Bandwidth efficient frame-based CMOS image sensor for edge and contour detection applications

Akhlesh Rajak[1,2,3], Vibhav Patel[2], and Mukul Sarkar[1,2]

[1]Indian Institute of Technology, Delhi, India
[2]DV2JS Innovation LLP (3rdiTech), Delhi, India
[3]Email: eez198335@ee.iitd.ac.in, Tel: +91-7415822490

***Abstract*— In this paper a bandwidth efficient frame-based CMOS image sensor is designed and fabricated for edge and contour detection applications. A novel in-pixel spatial contrast computation with non-overlapping of a 2 x 2 pixel kernel is used to reduce the pixel readout complexity and detect an edge and contour of an object without further processing. A 128 x 128 vision sensor prototype is fabricated in TSMC 0.18 μm CMOS CIS process with a pixel pitch of 11.2 μm, fill factor of 11% and frame rate of 90 fps.**

keywords- Spatial Contrast (SC), Dynamic Vision Sensor (DVS), CMOS Image Sensor (CIS), Event Detection (ED)

I. INTRODUCTION

Event-based CMOS image sensors are classified into asynchronous event-based readout [1-2] and synchronous frame-based readout [3-6]. In asynchronous event-based readout data redundancy reduces but at the cost of latency and readout complexity. The event-based readout depends on the imaging resolution and encoding scheme, for example, a 128 x 128 pixel array, each pixel event output contains a 14-bit address and 1-bit of event information [1]. As the pixel array size increases, the output data rate also increases. An address event representation (AER) arbiter and in-pixel asynchronous logic circuits are required to encode each pixel event [2]. The increased output data rate requires high speed processing circuitry to reduce the data collection time and increase the readout complexity.

Spatial contrast features such as edge and shape are widely used in various applications such as object detection or recognition. As pixel array size increases, the information needed to process is also increased. So, it is very difficult to process the large amount of data in limited bandwidth. To solve this bandwidth limitation and readout complexity, a frame-based readout is used. In frame-based readout, non-overlapping pixels are selected and only 1-bit per pixel is transmitted. The proposed sensor uses in-pixel spatial contrast computation for edge and contour detection applications which does not require further processing. This reduces the processing time and complexity of the system involved in the output data processing.

The rest of the paper is organized as follows: imager architecture and circuit level blocks are described in section II, measurement setup and results are described in section III and conclusion is presented in section IV.

II. SYSTEM ARCHITECTURE AND CIRCUITS

Fig. 1(a) shows the proposed non-overlapping technique in which a 2 x 2 pixel kernel is used to reduce the pixel readout complexity. In this scheme, each pixel is read only once to increase the bandwidth efficiency. Fig. 1(b) shows the sensor architecture of the proposed spatial contrast detection algorithm. The architecture has a pixel array of 128 x 128, row decoder, column decoder and a 1-bit column memory blocks. The row and column decoders select two rows and two columns at a time. These rows are first reset, integrated and then readout using column memory. The output bits corresponding to the first row are encoded as B_1 and B_2, while the second row are encoded as B_3 and B_4. These 4 bits (2 x 2) pixel outputs are stored in column memory and sent parallelly off chip.

A. Proposed pixel architecture

Fig. 2 shows the block-level diagram of the proposed pixel. A 4T pixel using a pinned photodiode (PPD) is used as a sensing element. The data from the neighbouring pixels are compared to detect the spatial contrast between the neighbouring pixels. The 2 x 2 pixels source follower outputs are compared using 4-input winner-take-all (WTA) and loser-take-all (LTA) circuits. The WTA circuit gives a winner output corresponding to the highest change among the inputs while the LTA circuit gives a loser output corresponding to the lowest change among the inputs. The WTA and LTA outputs act as a clock for the switches to find the maximum (v_w) and minimum (v_l) change values of the pixels. The spatial contrast is detected and compared using a comparator with the differential input (v_w and v_l) and the differential thresholds (v_{rp} and v_{rm}). The comparator takes a decision on

