# CDS Free Frame Differencing Event Vision Pixel with Lateral Overflow Capacitor for Dynamic Range Extension

Marko Jaklin, D. García-Lesta, P. López, V.M. Brea
*Centro Singular de Investigación en Tecnoloxías Intelixentes (CiTIUS)*
*Universidade de Santiago de Compostela*
Santiago de Compostela, Spain
email: marko.jaklin@usc.es

***Abstract*—This paper reports on a redesign of an event pixel by frame differencing with a lateral overflow capacitor for dynamic range extension. In the first prototype, the inclusion of an in-pixel correlated double sampling (CDS) and an algorithm for dynamic range extension resulted in a large pixel area, with a 32 $\mu$m pitch in 180 nm CMOS technology. This paper studies a CDS free version of our former pixel through post-layout simulations. The pixel pitch is shrinked to 22 $\mu$m, power consumption is lowered and speed is preserved.**

***Index Terms*—Event-based sensors, frame-differencing, HDR, CMOS**

## I. Introduction

Both dynamic vision sensors [1]–[4] and frame differencing based event vision sensors [5], [6] outperform conventional cameras in temporal resolution and power consumption. However, because of the extra circuitry embedded in the pixel they suffer from larger pixel size. Dynamic vision sensors achieve lower temporal resolution than frame differencing based event sensors due to their continuous asynchronous operation, and outperform them in dynamic range too due to their use of logarithmic photodiode. Frame based event sensors do not track the changes in the scene continuously but they compare the scene between frames, issuing events when the difference between two frames exceeds a certain threshold. The output of an event pixel is a digital information package called event, with the pixel location in the pixel array, polarity (positive or negative changes) and timestamp (when the event occurred).

This and our previous design [7] aim to narrow the gap in dynamic range between dynamic and frame based event vision sensors through a lateral overflow capacitor that collects photoelectrons that exit conventional photodiodes due to oversaturation [8].

The initial design (Fig. 1) exhibits certain limitations, as that of low fill factor- 2.8%- and a larger pixel size, with a

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 860370, from AEI through Grant PID2021-128009OB-C32, from the Xunta de Galicia-Consellería de Cultura, Educación e Ordenación Universitaria Accreditation 2019–2022 ED431G-2019/04 and Reference Competitive Group Accreditation 2021–2024, GRC2021/48 (FEDER). We also thank the **HISILICON** programme for funding for CMOS IC fabrication.

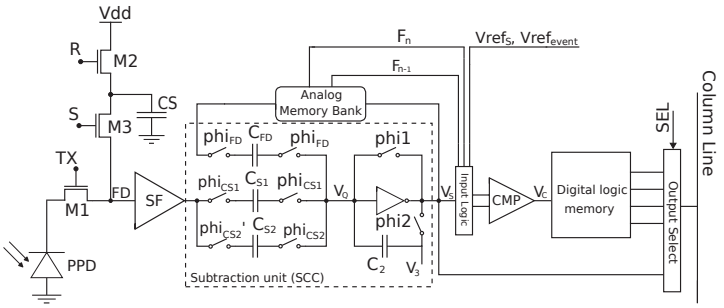


Fig. 1: Schematics of original design with CDS. The analog memory bank (AMB) stores the result of CDS increasing the complexity and size of the pixel( [7]).

32.3 $\times$ 32.3 $\mu m^2$ pitch. These characteristics are attributed to the numerous circuits required to facilitate the operation of various algorithms such as CDS (Correlated Double Sampling), HDR (High Dynamic Range), and frame differencing. Additionally, the incorporation of digital logic to synchronize these processes contribute to an increased number of circuits, ultimately increasing the complexity of the overall operation. Consequently, addressing these challenges is crucial for enhancing the performance and efficiency of the system.

The main change of our new pixel with respect to the previous one [7] is the lack of CDS, which, in turn, leads to less additional circuitry, as there is no need for the analog memory bank (AMB) shown in Fig. 1. This partially alleviates or compensates for the larger noise of a CDS free pixel. Besides, both the operation complexity of three operation stages, namely, CDS, HDR and frame differencing to provide events, is greatly simplified, and the pixel size shrinks, allowing for a higher fill-factor, which is beneficial for the overall operation. Finally, while the CDS cannot be done inside the pixel, this structure would still allow for CDS to be executed outside of the chip.

