# 64x64 Event-Driven Logarithmic Temporal Derivative Silicon Retina

*P. Lichtsteiner and T. Delbruck, Institute of Neuroinformatics, ETH/UNI Zurich, Winterthurerstr. 190, CH-8057 Zurich, Switzerland, tel. +41 1 636 30 38, tobi@ini.phys.ethz.ch*

Real time artificial vision is traditionally limited to the frame rate. In many scenarios most frames contain information redundant both within and across frames. Here we report on the development of an Address-Event Representation (AER) [1] silicon retina chip `TMPDIFF' that generates events corresponding to changes in log intensity. The resulting address-events are output asynchronously on a shared digital bus. This chip responds with high temporal and low spatial resolution, analogous to the biological magnocellular pathway. It has 64x64 pixels, each with 2 outputs (ON and OFF), which are communicated off-chip on a 13-bit digital bus. It is fabricated in a 0.35u 4M 2P process and occupies an area of $(3.3\ \text{mm})^2$. Each pixel has 28 transistors and 3 capacitors and uses a self-clocked switched-capacitor design to limit response FPN. Dynamic operating range is at least 5 decades and minimum scene illumination with *f*/1.4 lens is less than 10 lux.

## Basis of design

The majority of imager work has been aimed at high quality or low cost still imaging. These devices deliver frames of video data at constant rate. Biological vision is based on asynchronous events (spikes) delivered from the retina. The stream of events encodes scene contrasts and contrast changes rather than absolute illumination intensities. The retinal computation is optimized to deliver relevant information and to discard redundancy, enabling efficient and fast vision processing.

TMPDIFF reduces image redundancy by responding only to temporal changes in log intensity. Static scenes produce no output. Image motion produces spike event output that represents *relative changes in image intensity*. This operation in continuous form is represented mathematically by the operation (1) on the pixel illumination *I*.

$$\frac{d}{dt}\log I = \frac{dI/dt}{I} \qquad (1)$$

This temporal derivative is "self-normalized". By this normalization, the derivative encodes changes in contrast rather than absolute illumination differences. Contrasts are determined by differences in reflectance of objects independent of overall scene illumination.

The events generated by TMPDIFF are changes in (1) that exceed a threshold and are ON or OFF type depending on the sign of the change since the last event. Pixel output consists of the stream of ON and OFF events. Idealized operation is illustrated in Fig. 1.

## Pixel Design

The TMPDIFF pixel architecture overcomes previous limitations in [2, 3] by using the excellent matching between capacitors to form a self-clocked switched-capacitor change amplifier (Fig. 2). The front-end is an active unity-gain logarithmic photoreceptor that extends bandwidth to low intensities and can be self-biased by the average photocurrent [4, 5]. This photoreceptor is buffered to a voltage-mode capacitive-feedback amplifier with closed-loop gain $\approx 20$ that is balanced after transmission of each event by the AER handshake. ON and OFF events are detected by the comparators that follow. Mismatch of the event threshold is determined by only 5 transistors and is effectively further reduced by the gain of the amplifier. Much higher contrast resolution than in [2, 3] can be achieved, allowing for operation with realistic scene contrast. The complete pixel with AER handshaking circuits [1] is shown in Fig. 3, which includes one more analog cell that uses a starved inverter to generate an adjustable refractory period. Pixel layout is shown in Fig. 4 and a chip photo is shown in Fig. 5. Chip specifications are listed in Table 1.

## Results

We characterized TMPDIFF's responses and illumination requirements using the rotating fan-like stimulus shown at the upper left of Fig. 6. We designed this stimulus to test usable contrast and speed response. Features at the outer edge move faster, and the 6 wedge borders have varying contrast. The lightest gray segment has black density of 20%, so it has contrast 0.8:1 relative to the white background. TMPDIFF produces usable events over at least 5 decades of operating range and in response to scenes with only 20% contrast.

Although the fill factor (8.1%) is small, the photodiode area ($130\ \text{u}^2$) is large. Photodiode dark current can be inferred from the Fig. 7 DC response characteristics of the log*I* photoreceptor. Measured dark current level is consistent with vendor process specifications and is split about equally between area and perimeter leakage, even for this large area photodiode. For speedier response, we used n-well instead of ndiff for the photodiode, because capacitance of nwell is smaller by factor of ~10.

