A 1.2mW CMOS Temporal-Difference Image Sensor for Sensor Networks

Zhengming Fu, Eugenio Culurciello
Electrical Engineering Department, Yale University, CT USA

***Abstract*—In this paper we report a 64×64 active pixel image sensor (APS), *Mototrigger*, which targets low-power image sensing for sensor networks. The image sensor computes the temporal-difference between two continuous light-intensity frames and communicates them in address-event format. The image sensor is capable of computing 30 temporal-difference frame/s and consumes 1.2mW with a 3V power supply. Light-intensity images are also available as an output at a speed of 60 frame/s. The pixel occupies an area of 29×28μm^2 with a fill factor of 23%. The die size is 3mm by 3mm, including pads and test structures. The image sensor is fabricated in a 0.5μm bulk CMOS process.**

I. Introduction

Over the last decade, much research on CMOS image sensor has focused on increasing the imager resolution and speed [1], [2], [**?**]. A high-quality image delivers high fidelity visual information, but also requires large resources to compute and store them. In particular, when a video stream is communicated through a wireless channel, to understand the content of a scene, high-resolution images become cumbersome and inefficient to be handled [**?**]. Down-sampling techniques are usually employed to reduce the bandwidth requirements. A recent study indicates that a custom low-power image sensor can increase a sensor node's lifetime by two orders of magnitudes [**?**]. The proposed image sensor has two different properties from a COTS imager: First, it consumes 1mW of power or even less in active state. Second, it is able to compute the temporal-difference between continuous frames and to filter out redundant data. In fact, an iMote2 node consumes 80 times more power in an active state than in a sleep mode [5]. Microcontroller unit (MCU) and radio communication components on a sensor node stay in a sleep state when image sensor captures the same visual information. This low-power state continues until an exceptional event is detected, for example, the light intensity in a significant number of pixels has changed.

In this paper, we report a custom low-power image sensor, *Mototrigger*, which computes the temporal-difference between two continuous light-intensity frames and communicates them in address-event format. The light-intensity images are also available.

II. Image Sensor Architecture

In the *Mototrigger*, two continuous image-frames are subtracted, and the difference is compared with two predefined thresholds to generate a temporal-difference image. Figure 1 shows the block diagram of the image sensor architecture. The core pixel array consists of 64×64 photodiode pixels.

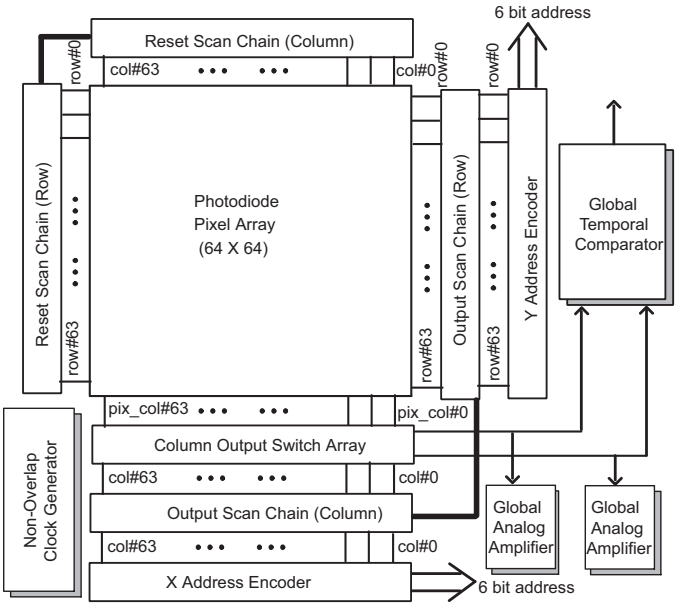

Fig. 1. Image sensor architecture block diagram

The pixel array is driven by two sets of shift registers: one resets the pixels, referred to as *reset scan chain*; the other accesses the pixels and outputs their integration voltages, referred to as *output scan chain*. The pixels in the array are accessed in a row-wise fashion when the pixels in the selected row are read out sequentially. When a pixel is selected by the output scan chain, the pixel outputs two analog signals: the integration voltage of the current-frame and that of the previous-frame voltage which is stored on a MOS capacitor in the pixel. A global comparator is outside the pixel array, and computes the temporal difference between two outputs. A digital pulse (event) is transmitted when the difference passes the predefined thresholds in pixel. Two address encoders are at the pixel's array boundary. They generate a 12-bit address associated with the pixel of the digital pulse. Since there is one global comparator in the sensor, Mototrigger performs sequential analog comparison pixel by pixel.

