A 100dB Dynamic Range Event-Driven Spatial Contrast Sensor with 100μs Response Time and Time-to-First-Spike Mode

J. A. Leñero-Bardallo, T. Serrano-Gotarredona, and B. Linares-Barranco

Instituto de Microelectrónica de Sevilla (IMSE-CNM-CSIC)

Ed. C/ Américo Vespucio s/n, 41092 Sevilla, Spain. Email: bernabe@imse-cnm.csic.es

Abstract– **Bio-inspired vision sensors have some inherent advantages over conventional sequential-still-image sensors. Some of them are high speed, low latency and reduced bandwidth and power consumption. In this paper, we present a new spatial contrast retina with signed output. Its output is zero if there is no contrast. The new sensor includes an optional Time-to-First-Spike mode (TFS) that combines the advantages of AER vision systems and frame-based ones. In TFS mode, times between consecutive frames can be adjusted dynamically by transmitting only relevant information. Both operation modes are ambient-light-independent, to first order. A 32x32 pixel prototype has been fabricated in 0.35um CMOS. Experimental results are provided.**

I. INTRODUCTION

AER (Address Event Representation) is a spike based signal representation hardware technique for communicating spikes between layers of neurons in different chips. A variety of AER visual sensors can be found in the literature, such as simple luminance to frequency transformation sensors, or sensors that include focal plane simple preprocessing such as temporal [1], spatio-temporal contrast [2], or combined luminance/contrast [3]. Spike based visual sensors can code their output signals using rate coding or TFS coding [4]. When using rate coding, each pixel is autonomous and continuously generates spikes at a frequency proportional to the signal to transmit. There are no video frames, so that sensing and processing is continuous and frame-free. When using TFS coding, a global system-wide reset is provided and each pixel encodes its signal by the time between this reset and the time of the only spike it generates (if it detects any contrast). Sensing and processing is frame-constraint. However, TFS is a highly compressed coding scheme (each pixel generates at the most one spike per frame) and frame time can be dynamically adjusted to an optimum minimum by subsequent processing stages. For example, a simple counter can be included at the chip output so that a fixed number of "contrast events" are transmitted, after which the sensor is reset. This TFS sensing mode is based on neuroscience studies that demonstrate single-spike-per-neuron recognition [5]. Spatial contrast AER retina sensors compute contrast on the focal plane reducing significantly data flow, while relevant information for shape and object recognition is preserved. In this paper we present a new spatial contrast retina chip. For each pixel, it computes a spatial contrast current $I_{cont}(x, y)$ using the pixel's locally sensed light intensity $I_{ph}(x, y)$ and a spatially weighted average $I_{avg}(x, y)$ sensed with a diffusive network. The contrast computation follows the equation

$$I_{cont}(x, y) = I_{ref}\left(\frac{I_{ph}(x, y)}{I_{avg}(x, y)} - 1\right) \qquad (1)$$

The design is based on Boahen's biharmonic contrast computation circuit [2]. The new sensor improves on mismatch by in-pixel calibration, on ambient light independence, on controllability, and reduces bandwidth communication. Chip spatial contrast output can be either a continuous frame-free asynchronous event flow, or a frame-reset burst of single-event-per-pixel with TFS coding.