## II. CDS Free Pixel Operation

The schematics of the new pixel is shown in Fig. 2. Switches are implemented as transmission gates, while the 4 transistor

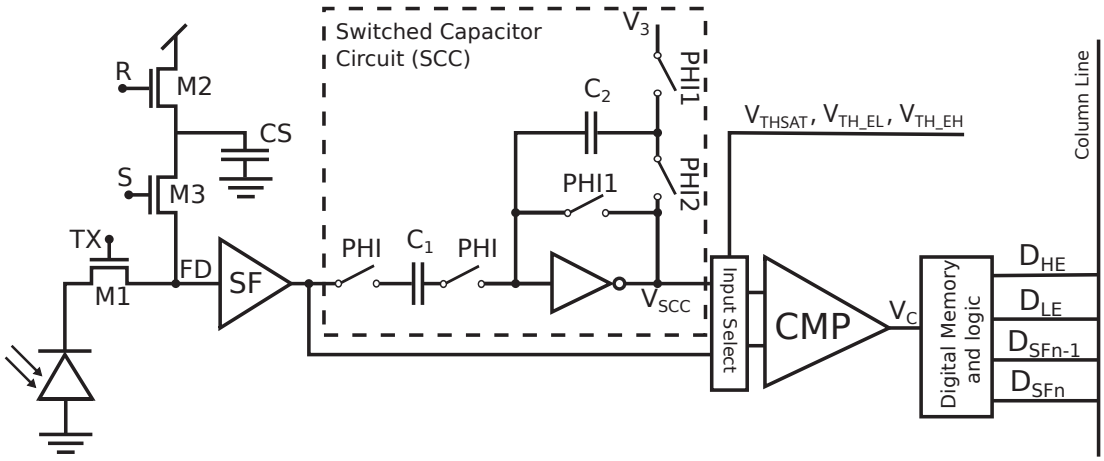


Fig. 2: Schematics of our CDS free pixel. Unlike the previous version with CDS ( [7]), this one does not include an analog memory bank. Also, the input select block and the digital logic are simpler, requiring less transistors.

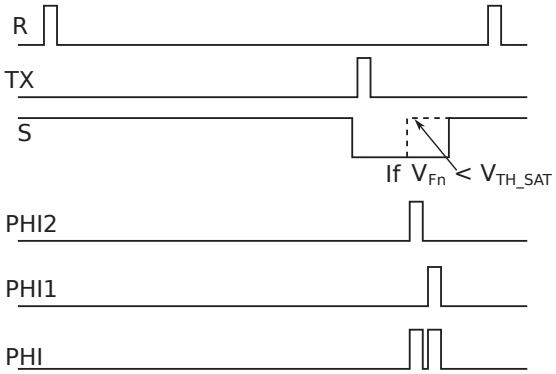


Fig. 3: Timing diagram of the new CDS free pixel with the waveform of the main control signals.

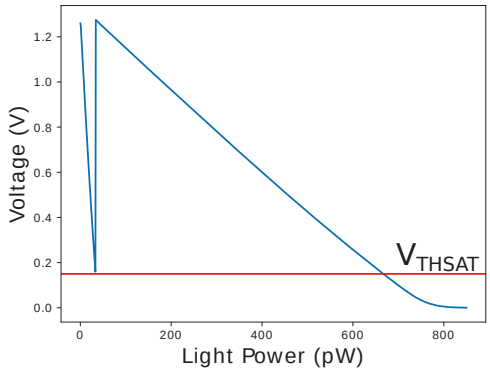


Fig. 4: Signal switching. If the signal generated by non saturated electrons falls below $V_{THSAT}$ the pixel will use signal generated by saturated electrons for frame differencing.

active pixel sensor (4T-APS) is made up of transistors M1, M2, M3, controlled by signals $TX$, $R$, $S$, and the source follower (SF) are powered to 3.3 V, while the rest of the circuitry is supplied with 1.8 V to diminish power consumption. Our pixel provides an event when the frame difference between two consecutive frames $\Delta F = F_{n-1} - F_n$, with $F_n$ being the current frame and $F_{n-1}$ the previous one, exceeds a user defined threshold, $V_{THSAT}$ in Fig. 2. The frame difference is executed by the switched capacitor circuit (SCC) of Fig. 2, while the comparator (CMP) implemented with a 5T-OTA configuration with PMOS as input transistors, determines if the frame difference is large enough to generate events. The removal of CDS as one of the three operation stages of the original pixel- namely, CDS, HDR and frame differencing to generate events, makes the input select block- Input Select in Fig. 2- that manages the inputs to the CMP also simpler, and thus with less number of transistors. Digital logic is simpler too.