Two analog buffers reside at the output of the pixel array and output integration voltages from the selected pixel. The timing logic generates proper non-overlapping timing sequences to drive the scan chains and the temporal comparator.

The image sensor simultaneously output one digital signal and two analog signals: the digital signal reports events when the light intensity in the pixels passes the thresholds; the

analog signals report the pixel voltages of the light-intensity images. The temporal-difference images can also be represented in address event format.

III. Pixel Analog Signal Path

Figure 2 shows the signal path from the photodiode to the global comparator. The APS pixel cell is composed of a well-substrate junction photodiode, a reset transistor (M1), three nMOS transistors (M2 to M4) as a source follower, a MOS storage capacitor (M7), two switch transistors (M5 and M6), and four pMOS transistor (M9 to M11) of the column-shared pMOS followers. A pMOS transistor (M1) is used in the reset path of the photodiode because it increases the amplitude of the pixel reset voltage to V_{dd}, and extends the pixel dynamic range. This also eliminates image lags, which is caused by incomplete resets and is evident when an nMOS reset transistor is used. However, the increased output swing comes at the expense of a larger pixel area. In the followers, the access transistors (M3, M9, M11) are used to control the row sequence to access the pixels. The access transistor (M3) is shared by the column of the pixels and controls the output sequence of the column. The global transistors (M14, M15) are the bias transistors which act as the current sources of the source-followers.

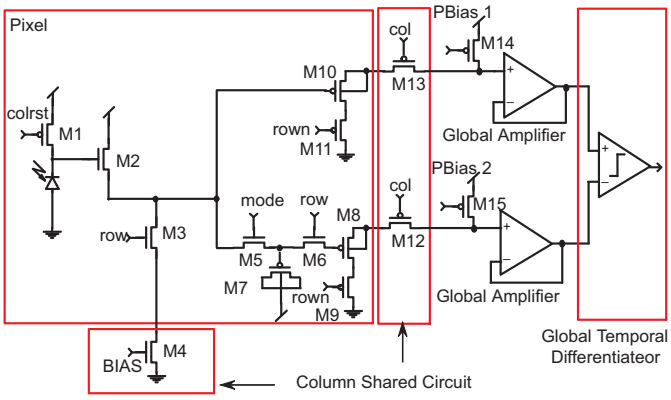

Fig. 2. Signal path from a pixel to the global temporal comparator

The analog storage structure in the pixel is composed of a MOS capacitor and two switch transistors (M5 and M6). The capacitor is implemented with a pMOS transistor by connecting the source and drain to ground, while the bulk is connected to V_{dd}. The pMOS capacitor is chosen because it is more compact than a linear metal-to-metal capacitor available in the bulk CMOS process. The storing voltage range for the pMOS capacitor is between V_{dd} and V_{th}, a range that fits the output of the nMOS source follower. Two column-shared pMOS followers are used to drive the global analog buffer and compensate the DC level-shift of the nMOS source follower. A global temporal comparator is placed at the pixel array's output to computes the temporal-differences.

Figure 3 shows the timing diagram of computing a temporal-difference image. The pixel starts the first-frame integration with a reset in the photodiode. After a certain