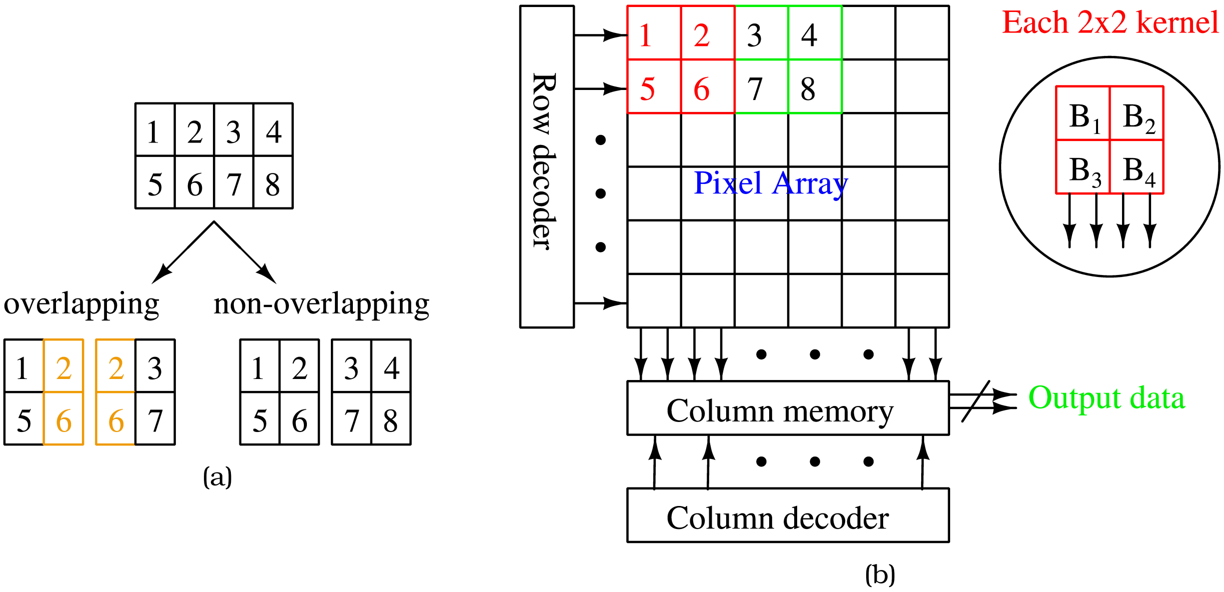

Fig. 1: (a) Proposed non-overlapping technique and (b) sensor block diagram.

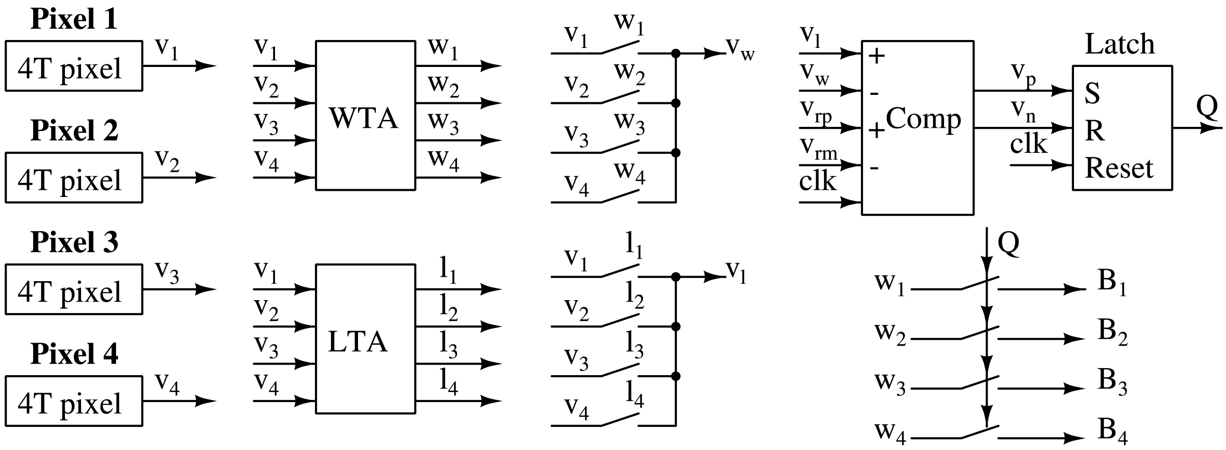

Fig. 2: Block level diagram of proposed pixel.

negative edge of the clock signal (clk) and the corresponding output is stored in the one-bit latch memory. If input difference is less than the threshold difference then there is no event and the latch output Q will be at logic 0 else the output will be logic 1. The latch output Q acts as a clock for the switches to pass the WTA outputs to the column memory.

B. Comparator, latch & memory circuits

Fig. 3(a), (b) and (c) show the schematic diagrams of comparator, latch and column memory respectively. The 4-input differential dynamic comparator is used to save power consumption of the pixel. The clock controlled cascode transistors are used to reduce the kickback noise generated from coupling between input and output through parasitic capacitances. The output of the comparator is stored by the reset-set (RS) latch. The response time of the comparator is 2 ns and resolution is 7 mV. The latch output is stored in a 1-bit column memory. The inverter-based memory circuit is used instead of latch-based memory to avoid the positive feedback of the latch which may amplify the noise event.

