

Global Shutter CMOS Vision Sensors and Event Cameras for On-Chip Dynamic Information

Marko Jaklin¹, Daniel García-Lesta¹, Paula Lopez¹, and Victor Brea¹

¹Universidade de Santiago de Compostela

January 13, 2023

Abstract

The on-chip extraction of dynamic information from a scene can be addressed with either frame-based CMOS vision, also called smart image sensors, or with dynamic vision sensors, also known as event cameras. When implemented with a pinned photodiode (PPD) as 4-transistor active pixel sensors (4T-APS) the former brings about the benefit of low temporal noise and dark current but without high dynamic range (HDR). The latter comes with the benefits of HDR and fast event detection rate at a low power consumption. The drawback is the background activity noise, which leads to additional hardware or algorithms to keep it low. In essence, the taxonomy of dynamic information extraction with image sensors is that of global shutter solutions and event cameras, each of which with their pros and cons. This paper digs in such differences and similarities focused on mismatch and noise through a global shutter 4T-APS implementation with local HDR incorporated in 180 nm CMOS technology vs conventional logarithmic event sensors found in the literature.

Global Shutter CMOS Vision Sensors and Event Cameras for On-Chip Dynamic Information

Marko Jaklin, D. García-Lesta, P. López, V.M. Brea

*^aUniversidad de Santiago de Compostela, Rúa de Jenaro de la Fuente Domínguez
Santiago de Compostela, 15782, A Coruña, Spain*

Abstract

The on-chip extraction of dynamic information from a scene can be addressed with either frame-based CMOS vision, also called smart image sensors, or with dynamic vision sensors, also known as event cameras. When implemented with a pinned photodiode (PPD) as 4-transistor active pixel sensors (4T-APS) the former brings about the benefit of low temporal noise and dark current but without high dynamic range (HDR). The latter comes with the benefits of HDR and fast event detection rate at a low power consumption. The drawback is the background activity noise, which leads to additional hardware or algorithms to keep it low. In essence, the taxonomy of dynamic information extraction with image sensors is that of global shutter solutions and event cameras, each of which with their pros and cons. This paper digs in such differences and similarities focused on mismatch and noise through a global shutter 4T-APS implementation with local HDR incorporated in 180 nm CMOS technology vs conventional logarithmic event sensors found in the literature.

Keywords: event camera, CMOS vision sensors, smart image sensors, high dynamic range, dynamic vision sensors

1. Introduction

Temporal and spatial redundancy of a static background scene do not call for continuous streaming or for continuous running of elaborate computer vision models as object detection, tracking or action recognition. Indeed, it is the entrance of an object in the field of view of the camera, or, in general, object motion, that triggers further computer vision processing and continuous streaming. Thus, it is clear that there are two main operation modes in video surveillance with different compute and power needs, which can be exploited to extend battery life towards stand-alone always-on visual edge computing nodes [1, 2, 3]. In this context, dynamic information extraction is key to separate both said operation modes.

Frame-based cameras can extract dynamic information through background subtraction or motion detection with smart image sensors implemented as either 3T- or 4T-APS sensors [4, 1, 5, 6]. Dynamic vision sensors or event cameras do it through time-stamped spikes of ON and OFF events that detect increasing and decreasing intensity changes in the scene above a certain threshold [7, 8, 9, 10]. In both models it is possible to provide both a 1.5 bit code signal array accounting for ON and OFF events, as well as a constant background, or the raw image itself.

In summary, the taxonomy of on-chip dynamic information extraction from a scene is that of global shutter solutions and event cameras, each of which with their pros and cons. This paper digs in such differences and similarities focused on mismatch and noise through a 4T-APS implementation with local HDR incorporated in 180 nm CMOS technology vs conventional logarithmic event sensors found in the literature.