As mentioned, for dynamic range extension, an overflow capacitor (CS) that stores saturated electrons is used. Whether to use the charge stored on CS or not is determined by checking if the photodiode is saturated or not. If saturation is detected, then signal $S$, which is a combination of the outside signal and in-pixel logic, connects the floating diffusion (FD) with the overflow capacitor CS, making the output of the source follower SF a combination of saturated and unsaturated electrons. Otherwise, signal $S$ is kept low and only non-saturated electrons are provided. We will refer to the signal generated by not-saturated electrons as $S1$, and by saturated and not-saturated electrons as $S2$.

The timing of operations is displayed in Fig. 3. We begin with a signal reset $R$ that is pulsed high to reset the FD and CS nodes. The signal $S$ is also high connecting FD and CS and remains high to keep transistor M3 on to serve as a path to CS for photo-electrons in case of saturation. Signal $S$ is pulled low just before $TX$ is pulled high. Pulsing $TX$ high activates M1 and photodiode transfers the charges to the FD node that is now isolated from CS. At this point, the SF is enabled and its output connects to the one of the inputs of the comparator CMP through the input select

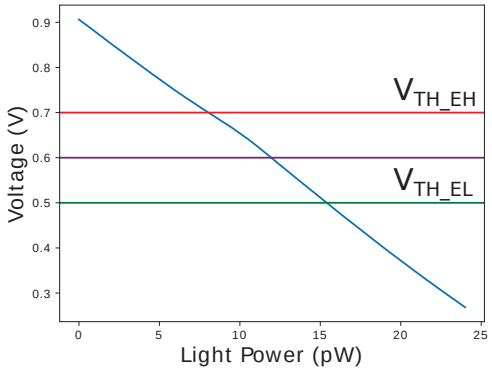


Fig. 5: Result of frame differencing. $V_{TH\_EH}$ and $V_{TH\_EL}$ represent upper and lower event threshold voltages. The graph shows the result of frame differencing while $F_n$ is kept constant at around 12 pW input light power. $V_3$ is a user programmable signal to adapt voltages between the inverter of the SCC and the capacitor CMP is set in this case to 600 mV.

block. Also, at this point, the comparison $V_{Fn} < V_{THSAT}$ is made, and if true, the result is interpreted as a saturation of photodiode, and consequentially pixel switches to use signal S2 by setting $S$ high and connecting FD with CS. The result of $V_{Fn} < V_{THSAT}$ is saved in a digital memory storing circuit, $D_{SF_n}$. Fig. 4 shows signal switching in effect. The difference in slopes is due to the difference in capacitance between CS and FD + CS nodes.

After determining the signal type to be used, the SCC block executes frame differencing. While $PHI2$ is high, the SCC frame differencing is completed, providing the output $V_{FD}$ given by the formula:

$$\mathrm{V_{FD}} = \frac{\mathrm{C_1}}{\mathrm{C_2}}(\mathrm{F_{n-1}} - \mathrm{F_n}) + \mathrm{V_3} \tag{1}$$

$F_{n-1}$ represents the analog value of the previous frame while $F_n$ is the value of the current frame. The result of frame differencing is shown in Fig. 5. As $PHI2$ is held high, the SCC output is sequentially compared to the user-defined thresholds; first to $V_{TH\_EH}$ and second to $V_{TH\_EL}$. If $V_{FD} < V_{TH\_EL}$ an OFF event is detected, and if $V_{FD} > V_{TH\_EL}$ an ON event is generated. If neither $V_{FD} > V_{TH\_EH}$ or $V_{FD} < V_{TH\_LH}$ are true, then no event is issued. These comparison results are stored in the digital memory circuits shown in Fig. 6.

After the comparisons are made, $PHI2$ is pulled low and shortly after, $PHI1$ is pulsed high. At $PHI1$ frame, $Fn$ is saved on capacitor C1 until $PHI2$ phase of the next frame cycle. Now the $R$ is pulsed high to reset FD and CS and the cycle starts anew.

