27.9 A 128×128 120dB 30mW Asynchronous Vision Sensor that Responds to Relative Intensity Change

P. Lichtsteiner[1], C. Posch[2], T. Delbruck[1]

[1]ETH, Zurich, Switzerland
[2]ARC Seibersdorf Research, Vienna, Austria

The frame-based architectures of most imagers are natural for making movies and pictures, they have significant drawbacks for machine vision. Short-latency vision problems require high frame rate, producing massive readout (e.g., >1GB/s from 352×288 pixels at 10kFrames/s [1]). Reducing the output to a manageable rate by using region-of-interest readout usually requires complex control strategies. Readout and processing of largely redundant data ultimately limit reductions in computational effort and power consumption. In this paper, a vision sensor is presented whose pixels asynchronously respond to events that represent *relative changes in intensity*. It operates largely independent of scene illumination, directly encodes object reflectance, and reduces redundancy while preserving precise timing information. Because output bandwidth is automatically dedicated to dynamic parts of the scene, the sensor is suitable for applications in surveillance and motion analysis. It improves on prior frame-based temporal difference detection imagers (e.g., [2]) by asynchronously responding to temporal contrast rather than absolute illumination, and on prior event-based imagers because they either do not reduce redundancy at all [3], reduce only spatial redundancy [4], have large FPN, slow response, and limited dynamic range [5], or have low contrast sensitivity [6].

The pixels of the sensor use a novel circuit design that combines an active (i.e., fast) continuous-time logarithmic photosensor with a well-matched self-timed switched-capacitor amplifier. Each pixel continuously monitors its photocurrent for changes. It responds with an ON or OFF event that represents a fractional increase or decrease in intensity that exceeds a tunable threshold. Events are communicated asynchronously off-chip on a self-timed bus using the address-event representation (AER) protocol [7].

Figure 27.9.1 shows the pixel circuitry. The photosensor (D, M_{fb}, M_n, M_{cas}, and M_{pr}) has a fixed contrast gain of $nU_T \approx$ 35mV/e. Its bandwidth is proportional to photocurrent and is larger than that of a passive logarithmic photosensor by a factor proportional to the loop gain, at the cost of increased power consumption and shot noise. M_{pr} can be self-biased by using a low-pass-filtered multiple of the average photocurrent (ΣI) [8]; this minimizes the power consumption while maintaining a constant resonance. This photoreceptor is buffered (via M_{b1}, M_{b2}) and capacitively coupled to a capacitive-feedback amplifier (C_1, C_2, M_{dn}, M_{dp}) with closed-loop gain $A \approx 20$. This amplifier is balanced by closing the switch M_r after transmission of each event by the AER handshake. ON and OFF events are detected by two comparators (M_{ONn}, M_{ONp}, M_{OFFn}, M_{OFFp}). The array can globally be held in reset by the switch M_{gr}. The remaining transistors implement the 4-phase AE handshaking with the peripheral AE arbiters. The row and column ON and OFF request signals (RR, CRON, CROFF) are generated individually, while the acknowledge signals (RA, CA) are shared. The combination of RA and CA resets the pixel using a starved inverter (C_3, M_{CA}, M_{RA}, M_{ref}) that balances the amplifier for an adjustable 'refractory' period which limits the maximum event rate. The key to the low FPN is that the mismatch of the event threshold referred to the input (contrast) is reduced by the well-matched capacitor loop gain $A = C_1/C_2$ of the amplifier, e.g., a 20mV comparator mismatch is reduced to 1mV at the photosensor, corresponding to a contrast of ≈ 3.5%. Charge injection by the balance switch M_r, which is nominally identical across pixels, is minimized by using a low overhead switch drive at rGND. Junction leakage in M_r causes a low rate of ON events; these can be eliminated by slightly turning on M_{gr}. Figure 27.9.2 shows the arrangement of a pixel in an array that is surrounded by all the functional blocks of the AER communication structure [7].