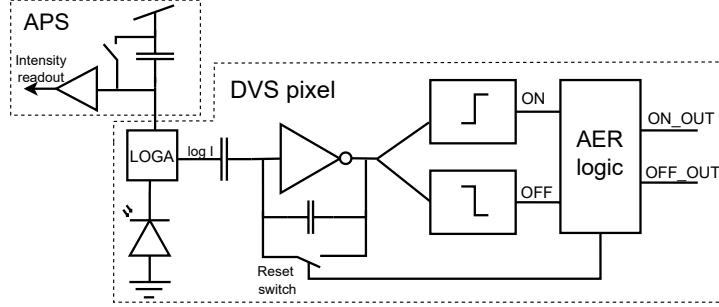


Figure 1: DAVIS solution to the dynamic vision sensor concept.

2. Global Shutter and Event Pixels for Dynamic Information Extraction

Dedicated global shutter pixels working in integration mode with frame differencing functionality are another approach to generate events[6][11]. In this case, frame differencing is performed through two consecutive frames with a fixed integration time and a global reset phase in between. Global reset keeps false event generation through leakage currents at bay, although at the cost of worse temporal resolution than that of event pixels. As a benefit, the pinned photodiode (PPD) in 4T-APS allows for lower dark current and temporal noise with correlated double sampling (CDS) techniques. Apart from their worse temporal resolution, global shutter pixels with frame differencing features lack HDR capability. In order to fill this gap, we have implemented a global shutter HDR 4T-APS pixel for event generation through frame differencing to be compared with event pixels

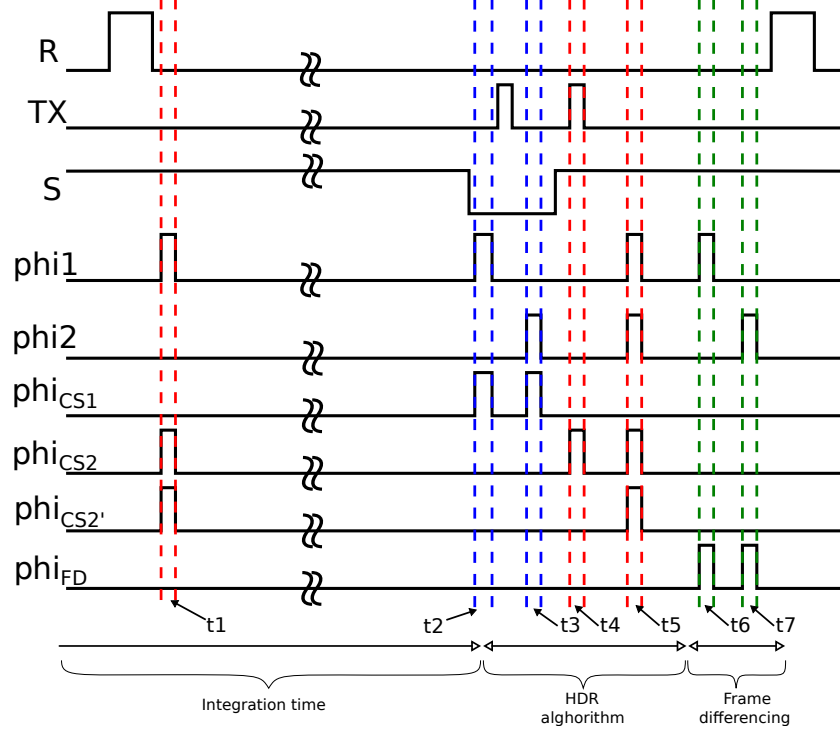


Figure 3: Timing diagram of our HDR 4T-APS for event generation.

store more electrons at $FD + CS$, extending the dynamic range at the price of smaller conversion gain.

The operation of our global shutter HDR 4T-APS pixel in more detail is as follows. First, right after reset at $t1$ in Fig. 3 and with signal S pulsed high, the noise level $N2$ of the node $FD + CS$ is read and stored on the capacitor C_{S2} driven by the source follower (SF). During the integration time signal S is pulsed high, keeping $M3$ transistor active and allowing for any saturated electrons that flow through $M1$ when TX is low at this time to be stored on the overflow capacitor CS . Just before integration time ends, signal S is pulled low, isolating FD from CS and immediately after, at $t2$, the noise level

$N1$ associated with the electrons of signal $S1$ of the now isolated FD node is read and stored on C_{S1} . Next, the switch TX is pulsed high, transferring the electrons from the PPD to the FD node, and thus, generating $N1 + S1$, and sending it to C_{S1} . The CDS is performed with the arrival of $N1 + S1$ at $t3$ by setting signal $\phi2$ high in the subtraction unit, yielding $S1$. $S1$ signal represents the voltage value generated by non saturated electrons only, while signal $S2$ which is calculated shortly after $S1$, represents the value generated by non saturated and saturated electrons. The $S1$ - $S2$ crossing is given by a user defined threshold $V_{THS1/S2}$. If the accumulated voltage $S1$ exceeds said threshold, the analog memory stores signal $S2$, otherwise it stores $S1$.