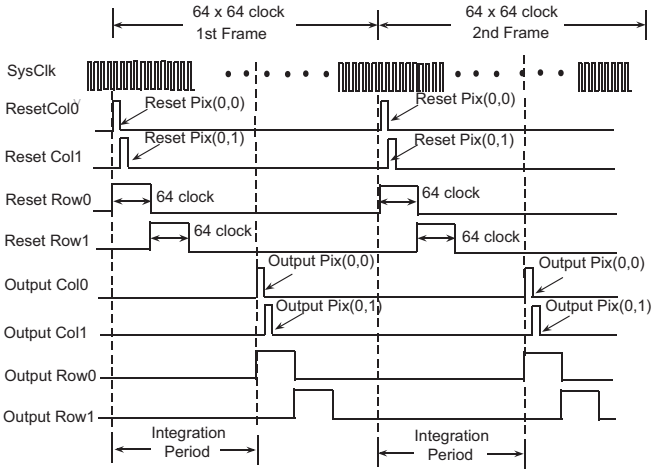

Fig. 3. Timing diagram of the mototrigger sensor

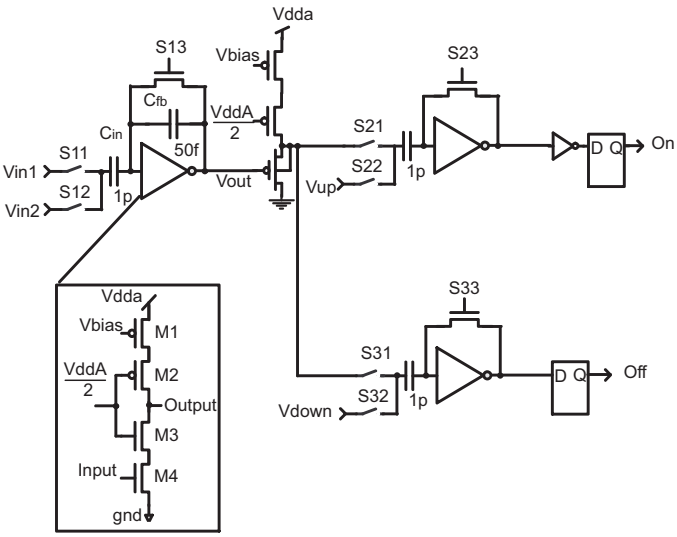

Fig. 4. Global Comparator

integration period, during which light shines on the photodiode, the pixel is addressed by the output scan chain. The integration voltage is stored on the MOS capacitor by asserting the signal of *mode*. After a reset, the pixel starts the second-frame integration. The pixel outputs the second-frame integration voltage and the stored first-frame voltage when the pixel is selected again. A global analog comparator subtracts two voltages and compares the difference with the thresholds. Two light-intensity images are also available from the global output buffers.

IV. Sample/Hold Circuit in Pixel

The sample/hold circuit in the pixel includes two switch transistors (M5 and M6) and one hold capacitor (M7). The purpose of the circuit is to hold the pixel value of the first frame while the pixel is integrating for the second-frame.

The switch-induced error voltage in the sample/hold circuit is caused by two factors: first, the charge injection into the hold capacitor when the switches (M5 and M6) turn off; second, the clock voltage fed through the gate-drain capacitance in

the switch transistors. In order to reduce this error voltage, the switch transistors turn off in the order of M5, M6 and M3 so that the capacitor is connected to the signal source while M6 is switching. Dummy transistors are not used in the pixel in order to minimize the pixel area.

The source and bulk of the transistors (M8 and M10) are connected in order to eliminate the non-linearity due to the body effect. However, this connection causes a charge coupling between the gate and bulk of the transistor M8 when the output of the pMOS follower changes. This can corrupt the stored voltage on the capacitor (M7) when the pMOS follower outputs the other pixel voltages and changes its value. A row-selected switch transistor (M6) is employed in order to shield this negative effect. Different biases are applied to transistor M14 and m15 to reduce the S/H offset in the pixel, which could be as much as 13 mV.

V. Temporal-Difference Comparator Design

The temporal-difference (TD) comparator subtracts the pixel values in two continuous frames and outputs the events when the frame difference exceeds a threshold. In Figure 4, the TD comparator contains three stages: (1) an AC signal amplifier, (2) an inter-stage pMOS follower, (3) two parallel comparators.

The AC amplifier in the first stage is a switched-capacitor (SC) circuit based on a cascoded inverter [**?**]. The cascoded inverter includes two pMOS (M1 and M2) and two nMOS (M3 and M4) connected in series. The input transistor of the inverter (M4) is separated by a capacitor with the input, which switches between the two inputs, V_1 and V_2. The biases are $\frac{V_{dda}}{2}$ for two middle transistors (M2 and M3) and V_{bias} for the top transistor (M1), where the V_{bias} can be varied to change the inverter gain. The cascoded topology is used in the amplifier because it can deliver two main advantages over a common-source amplifier: first, the cascoded inverter shields the input transistor from the output load; second, the cascoded structure delivers a bigger output impedance and a bigger gain. A correlated double sampling (CDS) technique was used in the circuit to eliminate the correlated noise, such as the mismatch in devices. The AC amplifier has a closed-loop gain of 20.

The pMOS follower is implemented as the second stage to drive the two 1pF input capacitor of the next stage.