This pixel has been integrated in a 128×128 array built in a standard 0.35µm CMOS process. The unique characteristics of this vision sensor—contrast coding under wide illumination variation, short-latency response to fast stimuli, and low output data rate—are illustrated in Fig. 27.9.3 to 27.9.5. Figure 27.9.3 shows how only contrast is encoded when a density-step target is moved through a field of view in which illumination varies by a factor of 135. Figure 27.9.4 demonstrates the high-speed capability of the image sensor in response to a 'contrast wedge' rotating at 1000rpm. The left images show the stimulus and the ON and OFF events and the right image shows the event time coded by color over a 5ms slice. The leading edge of each contrast step produces the youngest events. Figure 27.9.5 demonstrates how a low data rate can still maintain a good representation of a moving person.

Fig. 27.9.6 is a table of specifications that compares the vision sensor with other devices. The sensor pixels operate independently over a scene illumination range (f/1.4 lens) of >100klux down to under 1lux, limited at the low end by the dark current in the standard CMOS process used. As the ON and OFF event thresholds are decreased, the background firing increases. At a background firing rate of <1000events/s, more than 90% of the pixels respond to a 10% contrast.

Acknowledgements
This work is funded by EC grant CAVIAR (IST–2001-34124), ETH Zürich, the University of Zürich, and ARCS. We thank K. Boahen, S. Mitra, and G. Indiveri for AER circuit layout and S.C. Liu for editing.

References
[1] S. Kleinfelder, S. Lim, X. Q. Liu, and A. El Gamal, "A 10000 frames/s CMOS Digital Pixel Sensor," *IEEE J. of Solid-State Circuits*, vol. 36, no. 12, pp. 2049-2059, Dec., 2001.
[2] U. Mallik, M. Clapp, E. Choi, G. Cauwenberghs, and R. Etienne-Cummings, "Temporal Change Threshold Detection Imager," *ISSCC Dig. Tech. Papers*, pp. 362-363, Feb., 2005.
[3] E. Culurciello and R. Etienne-Cummings, "Second Generation of High Dynamic Range, Arbitrated Digital Imager," Int. Symp. on Circuits and Systems, vol. 4, pp. 828-831, May, 2004.
[4] P. F. Ruedi, et al., "A 128×128 Pixel 120-dB Dynamic-Range Vision-Sensor Chip for Image Contrast and Orientation Extraction," *IEEE J. Solid-State Circuits*, vol. 38, no. 12, pp. 2325-2333, Dec., 2003.
[5] K. A. Zaghloul and K. Boahen, "Optic Nerve Signals in a Neuromorphic Chip II: Testing and Results," *IEEE Trans. Biomedical Engineering*, vol. 51, no. 4, pp. 667-675, Apr., 2004.
[6] P. Lichtsteiner, T. Delbruck, and J. Kramer, "Improved ON/OFF Temporally Differentiating Address-Event Imager," *IEEE Int. Conf. on Electronics, Circuits, and Systems*, pp. 211-214, Dec., 2004.
[7] K. A. Boahen, "A Burst-Mode Word-Serial Address-Event Link-I Transmitter Design," *IEEE Trans. Circuits and Systems I*, vol. 51, no. 7, pp. 1269-1280, July, 2004.
[8] T. Delbruck and D. Oberhoff, "Self-Biasing Low-Power Adaptive Photoreceptor," Int. Symp. on Circuits and Systems, pp. 844-847, May, 2004.

