A Frame-Differencing CMOS Image Sensor for >120-dB HDR Temporal Event Detection Using Pixel-Grouped Multiple Exposure

John Carl Joel Salao Marquez*, Chih-Cheng Lu‡, Min Sun*, Meng-Fan Chang*, Kea-Tiong Tang* and Chih-Cheng Hsieh*
*Department of Electrical Engineering, National Tsing Hua University, Hsinchu, Taiwan
‡Industrial Technology Research Institute (ITRI), Hsinchu, Taiwan

***Abstract*—This paper presents a frame-differencing CMOS image sensor for high-dynamic-range temporal event detection. Conventional single-exposure frame-differencing sensors suffer from a limited event-detection dynamic range, which restricts reliable ON/OFF event detection under large illumination variations. The proposed approach spatially distributes multiple exposure times within a compact 2×2 pixel group, allowing event detection over a wide illumination range while preserving frame-compatible output and simple pixel circuitry. The proposed sensor employs an ECS-free subframe-differencing pixel with two local memory capacitors, which introduces similar parasitic loading during the reference and differencing phases and thereby minimizes parasitic-induced frame-differencing error without requiring an exposure compensation scheme. Fabricated in 0.18-µm CMOS, the prototype supports 256×256 raw-image and frame-difference modes and 128×128 ED/HDR-ED modes. In single-exposure ED mode, it achieves a maximum frame rate of 935 fps and an iFoM of 3.45 pJ/pix-frame. In HDR-ED mode, it achieves a measured temporal event-detection dynamic range of more than 120 dB. The proposed 2×2 grouping trades spatial resolution for wider illumination coverage and simple group-level readout.**

Keywords—CMOS image sensor, frame differencing, event detection, high dynamic range, multiple exposure

I. Introduction

Temporal event sensing is attractive for motion understanding and low-latency perception because only time-varying information is reported. This function is commonly realized by fully asynchronous event-based vision sensors, which generate ON/OFF events directly at the focal plane [1-2]. However, such sensors are generally less compatible with conventional frame-based imaging pipelines and often rely on additional in-pixel amplification or comparison circuits, which increase pixel complexity and may reduce power efficiency.

Frame-differencing image sensors provide a practical alternative for temporal event sensing because temporal differencing is performed at the sensor front end while the output remains frame compatible. Prior FD sensors have demonstrated simple event extraction using photodiode-based pixels, but several limitations remain. In [3], an exposure compensation scheme (ECS) is required to suppress parasitic-induced differencing error. This couples exposure time to readout timing and limits the maximum achievable frame rate. In [4], frame storage is performed during the exposure phase, which increases power consumption and reduces energy efficiency. More fundamentally, prior FD sensors typically perform event detection using a single exposure setting, which limits the event-detection dynamic range under large illumination variation and makes it difficult to detect events simultaneously in dark and bright regions, as illustrated in Fig. 1.

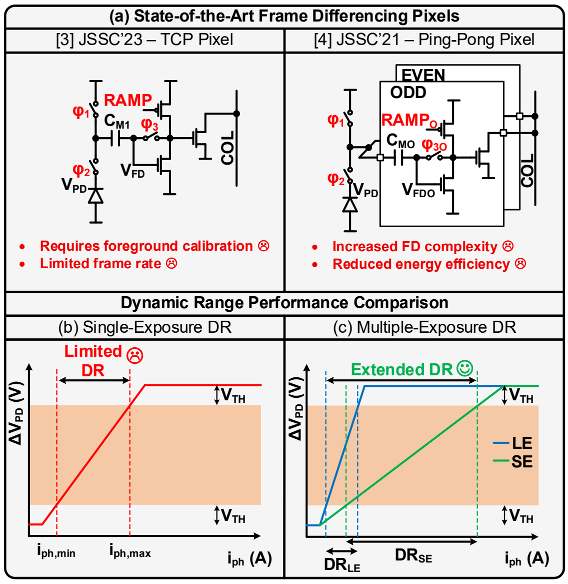

Fig. 1. Comparison of prior frame-differencing pixels and the proposed dynamic-range extension scheme. (a) Representative state-of-the-art frame-differencing pixels. (b) Transfer characteristic of conventional single-exposure operation, which results in limited event-detection dynamic range. (c) Dynamic-range extension achieved by the proposed multiple-exposure scheme.