II. CIRCUIT ARCHITECTURE

Fig. 1 shows the schematics of the pixel circuitry. Fig. 1(a) provides an overall block diagram. The pixel contains three main parts: (1) the photo sensing and contrast computation block, including calibration, which provides an ambient light independent signed-contrast current I_{cont}; (2) an integrate-and-fire block, which includes refractory circuitry, thresholding, and TFS mode; (3) and the pixel AER communication circuitry that sends out events to a periphery arbiter. Fig. 1(b) shows the new pixel contrast computation circuit. Contrast computation is based on solving a biharmonic computer vision equation. The operation of the circuit is mathematically intrigued and cannot be explained here, but is detailed elsewhere [6]. The modified pixel includes a current biasing scheme for tuning gate voltages of transistors M_c and M_h. This way, they tend to follow voltage excursions at nodes 'C' and 'H'. This biasing scheme shows three improvements over its original design: (a) biasing will adapt to ambient light conditions, (b) it attenuates mismatch, (c) and is much less critical to tune. The contrast computation block is connected to a calibration block. This block uses digitally-controlled-length transistors together with peripheral translinear circuits for global adaptation [7]. This way, the retina has only to be calibrated once and calibration degrades only slightly with ambient light. Fig. 1(c) shows the integrate-and-fire block. Input contrast current I_{cont} is integrated on capacitor C_{int}. Two comparators detect whether the capacitor voltage V_{cap} reaches an upper (V_{high}) or lower (V_{low}) threshold, triggering the generation of a positive (*pulse+*) or negative (*pulse-*) event, respectively. After event generation, capacitor C_{int} is reset to the central voltage V_{ref}. This is done by the reset circuit shown in Fig. 1(d). This reset mechanism includes a refractory timing circuit that inhibits the pixel from generating subsequent events before refractory capacitor C_{rfr} has been discharged by the DC current source MOS controlled by V_{rfr}. The reset circuit also includes a global TFS (Time-to-First-Spike) mode reset signal, which resets all pixel
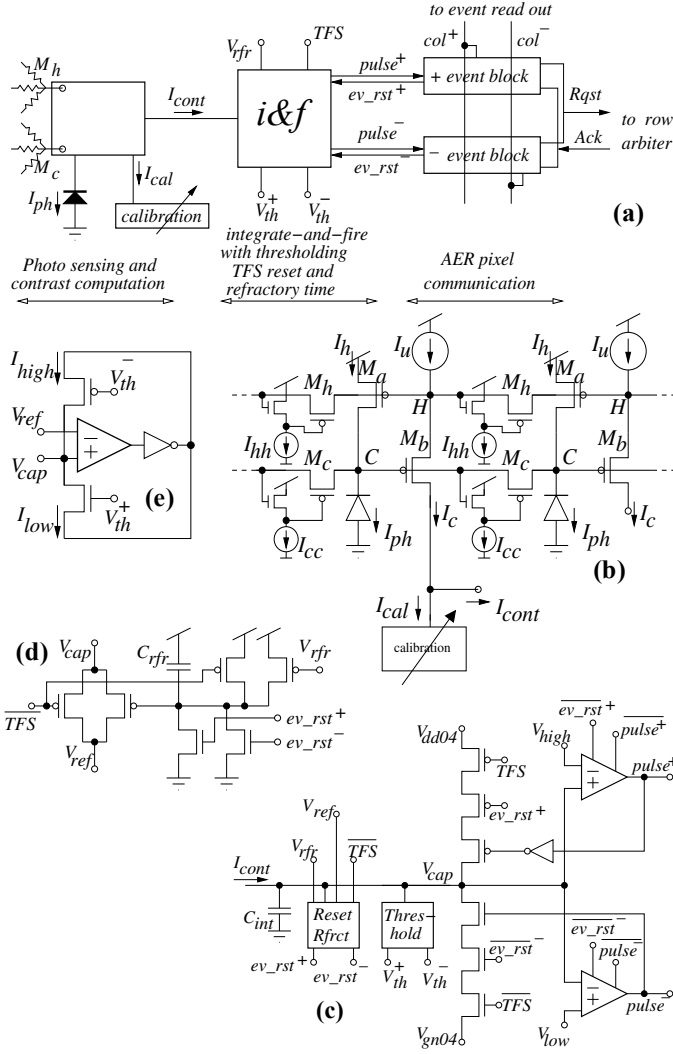
Fig. 1: Pixel schematics diagram. (a) Compact block diagram. (b) Detail of photo sensing and contrast computation circuit. (c) Detail of signed integrate-and-fire circuit. (d) Detail of reset and refractory circuit. (e) Detail of thresholding circuit.

capacitors C_{int} simultaneously by activating signal $\overline{TFS}$. Fig. 1(e) shows the minimum contrast thresholding circuit. With this block only pixels whose output current is higher than a preset threshold will generate events. A comparator detects whether capacitor voltage is above or below V_{ref} and turns on either a positive (I_{low}) or negative (I_{high}) threshold current, which I_{cont} needs to exceed for producing an event.

III. EXPERIMENTAL RESULTS

A 32 x 32 pixel test prototype AER signed spatial contrast retina chip has been designed and fabricated in a $0.35\mu m$ CMOS process. Table 1 summarizes the sensor specifications and its main features. Fig. 2 shows a micro photograph of the die, of size 3.7 x 3.5 mm^2. Pixel area is $81.5 \times 76.5\mu m^2$.