 1-4244-0079-1/06/$20.00 ©2006 IEEE

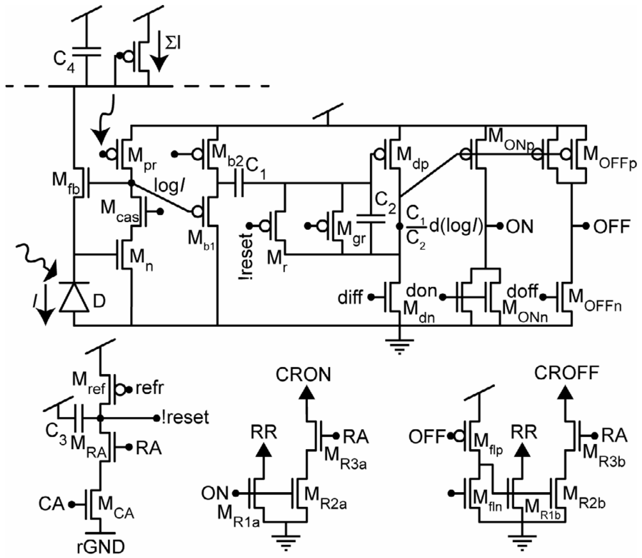

Figure 27.9.1: Pixel schematic.

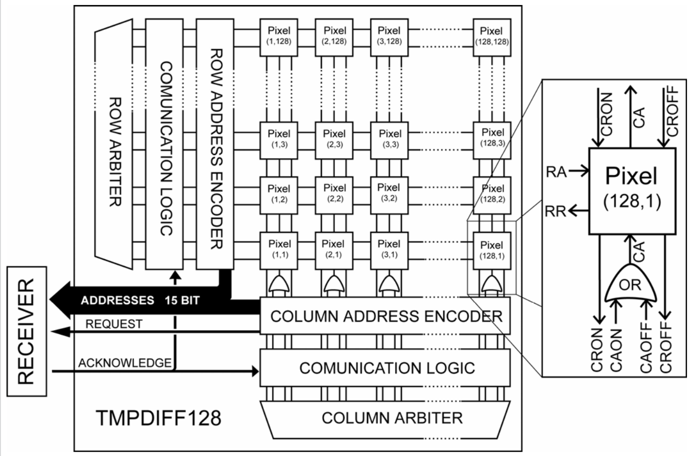

Figure 27.9.2: Chip architecture.

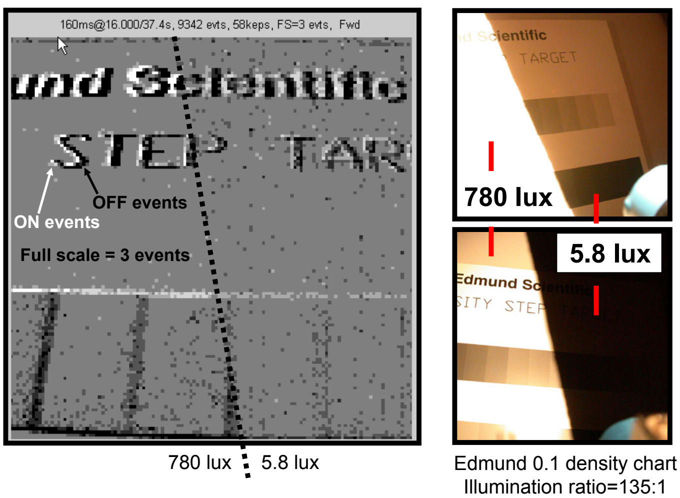

Figure 27.9.3: Contrast sensitivity under wide illumination.

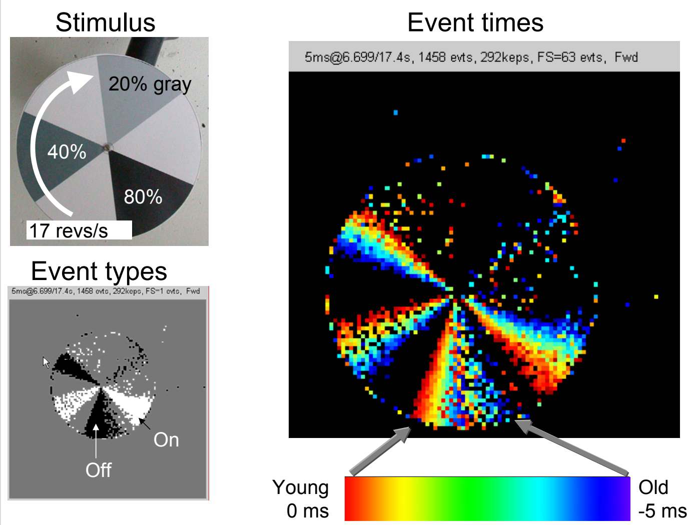

Figure 27.9.4: Time resolution of events.