This work addresses a key limitation of prior frame-differencing image sensors, which have primarily relied on single-exposure event detection and therefore offer limited illumination coverage under large scene variation. To enable high-dynamic-range temporal event detection in a frame-differencing sensor, this work combines an ECS-free subframe-differencing pixel with a pixel-grouped multiple-exposure scheme. Four neighboring pixels are operated with different exposure settings and share a group-level event readout, creating overlapping detection ranges while preserving frame-compatible output and compact circuit implementation. Compared with the fully-asynchronous event-based approach, the proposed HDR event detection mode targets edge-vision systems where precise asynchronous timestamps are not required, but low-power, frame-compatible event maps are preferred, such as always-on visual wake-up systems, object detection systems, and front-end motion processing for conventional vision systems. To the authors' knowledge, this is the first frame-differencing CMOS image sensor that demonstrates HDR temporal event detection using pixel-grouped multiple exposure.

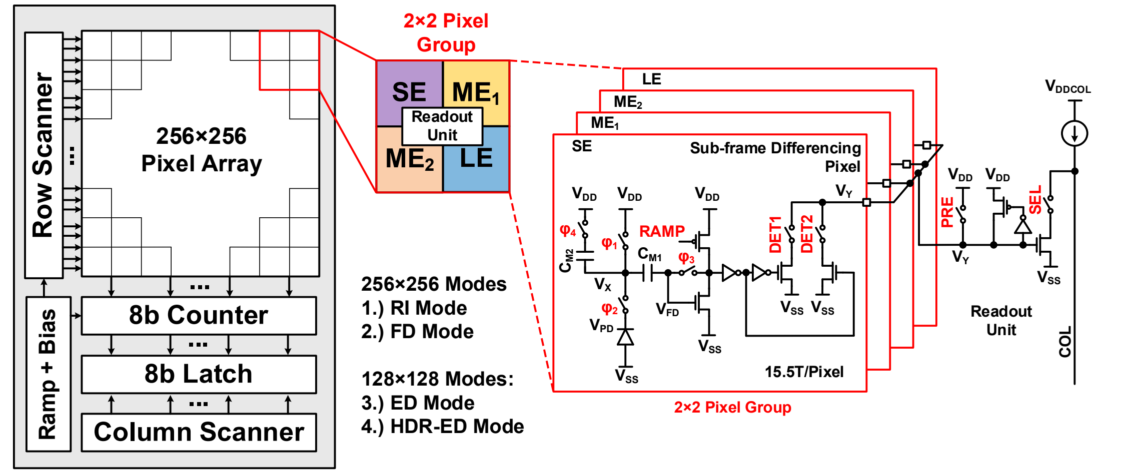

Fig. 2. Architecture of the proposed multimode image sensor and schematic of the 2×2 pixel group used for ED/HDR-ED operation.