A. Calibration

The low-current contrast computation circuit in Fig. 1(b) is highly sensitive to mismatch, both in DC offset and contrast gain. We introduce a calibration circuit for compensating DC offset, while contrast gain mismatch is reduced by using larger area for transistors M_a (since they dominate this mismatch

Table 1: Sensor Specifications

| technology	CMOS 0.35μm 4M 2P
power supply	3.3V
chip size	3.7 x 3.5 mm²
array size	32 x 32
pixel size	81.5 x 76.5 μm²
fill factor	2.0%
photodiode quantum efficiency	0.34 @ 450nm
pixel complexity	131 transistors + 2 caps
current consumption	65μA @ 10keps
dynamic range	1-100k lux
post-calibration FPN	1.6% over 5 decades of ambient light
contrast sensitivity	4400 Hz/WC
temporal latency	0.1ms @ 50k-lux
peak output event rate	22 Meps
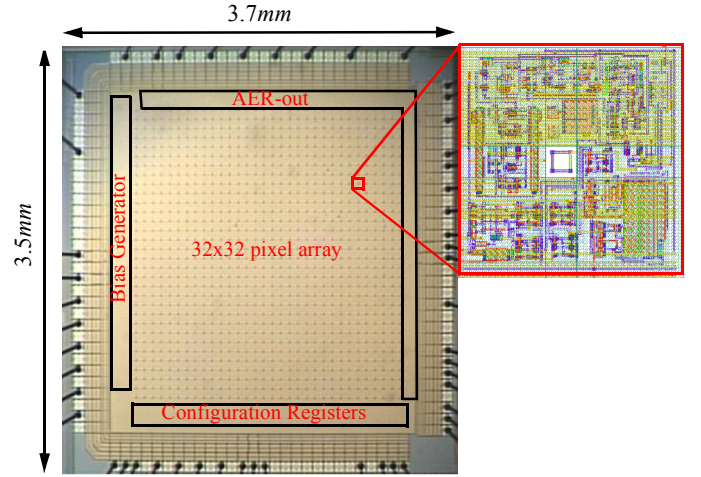
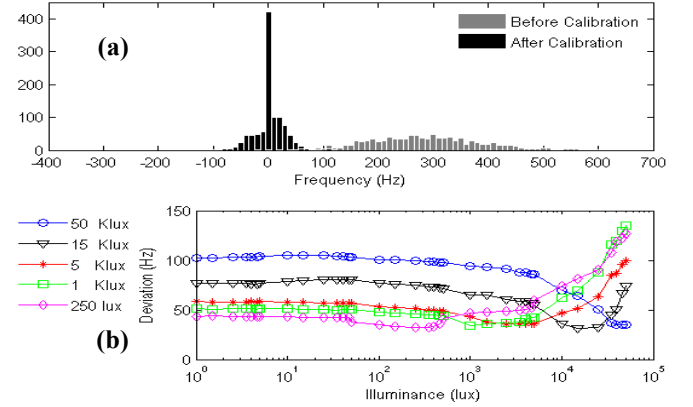Fig. 2: Die microphotograph and zoomout of pixel layout.

Fig. 3: (a) Histograms of pixel frequencies before and after calibration. (b) Measured standard deviation when changing light after calibration.

component). The retina is uniformly illuminated and each pixel I_{cal} current is adjusted to equalize all pixel frequencies. Fig. 3(a) shows the pixels histogram before calibration (grey color) and after calibration (black color). Residual inter-pixel standard deviation is $26Hz$, for a maximum contrast frequency excursion of $\pm4400Hz$ ($\sigma = 0.3\%$). The calibration process is only slightly sensitive to illumination. Fig. 3(b) shows how calibration degrades with illumination. The sensor was calibrated at different luminance values. The optimum situation corresponds to calibrating at 1k-lux which yields a worst case error over 5-decades of $\sigma_{max}= 1.6\%$.

Fig. 5: Time line of the Global Reset and the Request signal.