For event generation in the third stage, a comparison is obtained in two phases:

1) Reset Phase: Switches S_{13} (S_{23}, S_{33}) and S_{12} (S_{22}, S_{32}) are closed, while switch S_{21} (S_{21}, S_{31}) is opened. The reference voltages, V_{up} and V_{down}, are stored on capacitor C_s, together with any correlated noise. The inverter is also initialized to its logic threshold.
2) Compare Phase: Switch S_{11} (S_{21}, S_{31}) is closed, while switches S_{13} (S_{23}, S_{33}) and S_{12} (S_{22}, S_{32}) are opened. The pMOS follower's output is connected and the comparator changes its state when its voltage exceeds the reference voltages.

A digital register is placed after the comparator output in order to store the comparison results and to drive the following components. Similar cascoded inverter topology was used in recent 8-bit analog-to-digital converter publication for low-power purpose [**?**] [**?**].

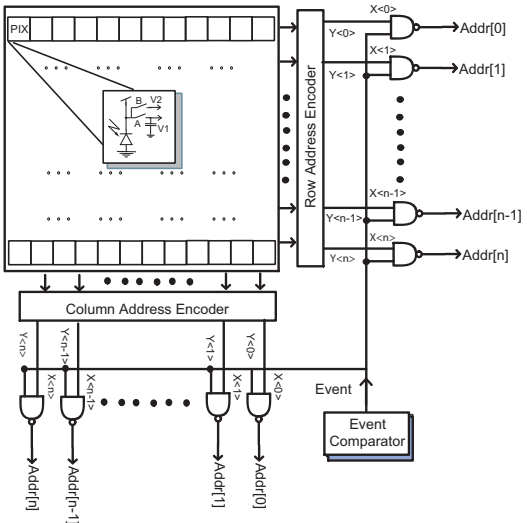

Fig. 5. Event generation in the sensor

VI. Output Buffer and Event Address Encoder

As an *output buffer*, the global amplifier in Figure 2 drives a capacitance load of 15pF outside the chip. A single-stage operational transconductance amplifier (OTA) is implemented because any compensation capacitor other than the load capacitance results in extra power consumption to drive them. The open-loop gain of the OTA is 300V/V in simulations. The amplifier has an input range between 0.5V and 3V which matches the output range of the array.

Mototrigger utilizes an event-based digital representation, known as *address-event representation* (AER) [**?**], [**?**] and outputs temporal-difference visual information. AER is an asynchronous communication protocol. An imager transmits events when changes in pixels exceed a threshold. In Mototrigger, collision detection circuit is unnecessary because pixels are polled sequentially, and only one pixel is addressed at a time. Figure 5 shows the system diagram of the event-address generation in the image sensor. When a pixel is selected, two integration voltages in the pixel, the previous-frame and current-frame value, are subtracted in the global comparator. An event is generated if the difference exceeds the thresholds. This event signal sets the *AND*-gates in the address encoder when the encoder outputs the selected pixel's address.

VII. System Experimental Performance

The sensor was fabricated on a 0.5μm bulk CMOS process (V_{tn}=0.75V and V_{tp}=$-$0.85V). The process features three metal layers and one polysilicon. The die size is 3$\times$3mm^2 including the pixel array, row/column logic, global output buffer, 25 pads and test structures. The micrograph of the chip is shown in Figure 6. The sensor operates at a power supply range between 2.5V and 3V.

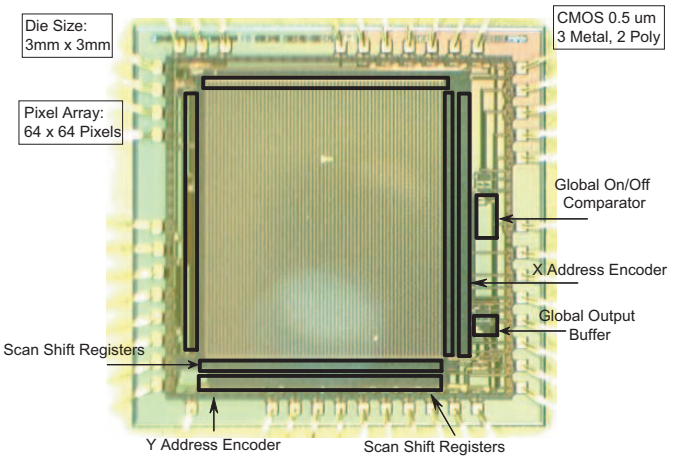

Fig. 6. Image sensor tapeout

(a) (b)

Fig. 7. Output images. Images are taken at normal room light condition. Figure (a) shows a 20 dollar bill. Figure (b) shows temporal-difference images of a waving hand.