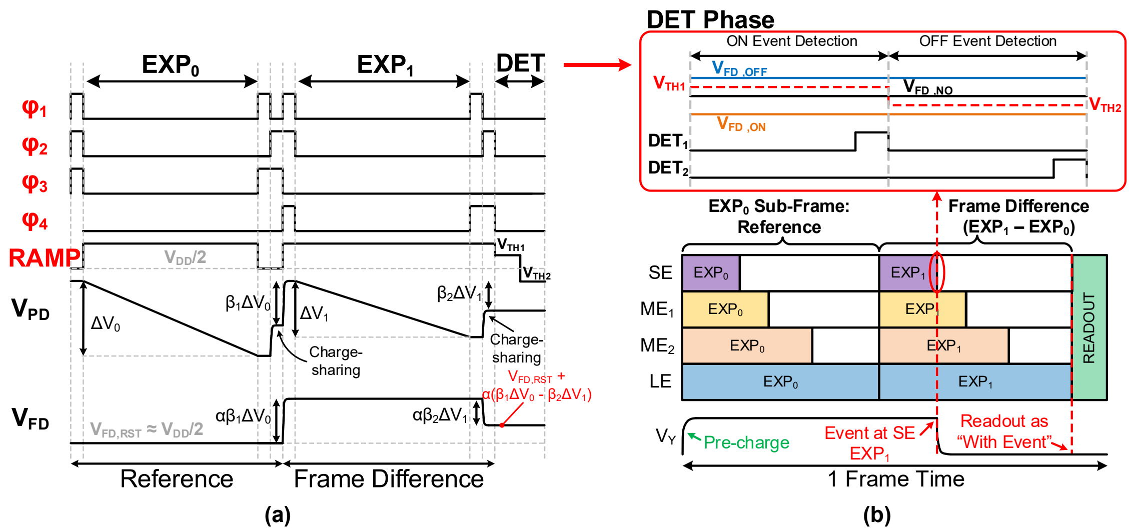

Fig. 3. Circuit operation of the proposed sensor. (a) Timing diagram of the proposed subframe-differencing pixel. (b) Simplified timing diagram of the

II. Sensor Architecture

Fig. 2 shows the overall architecture of the proposed sensor. The prototype supports a 256×256 8-bit raw-image (RI) mode, a 256×256 8-bit frame-difference (FD) mode, and 128×128 1-bit event-detection (ED) modes. In HDR event-detection operation, four neighboring pixels are merged into one 2×2 group and assigned different exposure settings. Group-level readout is then used to report whether any pixel within the group detects an ON or OFF event. This grouping strategy preserves compact pixel design while creating overlapping exposure windows across the scene.

The 2×2 grouping intentionally reduces the spatial resolution from 256×256 to 128×128 during ED and HDR-ED operation. In addition, because event readout is performed at the group level, multiple local changes within one group are merged into a single output event. This tradeoff was intentionally chosen to minimize circuit overhead and preserve compact frame-compatible readout while extending the detectable illumination range through multiple exposure. The proposed architecture therefore targets applications in which robust temporal event detection over wide illumination variation is more critical than full spatial resolution.

III. Circuit Operation

A. Frame-differencing Operation

Fig. 3(a) illustrates the timing diagram of the proposed subframe-differencing pixel. During the reference subframe, the photodiode integrates the incident illumination and stores the signal on the local memory capacitor C_{M1} through charge sharing. During the following differencing subframe, the stored signal is subtracted from the newly integrated signal, generating a voltage proportional to the temporal intensity change at the differencing node. Unlike the prior single-capacitor FD pixel in [3], the proposed pixel also uses C_{M2} during the differencing phase. This makes the effective loading of the reference and differencing operations more similar, which minimizes parasitic-induced FD mismatch without requiring exposure compensation. In addition, the charge-sharing operation during the reference subframe reduces the need for more active frame-storage operation within the pixel, supporting low-overhead implementation.

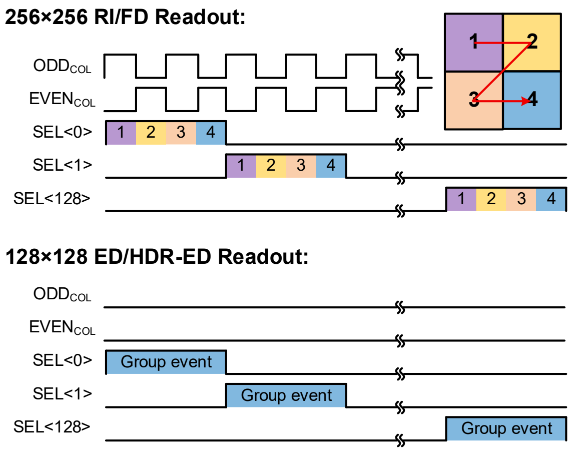

Fig. 4. Readout timing comparison between the full-resolution 256×256 RI/FD modes and the grouped 128×128 ED/HDR-ED modes.