C. Readout operation

Fig. 4 shows the timing diagram of the control signals. The row select, row reset, column memory reset, transfer gate and clock signals are represented as SEL, RES, Rst, TG and clk respectively. In each cycle, the rows are reset, integrated and readout in a rolling shutter mode. Since two rows and two columns are selected at a time, 64 cycles are needed to read a single frame. The maximum frame rate of the sensor is 90 fps with minimum integration time (t_{int}) of 100 μs and readout time (t_{ro}) of 64 μs.

III. MEASUREMENT RESULTS

A. Measurement setup and results

Fig. 5 shows a chip micrograph of the proposed sensor. A 128 x 128 vision sensor prototype is fabricated in TSMC 0.18 μm CMOS

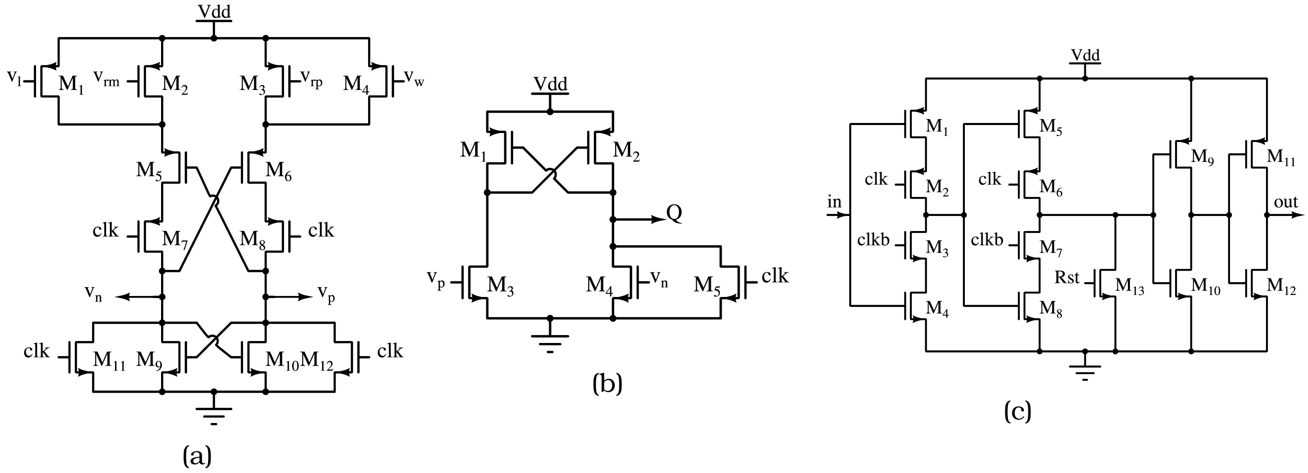

Fig. 3: (a) Comparator schematic, (b) latch schematic and (c) column memory schematic.