A case when two sequential frames are taken at distinct illuminance levels, namely high illuminance level and low illuminance level, needs to be taken into account. In low illuminance level the pixel will operate using $S1$ signal while

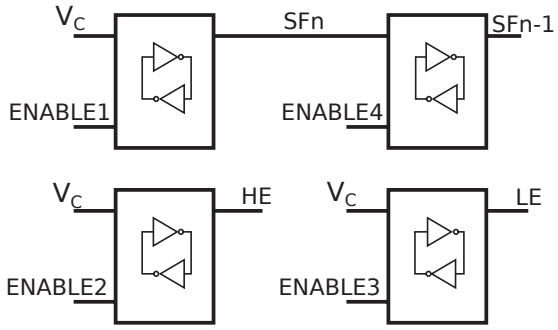


Fig. 6: Digital memory structure to categorize events and their polarity.

TABLE I: Event categorization logic

| Digital memory circuits | | | | Events | |
|---|---|---|---|---|---|
| $D_{EH}$ | $D_{EL}$ | $D_{SFn}$ | $D_{SFn-1}$ | ON | OFF |
| 1 | 0 | 0 | 0 | 1 | 0 |
| 1 | 0 | 1 | 1 | 1 | 0 |
| 0 | 1 | 0 | 0 | 0 | 1 |
| 0 | 1 | 1 | 1 | 0 | 1 |
| 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 1 | 1 | 0 | 0 |
| X | X | 1 | 0 | 1 | 0 |
| X | X | 0 | 1 | 0 | 1 |

in a high-illuminance environment it will switch to $S2$, and due to the non-linearity introduced by difference in sensitivity between signals $S1$ and $S2$ subtracting between them makes no sense as the result would not represent meaningful changes in the physical illumination of the scene.

To resolve the non-linearity issue, the pixel also remembers the type of signal that was used in the previous frame. If two sequential frames do not share the same signal type, the event is automatically declared. If frame $F_{n-1}$ used $S1$ and frame $F_n$ $S2$, an ON event is declared and if frame $F_{n-1}$ used $S2$ and $F_n$ used $S1$ an OFF event is declared.

The digital memory stores all the information needed to determine if there is an event and its polarity. In contrast to previous frame-based event sensors [5], [6] that output 2 bits, 1 to encode if there is an event and the other to encode the polarity of change, our pixel outputs 4 bits. This is to resolve the aforementioned non-linearity issue, which leads to a more complex logic to decide what kind of event occurred. The logic for event categorization can be seen in Table I. A logic circuit to categorize events in-pixel can be implemented but with the cost of increasing the pixel size.

## III. False events due to mismatch

False events will be generated due to mismatch. Fig. 7 shows the analysis of mismatch in false event generation by Monte Carlo simulations. It plots the percentage of false events vs $\Delta V_{TH}$. $\Delta V_{TH}$ is the difference between $V_{TH\_EH} - V_{TH\_EL}$, the thresholds seen in Fig. 5. The lower this difference is the more sensitive the pixel is to the change in the brightness of the scene, but more false events will be generated. For these simulations, the input light power was

TABLE II: Comparison with prior art.

| | This Work | Previous Work [7] | ISSCC 2020 [5] | ISCAS 2020 [3] | VLSI 2019 [4] | Finateu [9] |
|---|---|---|---|---|---|---|
| Process | 180 nm | 180nm | 180nm | 65nm | 65nm | 90nm |
| Resolution | n/a | 64×64 | 64×64 | 1280×960 | 132×104 | 1280×720 |
| Pixel Size | 22×22 $\mu$m$^2$ | 32.305×32.305 $\mu$m$^2$ | 15×15 $\mu$m$^2$ | 4.95×4.95 $\mu$m$^2$ | 10×10 $\mu$m$^2$ | 4.86×4.86 $\mu$m$^2$ |
| Fill Factor | 5% | 2.8% | 21% | 22% | 20% | >77% |
| Supply | 1.8 V/3.3 V | 3.3 V/1.8 V analog 1.8 V digital | 0.8 V | 2.8 V/1.8 V analog 1.0 V digital | 1.2 V | 2.5 V /1.1 V |
| Max event rate/ Frame rate | 1000 efps | 1000 efps | 510 fps | 1.3 Geps | 180 Meps | 1066 Meps |
| Power Consumption | 50-60nW | 505-515 nW | 18.1 nW | 122 nW | 357 nW | 35 nW |
| Readout | Sequential | Sequential | Sequential | Sequential | Sequential | Asynchronous |

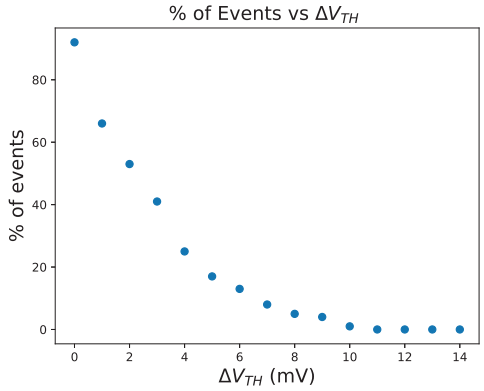


Fig. 7: Mismatch analysis when there is no brightness change in the scene.