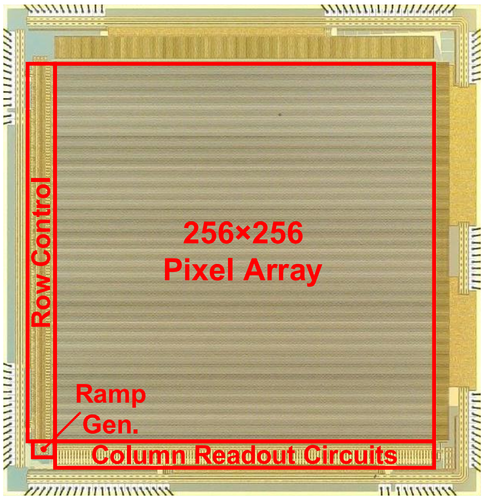

Fig. 5. Chip micrograph of the proposed 256×256 multimode image sensor fabricated in 0.18-µm CMOS.

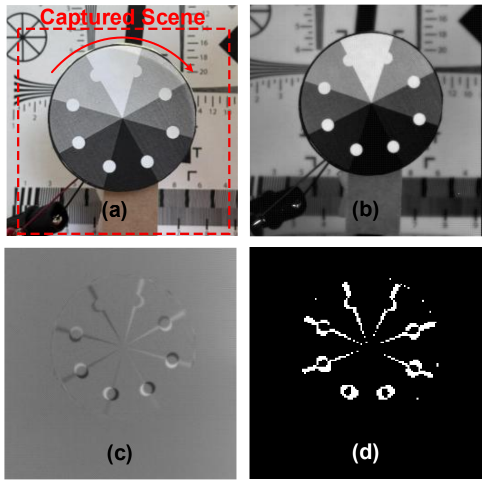

Fig. 6. Measured multimode operation using a test scene consisting of a rotating disc against a complex background. (a) Captured scene. (b) 256×256 RI mode output. (c) 256×256 FD mode output. (d) 128×128 ED mode output.

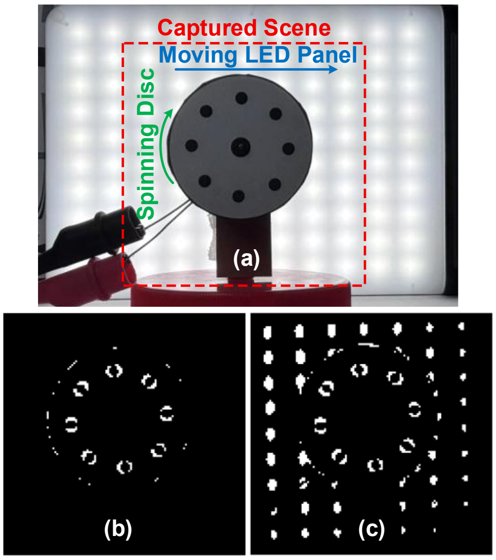

Fig. 7. Measured HDR event-detection demonstration using a spinning disc against a moving LED panel background. (a) Captured HDR scene. (b) Single-exposure ED mode output. (c) Multiple-exposure HDR-ED mode output.

B. High Dynamic Range Event Detection

To extend the ED dynamic range, the subframe-differencing operation is applied to all four pixels in the 2×2 group, with each pixel using a different exposure time, as illustrated in Fig. 3(b). After differencing, two thresholds are sequentially applied to an inverter-based comparator to detect ON and OFF events. A simple dynamic-OR logic then combines the outputs of the four pixels, and the pixel group is flagged as containing an event if any pixel detects an ON or OFF event. This group-level readout enables HDR event detection with minimal additional circuitry.