Events recorded from a public hallway are shown in Fig. 8. TMPDIFF produced about 200 events per 25ms 'frame', representing a data rate of 5-10% of equivalent resolution gray-scale imager.

To explore high-speed imaging capabilities, we used TMPDIFF to record motion of a drosophila fruit fly wing oscillating at the fly's natural wing beat frequency of ~230Hz (Fig. 9(a)). A normal sampled imager would require a frame rate of several times 230Hz to avoid substantial aliasing. Fig. 9(b) shows the ~100 events generated during a time slice of 310us. Fig. 9(c) shows the ON and OFF events during two cycles of the wing, along with a smoothed Y address $\bar{y}$ generated by a first order IIR filter with $\bar{y}_{k+1} = 0.9 * \bar{y}_k + 0.1 * y_k$ that operates directly on the Y addresses $y_k$. This simple measure seems sufficient for reliable extraction of wing beat frequency, phase, and amplitude. This filter has the interesting characteristic that its time constant varies inversely with the event rate.

## Discussion

AER systems are power-efficient and timely in their use of the communication channel between components [1]. TMPDIFF does significant focal-plane redundancy reduction. Simple digital processing on the AER outputs can solve some vision problems cheaply. Practical application targets for TMPIFF include high speed and power efficient vision systems. Event driven vision could be advantageous for high speed (low latency) manufacturing or inspection, and for low power surveillance, building automation, or visual prosthetic front ends.


### *Acknowledgements*

Funded by EC 5th Framework grant CAVIAR (IST–2001-34124), ETH Zürich, and the University of Zürich. Data was collected with help of M. Oster, M. Litzenberger, C. Posch, and S. Fry.



[1] K. Boahen, "Point-to-Point Connectivity Between Neuromorphic Chips using Address Events," *IEEE Trans.on Circuits and Systems II*, vol. 47 pp. 416-433, 2000.

[2] J. Kramer, "An On/Off transient imager with event-driven, asynchronous read-out," in 2002 International Symposium on Circuits and Systems, Phoenix, AZ, USA, 2002, pp. 165-168.

[3] P. Lichtsteiner, T. Delbruck, and J. Kramer, "Improved ON/OFF temporally differentiating address-event imager," in 11th IEEE International Conference on Electronics, Circuits and Systems (ICECS2004), Tel Aviv, Israel, 2004, pp. 211-214.

[4] T. Delbruck, "Analog VLSI phototransduction by continuous-time, adaptive, logarithmic photoreceptor circuits," California Institute of Technology, Pasadena, CA 1994/// 1994.

[5] T. Delbruck and D. Oberhoff, "Self-biasing low-power adaptive photoreceptor," in 2004 International Symposium on Circuits and Systems, 2004, pp. IV-844-847.


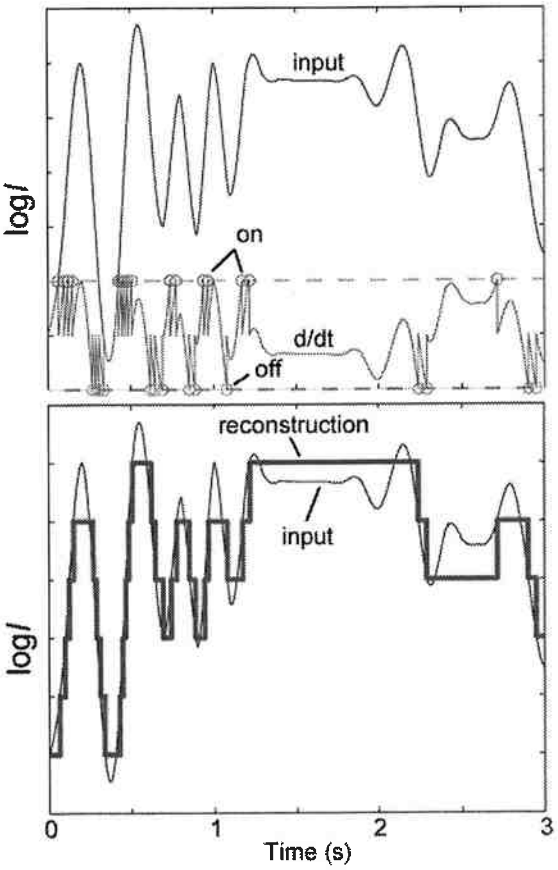


**Fig. 1** Idealized pixel encoding and reconstruction of resampled video data. The ON and OFF events represent significant changes in log*I*. Changes smaller than the thresholds do not generate events but the input signal is still represented internally in the differentiator. Reconstruction of the input signal is based on 30 events.