Figure 27.9.5: Dynamic scene example.

	TEMPDIFF128	Rüedi et al. [4]	Zaghloul, Boahen [5]	Kleinfelder et al. [1]	Mallik et al. [2]
Functionality	**Asynchronous temporal contrast**	Frame-based spatial contrast and gradient direction, ordered output	Asynchronous spatial and temporal contrast	In-pixel ADC APS imager	Temporal change detection APS imager
Pixel size um (lambda) Fill factor (%)	**40x40 (200x200) 8.1% (PD area 130μm²)**	69x69 (276x276) 9%	34x40 (170x200) 14%	9.4x9.4 (104x104) 15%	25x25 (100x100) 17%
Fabrication process	**4M 2P 0.35μm**	3M 2P 0.5μm	4M 2P 0.35μm	5M 2P 0.18μm	3M 2P 0.5μm
Pixel complexity	**26 transistors (14 analog), 3 capacitors**	> 50 transistors 1 capacitor	38 transistors	37 transistors	6 transistors, NMOS 2 capacitors
Array size	**128x128**	128x128	96x60	352x288	90x90
Die size mm^2	**6x6.3**	~ 10x10	3.5x3.5	5x5	3x3
Interface	**15-bit word-parallel AER**	8-bit bus, 16 x 24-bit FIFO,Non-arbitrated with collision detection	8-bit word-serial AER	64-bit (8-pixel) bus, 167 MHz	Serial, with event FIFO
Power consumption	**30mW @ 3.3V**	300mW @ 3.3V	62.7mW @ 3.3V	50mW @ 3.3V (10kfps)	30mW @ 5V (50 fps)
Operating range	**120dB 1 lux to > 100 klux with f/1.4 lens**	120dB	45dB	~45dB	51dB
Photodiode dark current	**20fA (~10nA/cm²) Nwell photodiode**	300fA	? Phototransistor	10nA/cm²	?
Response latency Frames/sec	**< 100μs @ 700mW/m² ~2M events/sec**	< 2ms 60 to 500 fps	? ~10M events/sec	100us 10k fps	< 5ms? 200 fps?

Figure 27.9.6: Summary and comparison of chip characteristics.

Continued on Page

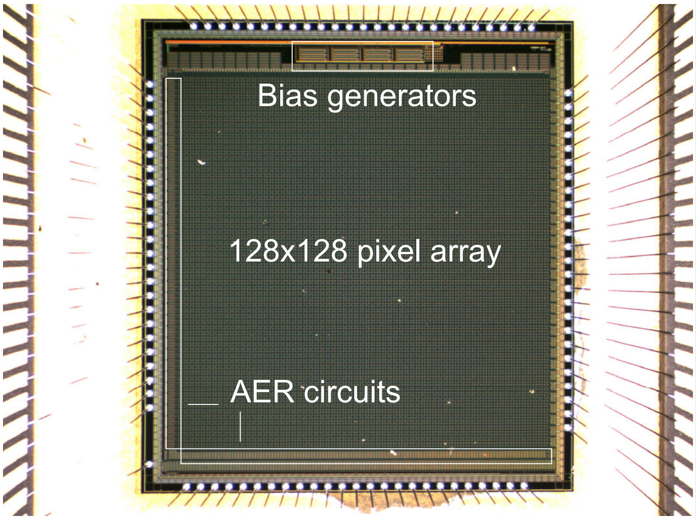

Figure 27.9.7: Chip micrograph.