Frame differencing is performed with subtraction and comparator circuits, which are also used for the decision making on the switch from $S1$ to $S2$. We apply circuit sharing techniques, as said circuits are the same for ON/OFF event generation, and for CDS operations to mitigate the effect of mismatch and noise levels $N1$ and $N2$ corresponding to signals $S1$ and $S2$, respectively.

3.2. Pixel Circuits

All our circuits feature power gating in order to decrease power consumption. The supply voltage of the 4T-APS is $V_{dd} = 3.3$ V, searching for a wide dynamic range. This voltage supply is set to 1.8 V for the rest of the circuitry, aiming at low power consumption.

3.2.1. 4T-APS

Our 4T-APS sensing structure with a PPD comprises a programmable overflow MIM capacitor CS set to 50, 100 and 150 fF for the HDR algorithm, which allows for different upper limits in the incident light. The aspect

ratios W/L ($\mu\text{m}/\mu\text{m}$) in the transistors of our 4T-APS implementation are: 5.4/0.8, 0.35/0.35 and 0.35/0.5 for $M1$, $M2$ and $M3$, respectively. The source follower is biased with a current $I_{SFAPS} = 0.5 \mu\text{A}$ with an aspect ratio of 0.22/0.9 ($\mu\text{m}/\mu\text{m}$), while the biasing transistor aspect ratio is 0.22/1.5, again in ($\mu\text{m}/\mu\text{m}$). The capacitance of the floating diffusion node FD has been estimated by post-layout simulations as 12 fF.

3.2.2. Subtraction Unit

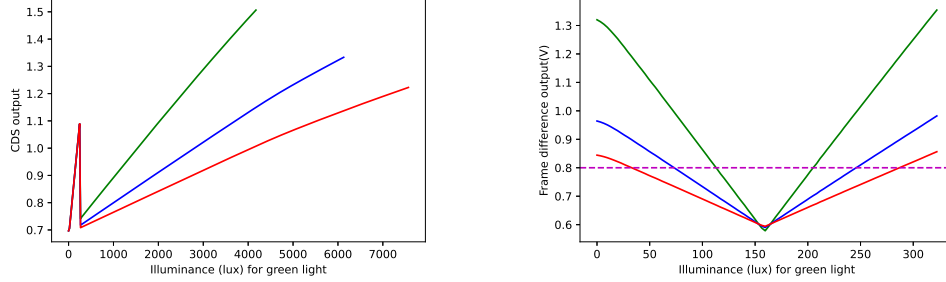
The subtraction unit is a double cascode inverting amplifier in feedback mode with the capacitor C_2 and the reset switches ϕ_{i1} and ϕ_{i2} . This unit runs CDS for signals $S1$ and $S2$ and frame differencing between the current F_n and the previous frame F_{n-1} . This structure is the same as that of the dynamic vision sensor [9] or other CMOS vision sensor solutions like [14].

The sensitivity of the frame differencing and $S1/S2$ signals is given by the $C_{FD}/C_{S1}/C_{S2}$ to C_2 ratios. We have set C_2 as a programmable device to $1\times$, $2\times$ and $3\times$ gains. Capacitors $C_{FD}/C_{S1}/C_{S2}$ have been sized to 90 fF, while C_2 is scaled accordingly. Signal V_3 adds a user-defined offset, which in our case sets the comparator input transistors in saturation. All the switches have been designed as NMOS transistors with minimum dimensions. The double cascode inverting amplifier has an open-loop gain of 60 dB.