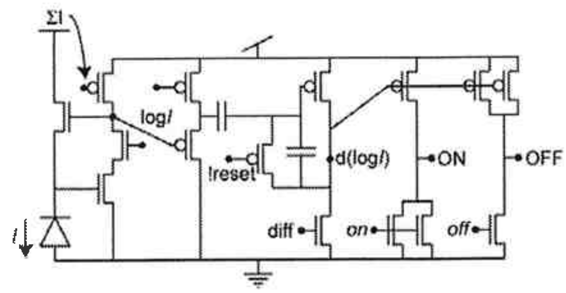


**Fig. 2** Core of pixel circuit. log*I* photoreceptor output is buffered to inverting differentiator, which is reset by each event. An ON or OFF event is triggered when the differentiator output exceeds the threshold set by *on* and *off*. The globally summed photocurrent can be used to adaptively set the photoreceptor bias [5].

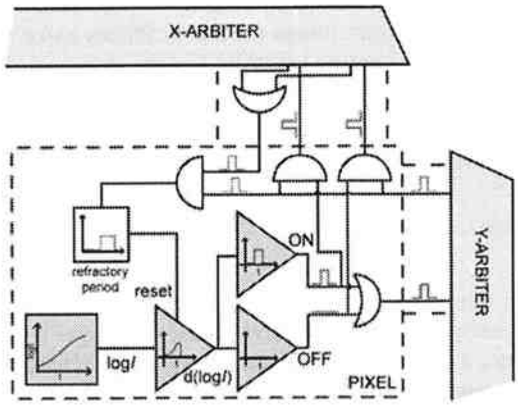


**Fig. 3** Abstracted pixel and handshake circuits. The shaded components (Fig. 2) transduce photocurrent *I* to log*I*, amplify, and differentiate, followed by thresholding. ON and OFF requests to the Y-arbiter are OR'ed together. The requests to the X-arbiter are separate for each channel and are made only if the pixel has been acknowledged by the Y-arbiter. The AND of the Y- and X-acknowledge triggers the reset of the differentiator. The reset is active for an adjustable "refractory period"

to limit maximum spike rate and prevent spike storms from pixels with anomalously low thresholds.

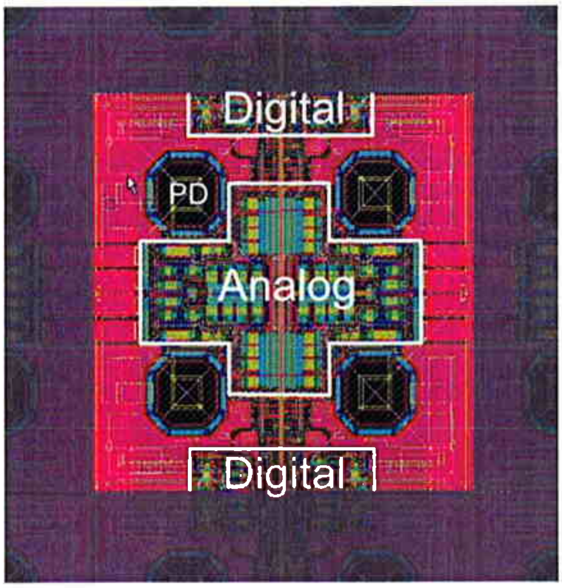


**Fig. 4** Point-symmetric layout of four pixels. Analog circuits and digital circuits are separated as far as possible and bias wires are shared between neighboring pixels.