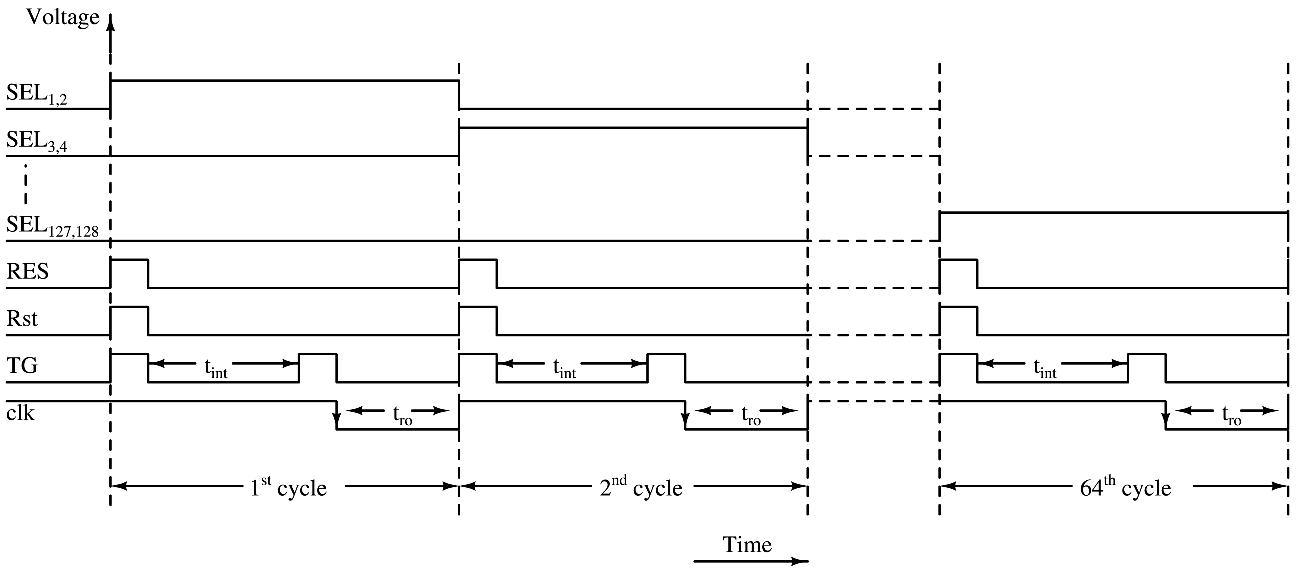

Fig. 4: Timing diagram of the control signals.

CIS process with pixel pitch of 11.2 μm and fill factor of 11%. The measurement setup used to test the prototype sensor is shown in Fig. 6. A printed circuit board (PCB), field programmable gate array (FPGA) board and optical wide-angle lens are used to test the fabricated sensor. The sample images are captured with the help of a camera link interface to the laptop using USB cable.

Fig. 7(a) and (c) show the sample images while Fig. 7(b) and (d) show the captured images in 1-bit edge detection (ED) mode. The alphabets A and B and numbers 2 and 5 are taken as reference images for detecting the different edges for bandwidth constrained environment applications. Similarly, panda, panda face, right arrow and star symbol are taken as references to observe the object edge and shape. The images are captured in the indoor light conditions with an integration time of 500 μs and threshold voltage of 50 mV. Output images clearly detect the alphabets, numbers, sign and face shapes.

B. Performance and comparison table

The performance and comparison table of the sensor is shown in table I. The sensor uses two supply voltages 3.3 V and 1.8 V. The supply voltage of 3.3 V is used to drive the buffer of the transfer gate (TG) signal. While the supply voltage of 1.8 V is used in the rest of the core area. The sensor consumes a total power of 92 mW @ 90 fps. Most of the sensor power consumed by the WTA and the LTA circuits are used in the pixel to reduce the readout time. The measurement parameters are compared with the state-of-the-art results shown in the comparison table. The chip occupies lesser area including pads is 2.65 mm x 2.52 mm as compared to the other sensors.

IV. CONCLUSION

A 128 x 128 in-pixel spatial contrast vision sensor prototype is designed and fabricated in

TABLE I: Comparison table

Parameters	JSSC 2015 [3]	VLSI symp. 2017 [4]	VLSI symp. 2018 [5]	TCAS-I 2019 [6]	**This work**
Technology	0.35 μm	0.18 μm	0.18 μm	0.35 μm	0.18 μm
Supply (V)	3.3/1.5	3.3/1.8	1.2/0.8	3.3/1.5	3.3/1.8
Pixel pitch (μm^2)	26 x 26	31 x 31	7.9 x 7.9 (1.75T pinned)	17 x 17	**11.2 x 11.2**
Fill factor (%)	15	19	55	15	11
Array size	110 x 110	256 x 256	128 x 108	64 x 64	128 x 128
Frame rate (fps)	30	30	30	15	**90**
Feature	SC[1] (LBP[2])	Spatial-temporal	SC (ED[3])	SC (LBP)	SC (ED)
Power	30 μW	2.18 mW	12.7 μW	35 μW	92 mW
Chip size (mm^2)	13	6.2 x 4	2.5 x 2	1.2 x 1.4[#]	**2.65 x 2.52**

[1]spatial contrast, [2]local binary pattern, [3]edge detection [#]core area only

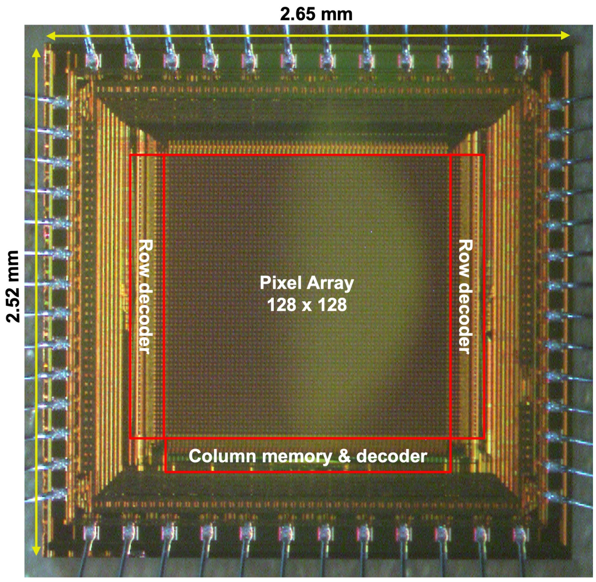

Fig. 5: Chip micrograph of the sensor.