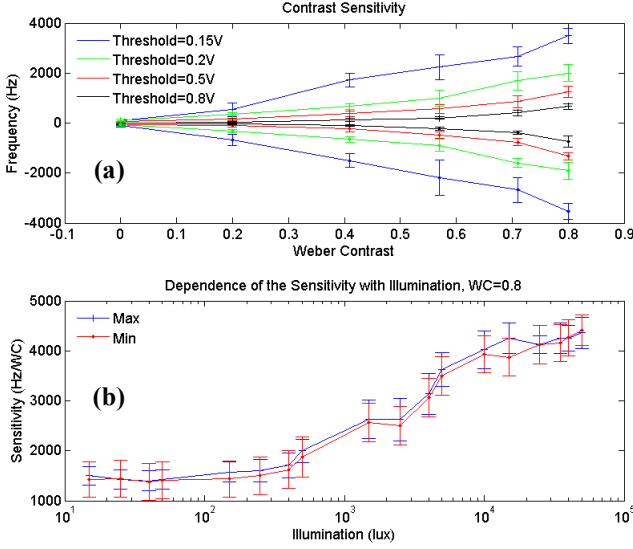Fig. 6: Effect of illumination on T_{frame} and T_{first}

Fig. 4: Contrast sensitivity measurements. A stimulus step was applied and max and min frequencies were recorded. (a) Top panel shows max and min frequencies for different stimulus step contrasts and different threshold values. (b) Bottom panel shows how the maximum and minimum frequencies depends on illumination (*WC*=0.8).

B. Contrast Sensitivity

Contrast sensitivity is the output event rate for a given input contrast stimulus. A gray level step stimulus of different contrast values was swept over the array. This process was repeated for different bias values for V_{high} and V_{low}, with V_{ref} = 1.65V. The results are shown in Fig. 4(a). The measured maximum contrast sensitivity was 4400*Hz/WC* (*Hz* per Weber Contrast[1]) for *Threshold* = V_{high} - V_{ref} = V_{ref} - V_{low} = 0.15*V*. Error bars indicate inter-pixel variability.

To show the sensitivity dependence with illumination, the maximum output frequency for a Weber Contrast of *WC* = 0.8 was measured (for both signs of contrast) with different illumination levels. As shown in Fig. 4(b), sensitivity remains almost constant over the first two decades, and approximately doubles over the second two decades.

C. Latency Characterization

Latency was characterized by stimulating a LED with a step signal to turn it ON, focusing it over a central region of the sensor array, and recording the time delay between the step signal and the first event *Rqst* coming out of the chip from that region. The measurements were repeated for different light intensities from about 50k-lux down to 2 lux. Latency Δt changes approximately linearly in log-log space from about 10*ms* down to about 0.1*ms* when illumination varies over almost 5 decades. A first order fit results in $\Delta t = 20ms / \sqrt{lx}$, where 'lx' is ambient light in lux.

D. TFS Output Mode

The integrate-and-fire circuit of the retina pixel can be configured to operate in TFS mode. In this mode, the refractory period allows each pixel to fire at the most one single event.
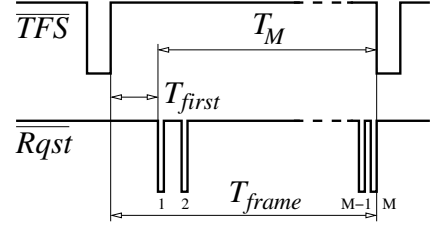
---

1. $WC = (I_1-I_2)/(I_1+I_2)$ between two adjacent pixels or regions.
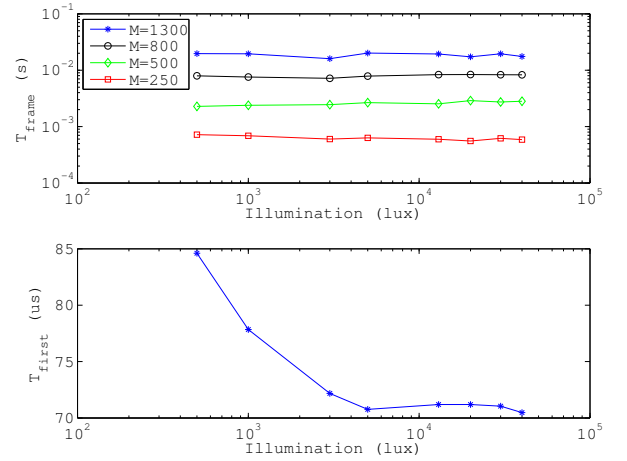
Then a periodic reset pulse has to be provided for global signal $\overline{TFS}$. This can be done in several ways. One option is to reset at a fixed preset frequency. Another option is by counting the output events. Since output events are coming out in decreasing order of pixel contrast, high contrast pixels (either positive or negative) come out first. These are the pixels carrying more relevant information. Consequently, one could add a simple counter at the *Rqst* line and have it generating a reset pulse for $\overline{TFS}$ after each *M* events. This way, a dynamic "frame time" T_{frame} would be produced which self adjusts to the contrast level of the scene, independent of ambient light. High contrast scenes would self-tune to faster frame rates, while low contrast scenes would self-tune to slower frame rates for the same amount of contrast information.