Figure 7 shows the light-intensity images (upper) and temporal-difference images (lower) from the sensor, which is running at 60 light-intensity frame/s. Experimental video outputs of the imaging system can be viewed at http://pantheon.yale.edu/∼zf5/MT.html

Table I summarize the performance of the mototrigger. The image sensor is capable of computing 30 temporal-difference frame/s and consumes 1.2mW at a power supply of 3.3V.

VIII. Conclusion

We report a low power temporal feature extracting image sensor, *Mototrigger*. The sensor is a 64×64 active pixel image sensor (APS) with temporal-difference computation capability. The sensor targets low-power image sensing in sensor networks. The sensor computes temporal-difference between two light-intensity frames and communicates in address-event format. The chip is capable of computing 30 temporal-difference frames per second, consuming 1.2mW at 3V. Intensity image frames are also available at the speed of 60 frames per second.

In comparison to the temporal-contrast imagers using parallel comparing [**?**] [**?**] , mototrigger shares a comparator among all pixels. The power cost is reduced because it avoids the scenario where each pixel has a comparator and continuously performs analog comparison even when sparse events are triggered. Also Mototrigger outputs light-intensity snapshots, which could be important in surveillance applications.

TABLE I
Summary of Image Sensor Performance

Pixel Array	64×64
Pixel Size	29×28 μm^2
Technology	0.5μm 1P, 3M CMOS
Pixel Fill Factor	23%
Chip Size	3×3 mm^2
Dark signal	15mV/s
Reset Fixed Pattern Noise	0.07%
Fixed Pattern Noise	0.4%
Operating voltage	2.5∼3V
Maximum light intensity frame rate	60 frame/s
Maximum pixel readout rate	245k pixel/second
Power Consumption	1.2mW

Mototrigger also differs from the previous APS temporal-difference imager design [**?**] in the following ways: First, it includes a cascaded inverter as a comparator structure for further power benefit. Second, pixels in Mototrigger includes one storage capacitor in order to simplify the pixel design and save pixel area. Third, it uses a source follower between the photodetector and voltage storage structure so that the photodector does not share its charge with the capacitor when pixel stores its integration voltage. In this way, image pixel can operate at a faster speed and with a higher light sensitivity.

References

[1] E. Fossum, "CMOS image sensors: Electronic camera-on-a-chip," *IEEE Transactio ns on Electron Devices*, vol. 44, no. 10, pp. 1689–1698, 1997.

[2] N. Stevanovic, "A CMOS image sensor for high-speed imaging," in *IEEE International Solid-State Circuits Conference, 2000. Digest of Technical Papers. ISSCC. 2000*, San Francisco, CA, February 2000, pp. 104–105.

[3] P. Lichtsteiner, C. Posch, and T. Delbruck, "A 128$times$128 120db 30mw asynchronous vision sensor that responds to relative intensity change," in *2006 International Solid State Circuits Conference (ISSCC 2006)*, 2006, pp. 508–509.

[4] E. Park and E. Culurciello, "Address-event video streaming over wireless sensor networks," in *IEEE International Symposium on Circuits and Systems, 2007*, 2007.

[5] D. Jung, T. Teixeira, and A. Savvides, "Sensor node lifetime analysis: Models and tools," *IEEE ACM Transactions on Sensor Networks*, March 2007.

[6] Intel, *Intel Mote 2 Engineering Platform*, Intel Co., 2006.

[7] Z. Fu and E. Culurciello, "An ultra-low power silicon-on-sapphire adc for energy-scavenging sensors," in *IEEE International Symposium on Circuits and Systems, ISCAS*, Kos, Greece, May 2006.

[8] ——, "Nano-watt silicon-on-sapphire adc using 2c-1c capacitor chain," *IEE Electronics Letters*, vol. 42, no. 6, pp. 590–592, March 2006.

[9] K. A. Boahen, "Point-to-point connectivity between neuromorphic chips using address-events," *IEEE Tracsactions on Circuits and Systems*, vol. 3, pp. 100–117, 2000.

[10] K. Boahen, "Retinomorphic vision systems," *IEEE MicroNeuro*, pp. 2–14, 1996.

[11] M. C. R. Etienne-Cummings, "A dual pixel-type array for imaging and motion centroid localization," *IEEE Sensor Journal*, vol. 2, no. 6, pp. 529 – 548, December 2002.

[12] V. Gruev and R. Etienne-Cummings, "A pipelined temporal difference imager," *IEEE Journal of Solid-State Circuits*, vol. 39, no. 3, pp. 538 – 544, March 2004.