The sequence of operations of the different switches can be seen in Fig. 3. The CDS operation for signals $S1$ and $S2$ to mitigate the effect of noise $N1$ and $N2$ is given by the formula:

$$Si = \frac{C_{Si}}{C_2}(Ni - (Ni + Si)) + V_3 \quad (1)$$

with subindex i referring to either $S1/S2$ or $N1/N2$. This result is stored



(a) HDR extension up to the linear limit of S2. At this point N1 is saturated resulting in nonlinearity and false threshold $V_{THS1/S2}$ detection.

(b) Frame difference gain programability.

Figure 4: Simulations showing HDR extension and gain programability.

in the analog memory bank, which holds the value of the current frame F_n and the previous one F_{n-1} .

The programmability of the overflow capacitor CS offers flexibility in the HDR mode. Fig. 4a collects a simulation for monochromatic light at wavelength of 550 nm (green) for the three gains defined in our pixel. The Y-axis is the signal V_S in Fig. 2, after running CDS. The pixel operates only with the floating diffusion capacitance, and thus with high conversion gain (signal $S1$), for low illumination levels. At a given illumination level, the pixel enters the HDR region, with the excess of electrons being collected on the overflow capacitor CS . This is the region of signal $S2$, where the conversion gain decreases due to the capacitance of two shunted nodes, $FD + CS$. The upper the curve in the HDR region, the lower the CS capacitance. The three curves in Fig. 4 correspond to our three cases of CS ; 50, 100 and 150 fF.

The frame differencing is calculated as:

$$V_S = \frac{C_{FD}}{C_2} |F_n - F_{n-1}| + V_3 \quad (2)$$

This value is compared to a user-defined threshold voltage $V_{THevent}$ to make a decision on whether or not there is an event. We have implemented the absolute difference operation in Eq. (2) in order not to yield negative voltages, and to have only one comparator instead of two dedicated comparators for ON and OFF events, as is the case of the classical dynamic vision sensor[10]. The operation of our subtraction unit obliges to store on C_{FD} the highest voltage of the current F_n and the previous frame F_{n-1} . This is carried out with the comparator labeled *CMP* in Fig. 2. This comparison can also be used as polarity flag p_k , latching this result onto the digital memory block of our pixel (see Fig. 2). In terms of the timing diagram of Fig. 3, when ϕ_{i1} is pulsed high at $t6$, the higher value frame arrives on C_{FD} . The arrival of the second frame with ϕ_{i2} pulsed high completes the operation of frame differencing. Fig. 4b conveys simulations showing the absolute value of the frame differencing operation for our three different gains. According to Eq. (2), higher slopes come from lower capacitance values of C_2 . The dashed line represents an example of the user-defined threshold $V_{THevent}$ to trigger an event.

3.2.3. Analog memory bank

The analog memory bank of Fig. 2 stores the previous F_{n-1} and current F_n frames. It is implemented as an open-loop sample and hold configuration. Open-loop sample and hold architectures are up to the challenge of keeping the image for hundreds of ms with an acceptable accuracy degradation due

to long term storage losses in more demanding solutions for on-chip dynamic information extraction [15].

3.2.4. *Comparator for Event Generation and S1/S2 Crossing with its Input Logic*

The decision making on when to switch from conventional integration mode with signal $S1$ to HDR extension through signal $S2$ and on whether or not there is an event is carried out by the comparator labeled CMP in Fig. 2. This comparator takes the difference given by the subtraction unit V_S , i.e., either the CDS output of $S1$ or the frame differencing value, or signal $S2$ as input $IN1$, and a user programmable threshold $V_{THS1/S2}$ or $V_{THevent}$ for the $S1$ - $S2$ crossing or the event generation, respectively as input $IN2$. The block labeled *Input Logic* in Fig. 2 sets the appropriate input at the right time instant. Such an input logic is implemented with a bank of switches realized with NMOS transistors of minimum dimensions.

This comparator has been implemented as a two-stage open-loop amplifier with a 5 NMOS OTA architecture with a differential input driving an inverter. Noise and mismatch can cause incorrect polarity when two frames are close, so, although it does not feature offset cancellation, it has been designed with large transistors in order to make it mismatch resilient.