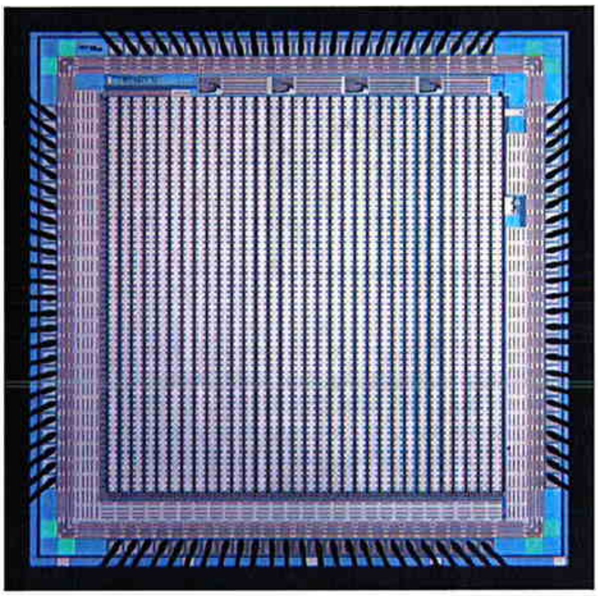

**Fig. 5** Microphotograph of TMPDIFF.

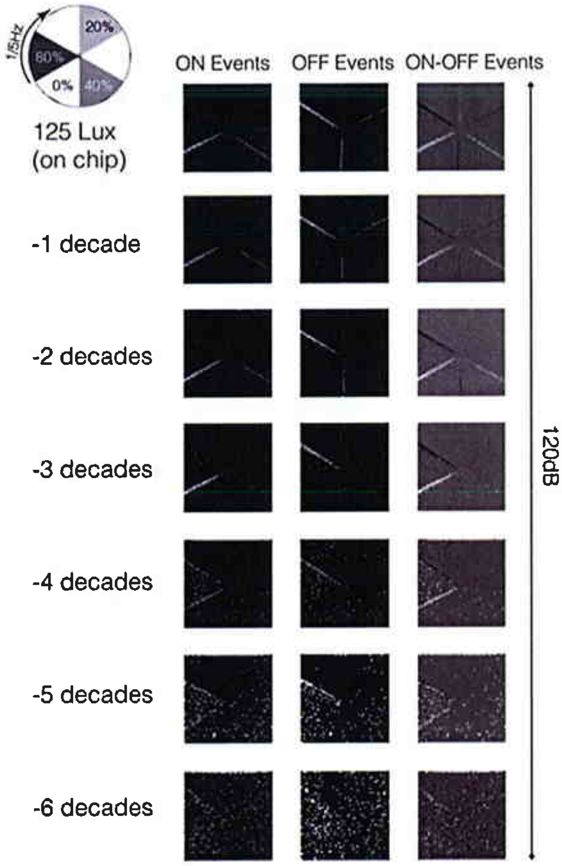


**Fig. 6** Grayscale images of 20 ms integrated AER retina output observing a rotating fan stimulus that was illuminated by a DC-powered halogen 20W 12V lamp. The disk was irradiated with approximately 35W/m$^2$. The luminance of the white part was 1 knit=1000 cd/m$^2$. The lens had aperture f/1.4. Full gray-scale is 6 events. For each row starting from the top, we added another decade of Kodak Wratten neutral-density filtering.

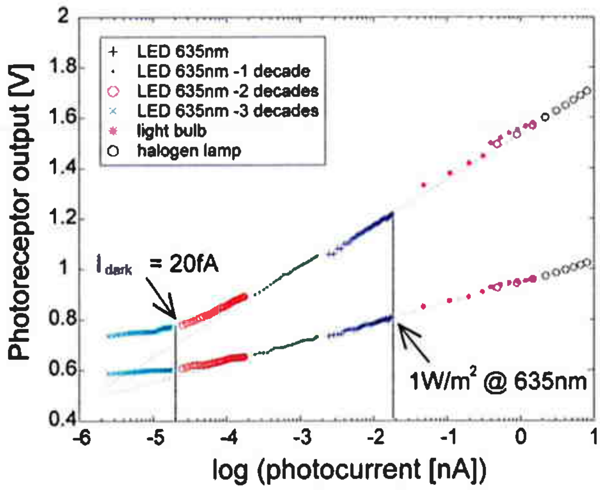


**Fig. 7** DC response of front-end log photoreceptor circuits for two alternative pixel types (the top curve's circuit includes an extra diode-connected feedback transistor). Pixel photocurrent was computed directly from sum of array photocurrents to calibrate different illumination sources. The response is logarithmic over at least 5 decades.

