac{i}{I_{\text{bg}}}\right)^2. \quad (1.5)$$

The maximum response is when half the pixels are stimulated, and the response drops to zero when all are stimulated.

4. RESULTS

In order to have a well-controlled test environment, we used an LCD screen to display images for the chip. We used an 8mm lens on the chip. We explored the responses to simple synthetic stimuli and captured video.

Using a fixed uniform background illumination, we verified the response to stimulus contrast by applying a stimulus patch covering one sector with varying contrast and measuring the voltage of the step response. **Figure 3** shows the output from one ring's readout circuit along with a fit to (1.5). We believe that the shift in measured contrast relative to the fit is due systematic offset in the follower aggregator transconductance amplifier.

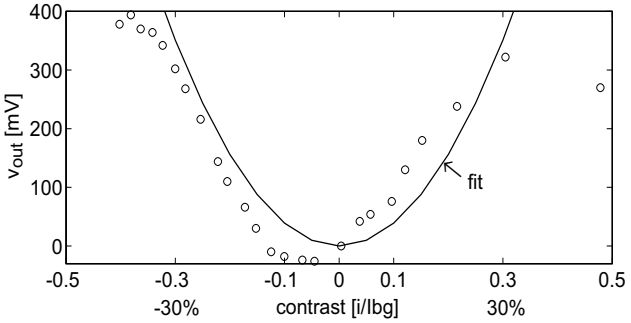

Figure 3. Measured response to steps with positive or negative contrast applied to one of eight sectors.

The effect of varying the number of stimulated sectors is shown in **Figure 4**; it also agrees with (1.5).

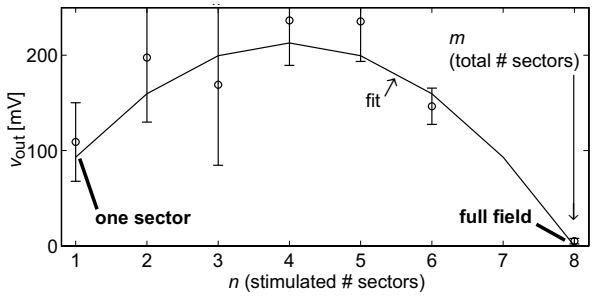

Figure 4. The effect of object size on measured response. We used a stimulus with a contrast of 10%.

We directly verified global illumination change rejection by comparing the response to stimulation of a single sector with the response to global stimulation (**Figure 5**). The response to single sector stimulation is much larger than the response to global stimulation, even though V_{avg} changes much more for the global stimulation; this comparison illustrates the rejection of global illumination changes.

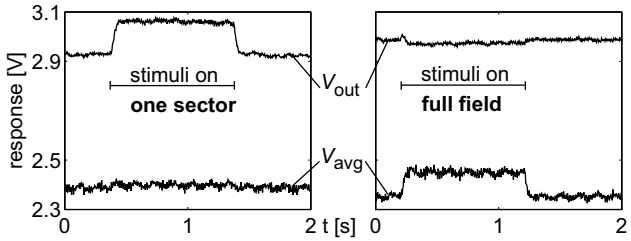

Figure 5. Measured responses comparing one sector with full-field stimulation. Stimulus contrast was 10%.

Figure 6 shows the response to a movie of a person entering the field of view. The viewpoint is from the ceiling through a fisheye lens. The temporal separation of the responses from the outer and inner rings is clearly evident.

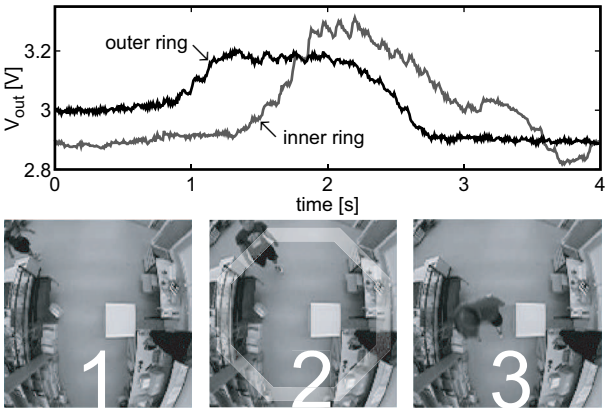

Figure 6. Chip response to a typical move-in event. A person enters at the top of the picture and moves to the middle.

5. DISCUSSION

Measurements have shown that the chip is largely functional. The chip functions over a wide range of illumination and can detect crossing objects with contrasts down a few percent. The output is nearly immune to global illumination changes. However, there are several problems with the existing chip. The input node of the readout circuit is too far from ground, which results in a limited dynamic output range. Interposing a current mirror between antibump current sum and readout amplifier would solve this problem. The on-chip generated bias currents were not optimal. A systematic voltage offset in the follower aggregator transconductance amplifier leads to an asymmetrical contrast response. The photoreceptor 50Hz flicker response necessitates an additional lowpass filter. We have an improved design in fabrication to deal with these problems.

6. REFERENCES

[1] Harrison, R. R. and Koch, C., "A robust analog VLSI Reichardt motion sensor," *Analog Integrated Circuits and Signal Processing*, vol. 24, no. 3, pp. 213-229, Sept.2000.

[2] Liu, S. C., "A neuromorphic aVLSI model of global motion processing in the fly," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 47, no. 12, pp. 1458-1467, Dec.2000.

[3] Delbrück, T. and Mead, C. A. Analog VLSI adaptive logarithmic wide-dynamic-range photoreceptor. 4, 339-342. 1994. London. 1994 IEEE International Symposium on Circuits and Systems. 5-30-1994. http://www.pcmp.caltech.edu/anaprose/tobi/recep/

[4] Mead, C. A., *Analog VLSI and Neural Systems* Reading, MA: Addison-Wesley, 1989.

[5] Delbrück, T. "Bump" circuits for computing similarity and dissimilarity of analog voltages. 475-479. 1991. Proceedings of the International Joint Conference on Neural Networks. 1991. http://www.pcmp.caltech.edu/anaprose/tobi/bump/