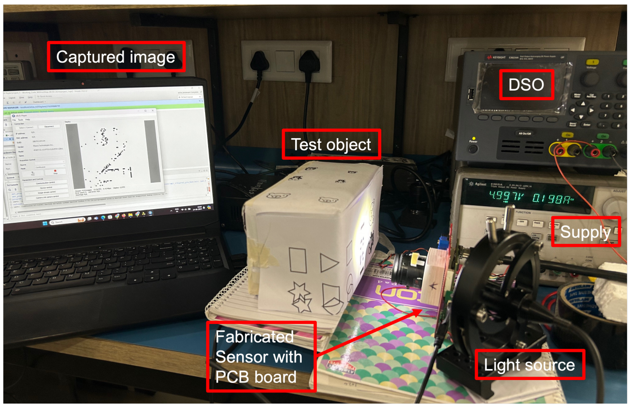

Fig. 6: Measurement setup.

TSMC 0.18 μm CMOS CIS process with a pixel pitch of 11.2 μm and fill factor of 11%. To process the large amount of data in limited bandwidth, a non-overlapping of pixels with frame-based readout is used. The sensor achieved a maximum speed of 90 frames/second using 1-bit per pixel readout.

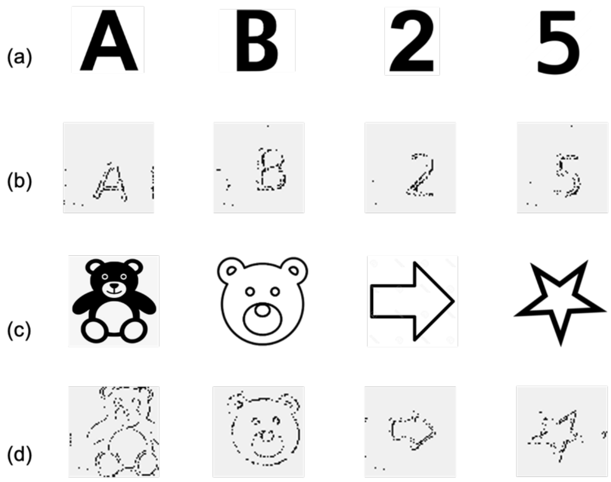

Fig. 7: Reference images in (a) & (c) and captured images in (b) & (d) using fabricated sensor.

References

[1] P. Lichtsteiner, C. Posch and T. Delbruck, "A 128 × 128 120 dB 15 μs latency asynchronous temporal contrast vision sensor," in *IEEE Journal of Solid-State Circuits*, vol. 43, no. 2, pp. 566-576, Feb. 2008, doi: 10.1109/JSSC.2007.914337.

[2] M. Yang, S. -C. Liu and T. Delbruck, "A Dynamic Vision Sensor With 1% Temporal Contrast Sensitivity and In-Pixel Asynchronous Delta Modulator for Event Encoding," in IEEE Journal of Solid-State Circuits, vol. 50, no. 9, pp. 2149-2160, Sept. 2015, doi: 10.1109/JSSC.2015.2425886.

[3] A. Berkovich, M. Lecca, L. Gasparini, P. A. Abshire and M. Gottardi, "A 30 μW 30 fps 110 × 110 pixels vision sensor embedding local binary patterns," in *IEEE Journal of Solid-State Circuits*, vol. 50, no. 9, pp. 2138-2148, Sept. 2015, doi: 10.1109/JSSC.2015.2444875.

[4] K. Lee, S. Park, S. Y. Park, J. Cho, and E. Yoon, "A 272.49 pJ/pixel CMOS image sensor with embedded object detection and bio-inspired 2D optic flow generation for nano-air-vehicle navigation," in *Proc. Symp. VLSI Circuits*, Kyoto, Japan, Jun. 2017, pp. 294–295.

[5] X. Zhong, Q. Yu, A. Bermak, C. Tsui and M. Law, "A 2 pJ/pixel/direction MIMO processing based CMOS image sensor for omnidirectional local binary pattern extraction and edge detection," 2018 *IEEE Symposium on VLSI Circuits*, 2018, pp. 247-248, doi: 10.1109/VLSIC.2018.8502214.

[6] M. Gottardi and M. Lecca, "A 64 x 64 pixel vision sensor for local binary pattern computation," in *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 66, no. 5, pp. 1831-1839, May 2019, doi: 10.1109/TCSI.2018.2883792.