**Fig. 8** Example natural data. The woman on the right is moving to the right and the man on the left is moving to the left; both are darker than the background so the leading edges of each person (the outside edges) tend to produce OFF spikes, while the trailing edges tend to produce ON spikes. The right grayscale image shows the integrated spikes during 25ms; on average in this scene there are 200 spikes per 25ms 'frame'. The left camcorder image was not aligned perfectly with the retina chip image, therefore, the head of a third person, who is walking behind the other two towards the left is only visible in the retina image.

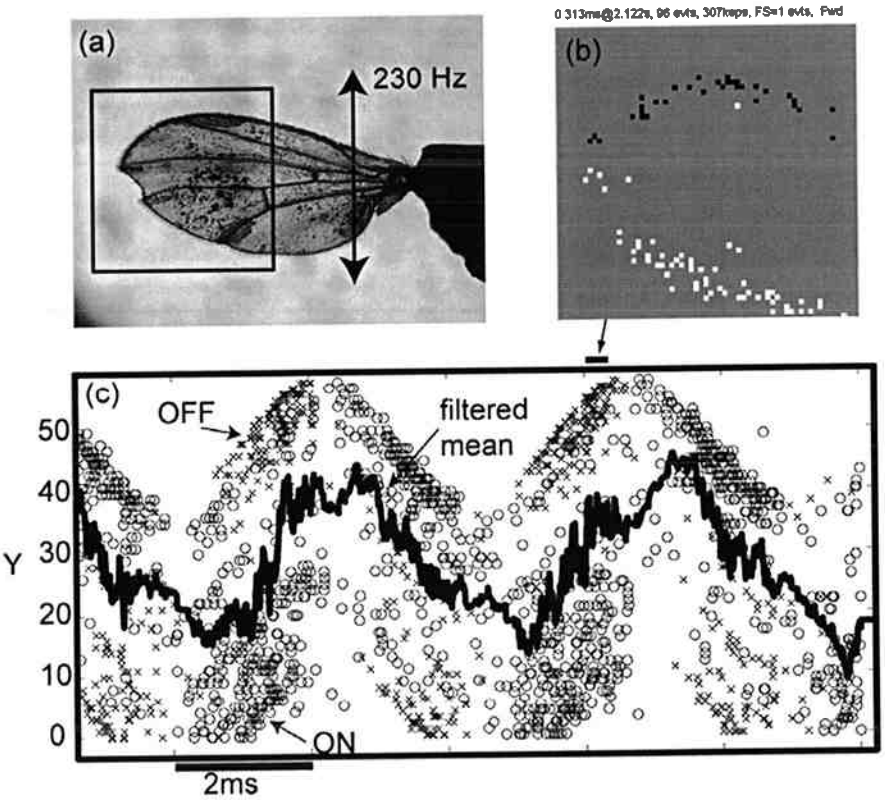


**Fig. 9** Wing beat data. **(a)** shows the wing mounted on vibrating rod as viewed through the microscope. The oscillation frequency was 230 Hz as measured independently using a microphone recording. TMPDIFF was mounted on the microscope camera mount and imaged approximately the rectangular area. **(b)** shows a single 'frame' consisting of the events during a 313 us time window; on average there were about 50 events per frame corresponding to a mean rate of 150k events/second. **(c)** shows plots of the Y event-addresses. Superimposed on the Y addresses is a running average that represents a measure of the vertical location of the wing.

**Table 1** Specifications of TMPDIFF.

| | |
|---|---|
| Pixel size | 40x40 $um^2$ |
| Pixel complexity | 28 transistors (16 analog), 3 caps |
| Array size | 64x64 |
| Chip size | 3.3x3.3 $mm^2$ |
| System supply current | 1-10mA @3.3V |
| Fabrication process | 4M 2P 0.35 um |
| Interface | AER with X+Y+Polarity=6+6+1=13 bits of address, Request&Acknowledge 4-phase handshake |
| Biasing requirements | 12 generated gate biases |
| Fill factor | 8.1% (PD area 130$um^2$) |
| Minimum scene irradiance/illumination using f/1.4 lens | Approx. 1mW/$m^2$ or 1 lux. |
| Photodiode dark current (estimated) | 20fA (~10nA/$cm^2$) |
| Burst event rate | Approx. 4 M event/sec |