The TFS output mode is also insensitive to illumination, in first order. Several snapshots of the same object were taken under different illumination conditions. As shown in Fig. 5, T_{frame} is the sum of T_{first} (the time the retina needs to generate the first spike after the reset) and T_M (the time between the first and *M-th* spike). Fig. 6 top shows the value of T_{frame} for different values of *M* and illumination levels. T_{frame} is almost independent on illumination and is approximately constant for a given *M*. Fig. 6 bottom shows how T_{first} changes with illumination. For low light (5 lux) T_{first} degrades to 1.3ms.

E. Power Consumption

Chip power consumption depends on static bias conditions and output event rate. However, it is dominated by the last one. For normal operation regimes (between 100keps and 1Meps) current consumption varies between 660μ*W* - 6.6*mW* for nominal bias settings. The dependence is approximately linear in log-log space, resulting in $I_{vdd} \approx I_o \times eps^m$, where I_{vdd} is
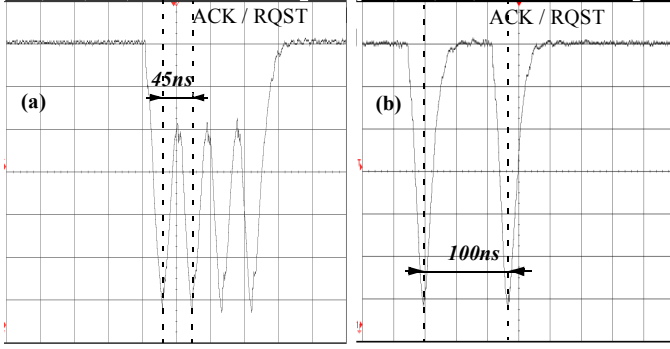
Fig. 7: Delay between events when short cutting the signals _Ack_ and _Rqst_. (a) Minimum delay between consecutive events generated by the pixels of one row (b) Minimum delay between consecutive events generated by two pixels of different rows.

supply current, $I_o = 50nA$, $m = 0.77$, and eps is output "events per second". Pixel output frequency range is directly controlled by bias current I_u (see Fig. 1(a)).

F. Maximum Output Event Rate

Under normal stimuli the sensor provided output event rates up to slightly higher than 2*Meps*. In practical situations, it is difficult to stimulate the sensor for higher output rates. The maximum output event rate was determined by shorting *Ack* and the *Rqst* signals of the chip and measuring the delay between consecutive events. The chip used row-parallel event read-out, where simultaneous events in the same row are sent off chip in a fast burst mode, while inter-row events have a slower rate [8]. Fig. 7(a) shows an event burst generated by the pixels in the same row. The delay between consecutive events is 45*ns*. This indicates that the sensor can provide peak output rates of up to 22*Meps*. Fig. 7(b) shows two events originated at different rows. In this case, the minimum delay between the consecutive events was about 100*ns*, which corresponds to a rate of 10*Meps*. Maximum sustained rate could be somewhere in between, depending on the statistics of the data.

Fig. 8 shows some captured images when observing a ball pen. There are some snapshots of the different parts of the object. High contrast negative events are represented in white and high contrast positive events in black. Central grey color represents zero contrast.