C. Readout Procedure

Fig. 4 compares the readout timing of the full-resolution 256×256 modes and the grouped 128×128 modes. In the 256×256 modes, the four pixels in each 2×2 group are read out sequentially within one row-select period. In contrast, the 128×128 modes use 2×2 pixel grouping with shared group-level readout, such that only one event result is reported for each pixel group. This reduces readout overhead and enables higher-speed ED/HDR-ED operation at the cost of spatial resolution.

IV. Measurement Results

The proposed image sensor was fabricated in TSMC 0.18-µm CMOS and occupies a total chip area of 4.77×4.91 mm², as shown in Fig. 5. To demonstrate multimode operation, a scene consisting of a rotating disc against a complex background was used as the test target. Fig. 6 shows the measured outputs in the 256×256 RI mode, the 256×256 FD mode, and the 128×128 ED mode, confirming correct multimode operation. In single-exposure ED mode, the sensor achieves a maximum frame rate of 935 fps. This measured value is higher than the 819-fps result reported in [3], while the proposed architecture also avoids ECS-imposed exposure-to-readout timing constraints.

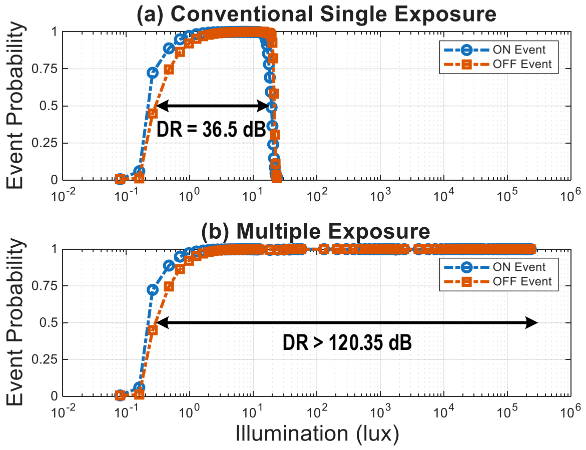

Fig. 8. Measured event-transfer characteristics of the proposed image sensor in (a) single-exposure ED mode and (b) HDR-ED mode. The dynamic range is extracted from the illumination level at 50% event probability.

For HDR event detection, the sensor operates in the 128×128 HDR-ED mode using pixel-grouped multiple exposure. In this mode, different exposure times are assigned within each 2×2 pixel group to extend the detectable illumination range. To preserve meaningful exposure ratios, the operating frame rate in the HDR-ED mode is reduced to 200 fps. This mode therefore prioritizes extended event-detection dynamic range rather than maximum temporal resolution. Fig. 7 shows the measured HDR-ED demonstration using a rotating disc against a moving LED-panel background. Compared with the conventional single-exposure ED mode, the proposed HDR-ED mode maintains event detection over a much wider illumination range.

The measured event-transfer characteristics are shown in Fig. 8. Illumination is swept while applying a constant absolute difference ($|\Delta I| > I_{TH}$). The dynamic range is extracted from the illumination corresponding to 50% event probability. Under the same measurement conditions, the conventional single-exposure mode provides a temporal event-detection dynamic range of 36.5 dB, whereas the proposed pixel-grouped multiple-exposure scheme extends the measured range to more than 120.35 dB. The upper limit in this measurement is set by the available illumination system rather than by the sensor itself.

Finally, Table I summarizes the prototype performance and compares it with prior frame-differencing and event-based image sensors. The proposed sensor achieves a maximum single-exposure ED frame rate of 935 fps and an iFoM of 3.45 pJ/pix-frame. In addition, unlike prior frame-differencing sensors that primarily report single-exposure operation, the proposed sensor demonstrates HDR temporal event detection using pixel-grouped multiple exposure with measured dynamic range exceeding 120 dB.