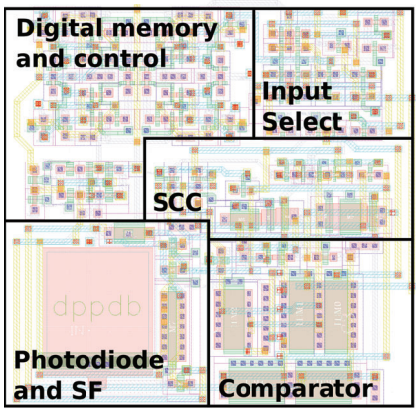


Fig. 8: CDS free pixel layout.

kept constant between sequential frames, thus simulating no change in the scene and ideally, this would never trigger any events. From Fig. 7 we can see that if we want to avoid false event generation due to mismatch in the absence of intensity change, we should set $\Delta V_{TH}$ at least at 11 mV.

## IV. Pixel Data and Comparison with Prior Art

The pixel can generate events at 1000 efps. Pixel pitch is 22 $\times$ 22$\mu$m$^2$. The photodiode area is 5 $\times$ 5$\mu$m$^2$. The pixel layout is shown in Fig. 7 with its main parts highlighted.

Simulations show power consumption in the range of 50-60 nW per pixel depending on the light power. The peak power consumption happens when power input generates $S1$ signal that is similar to the $V_{THSAT}$. Compared to prior art- Table II- redesign still suffers from low fill-factor and larger pixel pitch. However, no other frame differencing event pixels feature dynamic range extension, while speed and power consumption are competitive.

## V. Conclusions and Outlook

The paper has described a CMOS event pixel by frame differencing with dynamic range extension method and no CDS. Post-layout simulations show that power consumption metrics are comparable to those of previous art of ON/OFF event generation through frame differencing. Compared to a previous version this redesign shrinks the pixel pitch to 22 $\mu m$ from 32 $\mu m$ and cuts power consumption. Noise analysis to estimate the dynamic range and subsequent fabrication are future steps to take.

## References


[1] G. Gallego *et al.*, "Event-Based Vision: A Survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 1, pp. 154–180, 2022.

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

[3] Y. Suh *et al.*, "A 1280×960 Dynamic Vision Sensor with a 4.95-$\mu$m Pixel Pitch and Motion Artifact Minimization," in *2020 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2020, pp. 1–5.

[4] C. Li *et al.*, "A 132 by 104 10$\mu$m-Pixel 250$\mu$W 1kefps Dynamic Vision Sensor with Pixel-Parallel Noise and Spatial Redundancy Suppression," in *2019 Symposium on VLSI Circuits*, 2019, pp. C216–C217.

[5] T.-H. Hsu *et al.*, "A 0.8V Multimode Vision Sensor for Motion and Saliency Detection with Ping-Pong PWM Pixel," in *2020 IEEE International Solid- State Circuits Conference - (ISSCC)*, 2020, pp. 110–112.

[6] Y. M. Chi *et al.*, "CMOS Camera With In-Pixel Temporal Change Detection and ADC," *IEEE Journal of Solid-State Circuits*, vol. 42, no. 10, pp. 2187–2196, 2007.

[7] M. Jaklin *et al.*, "HDR 4T-APS Pixel for Event Generation by Frame Differencing," in *2021 IEEE International Midwest Symposium on Circuits and Systems (MWSCAS)*, 2021, pp. 921–924.

[8] N. Akahane *et al.*, "A Sensitivity and Linearity Improvement of a 100 dB Dynamic Range CMOS Image Sensor using a Lateral Overflow Integration Capacitor," in *Digest of Technical Papers. 2005 Symposium on VLSI Circuits, 2005.*, 2005, pp. 62–65.

[9] T. Finateu *et al.*, "5.10 A 1280×720 Back-Illuminated Stacked Temporal Contrast Event-Based Vision Sensor with 4.86µm Pixels, 1.066GEPS Readout, Programmable Event-Rate Controller and Compressive Data-Formatting Pipeline," in *2020 IEEE International Solid- State Circuits Conference - (ISSCC)*, 2020, pp. 112–114.