IV. CONCLUSIONS

A new AER signed spatial contrast retina has been presented. The new retina has signed output. A calibration scheme is included to partially compensate for pixel mismatch. An optional TFS coding scheme is also available. Very low latency (100μs), FPN (1.6%), power consumption (65μA @ 10keps) and bandwidth consumption are its main advantages. Experimental results from a test prototype of 32 x 32 pixels, fabricated in a 0.35μm CMOS technology, are provided. Further details will be provided elsewhere [9].
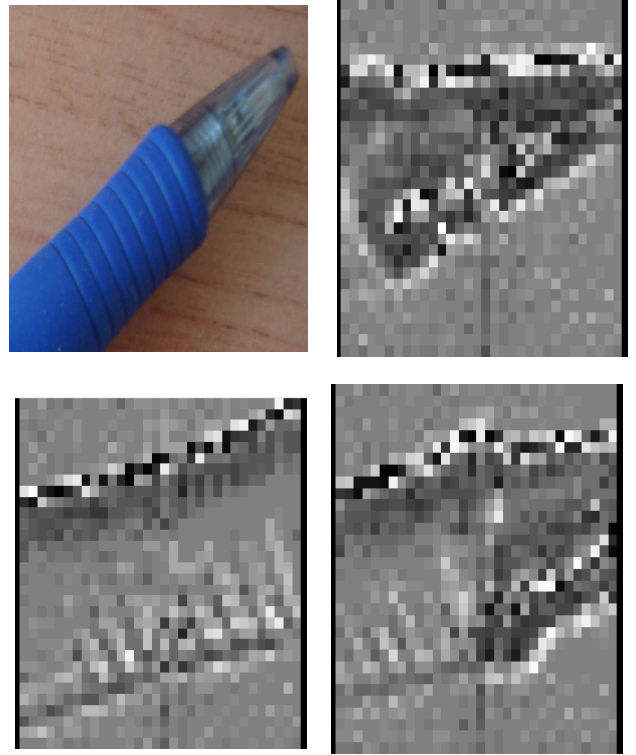
V. ACKNOWLEDGEMENTS

Fig. 8: Snapshot of a ball pen taken with a conventional camera and different detailed captures of the object taken with the spatial contrast retina.

VI. REFERENCES

[1] P. Lichsteiner, C. Posh, and T. Delbrück, "A 128 x 128 120dB 15μs Latency Asynchronous Temporal Contrast Vision Sensor", *IEEE J. Solid State Circ.* vol. 43, No. 2 pp. 566-576, Feb. 2008.

[2] K. Boahen and A. Andreou, "A contrast-sensitive retina with reciprocal synapses", in J. E. Moody (Ed.), *Advances in neural information processing,* vol. 4, pp. 764-772, San Mateo CA, 1992. Morgan Kaufman.

[3] C. Posch, et al. "A QVGA 143dB dynamic range asynchronous address-event PWM dynamic image sensor with lossless pixel-level video-compression," *IEEE Int. Solid-State Circ. Conf.* (ISSCC) 2010.

[4] S. Chen, et al., "Arbitrated Time-To-First Spike CMOS Image Sensor," *IEEE Trans. VLSI Systems,* vol. 15, No. 3, pp. 346-357, March 2007.

[5] S. Thorpe, D.Fize, C. Marlot, "Speed of processing in the human visual system", *Nature* 381:520-2, 1996.

[6] K. Boahen, "Retinomorphic Vision System," *5th Int. Conf. on Microelectronics for Neural Networks and Fuzzy Systems* (MicroNeuro '96), pp. 2-14, 1996.

[7] J. A. Leñero-Bardallo, et al., "A Calibration Technique for Very Low Current and Compact Tunable Neuromorphic Cells. Application to 5-bit 20nA DACs", *IEEE Trans. on Circuit and Systems, Part-II: Brief Papers,* vol. 55, No. 6, pp. 522-526, June 2008.

[8] K. Boahen, "Point-to-Point Connectivity between Neuromorphic Chips using Address Events," *IEEE Trans. Circ. Syst., Part-I*, vol. 53, No. 12, pp. 2548-2566, Dec. 2006.

[9] J. A. Leñero-Bardallo, et al., "A 5-decade Dynamic Range Ambient-Light-Independent Calibrated Signed-Spatial-Contrast AER Retina with 0.1ms Latency and Optional Time to First Spike Mode," *IEEE Trans. Circ. Syst., Part-I*, accepted for publication.

137