V. Conclusion

This paper presented a multimode CMOS image sensor that employs pixel-grouped multiple exposure for high-dynamic-range temporal event detection. By assigning different exposure times within each 2×2 pixel group, the proposed architecture extends the usable event-detection illumination range while maintaining frame-compatible output and modest circuit overhead. Fabricated in 0.18-µm CMOS, the prototype supports 256×256 raw-image and frame-difference modes and 128×128 ED/HDR-ED modes. In single-exposure ED mode, the sensor achieves a maximum frame rate of 935 fps with an iFoM of 3.45 pJ/pix-frame. In HDR-ED mode, it achieves a measured temporal event-detection dynamic range of more than 120.35 dB. These results demonstrate a practical multimode frame-differencing sensor architecture that combines competitive ED performance with HDR temporal event detection enabled by pixel-grouped multiple exposure.

TABLE I. Comparison Table

	This Work	JSSC'23 [3]	JSSC'21 [4]	ISSCC'23 [5]
Tech.	**0.18 µm CMOS**	0.18 µm CMOS	0.18 µm CMOS	90 nm CIS + 22 nm CMOS
Array Size	**256×256**	126×126	64×64	640×640
Pixel Pitch	**15 µm**	12 µm	15 µm	2.97 µm
Supply	**0.8V**	0.56V/ 0.8V	0.8V	2.8V/0.8V
ED Method	**Frame Difference**	Frame Difference	Frame Difference	Frame-based EVS**
Max. ED Frame Rate	**935 fps**	819 fps	510 fps	1724 fps
ED-Mode Power	**211.4 µW**	162.6 µW	74.4 µW	79.8 mW
ED-Mode iFoM*	**3.45 pJ/pix-frame**	12.5 pJ/pix-frame	35.6 pJ/pix-frame	113 pJ/pix-frame
DR	**> 120 dB (HDR-ED)**	-	-	> 120 dB (ED Mode)

*iFoM = Power / (Array Size × Frame Rate)
**EVS architecture with a frame-compatible readout

Acknowledgment

The authors would like to thank Taiwan Semiconductor Research Institute (TSRI) and Taiwan Semiconductor Manufacturing Company (TSMC) for the fabrication of the test chip.

References

[1] P. Lichtsteiner, C. Posch and T. Delbruck, "A 128 × 128 120 dB 15 µ s Latency Asynchronous Temporal Contrast Vision Sensor," in *IEEE Journal of Solid-State Circuits*, vol. 43, no. 2, pp. 566-576, Feb. 2008

[2] R. Graca, S. Zhou, B. McReynolds and T. Delbruck, "SciDVS: A Scientific Event Camera with 1.7% Temporal Contrast Sensitivity at 0.7 lux," *2024 IEEE European Solid-State Electronics Research Conference (ESSERC)*, Bruges, Belgium, 2024, pp. 205-208.

[3] M. -Y. Chiu et al., "A Multimode Vision Sensor With Temporal Contrast Pixel and Column-Parallel Local Binary Pattern Extraction for Dynamic Depth Sensing Using Stereo Vision," in *IEEE Journal of Solid-State Circuits*, vol. 58, no. 10, pp. 2767-2777, Oct. 2023.

[4] T. -H. Hsu et al., "A 0.8 V Multimode Vision Sensor for Motion and Saliency Detection With Ping-Pong PWM Pixel," in *IEEE Journal of Solid-State Circuits*, vol. 56, no. 8, pp. 2516-2524, Aug. 2021.

[5] A. Niwa et al., "A 2.97µm-Pitch Event-Based Vision Sensor with Shared Pixel Front-End Circuitry and Low-Noise Intensity Readout Mode," *2023 IEEE International Solid-State Circuits Conference (ISSCC)*, San Francisco, CA, USA, 2023, pp. 4-6.