3.2.5. *Digital Memory Block: Event Generation and Read-Out*

The digital memory block shown in Fig. 5 contains 4 D-latches to: i) provide events, ii) set the polarity of the event, and iii) assess the signal uniformity of $S1$ or $S2$ between frames.

The behavior of the 4 D-latches is conveyed in Fig. 5. If D_e is set to

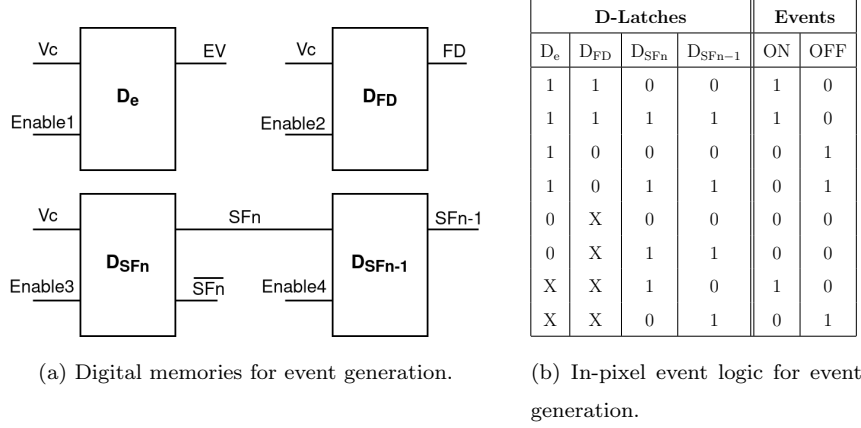


Figure 5: Digital logic structure and logic.

logical '1' and D_{SF_n} and $D_{SF_{n-1}}$ have the same logical values, either '0' or '1', an event is issued. In this case, the local value at D_{FD} determines the polarity of the event. A logical '1' means that the voltage variation along the integration time at the current frame exceeds that of the previous frame, providing an ON event, and vice versa. The fact of two consecutive frames coming from two different sensitivities ($S1$ and $S2$) is accounted for with D_{SF_n} and $D_{SF_{n-1}}$ issuing different logical states. In this situation, there is an ON event if the state from D_{SF_n} is a logical '1' and that from $D_{SF_{n-1}}$ is a logical '0' regardless of D_e and D_{FD} . There is an OFF event if the state from D_{SF_n} is a logical '0' and that from $D_{SF_{n-1}}$ is a logical '1' regardless of the states of D_e and D_{FD} . D_{SF_n} and $D_{SF_{n-1}}$ holding the same logical values means that the change from the previous to the current frame generates events.

Finally, our chip provides both raw images and events. This is managed by the "Output select" block in Fig. 2. The raw image is read out with an unity gain buffer as an analog signal, as the result of the CDS operation.

The events are sent directly from the pixel to the outside of the chip. Event readout speed is estimated at around 1000 efps by post-layout simulations.

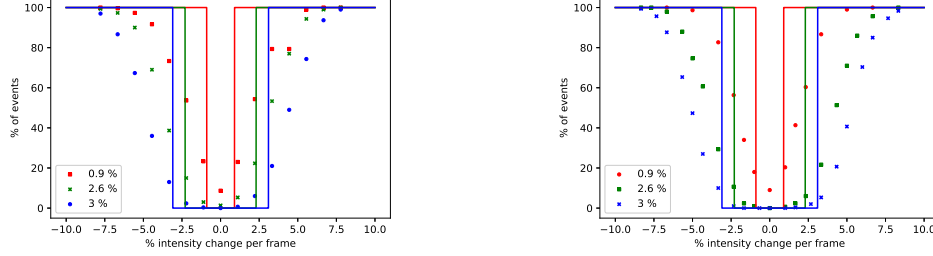
4. Performance Metrics

4.1. Spatial Accuracy

The spatial uniformity of a pixel array in our global shutter HDR 4T-APS is given by the mismatch of every pixel along the data path. Our solution comprises the two integration modes $S1$ and $S2$, so that we have run Monte Carlo simulations for both cases in order to assess the sensitivity to intensity changes of our approach.

Fig. 6a and Fig. 6b collect the effect of mismatch on event generation for $S1$ and $S2$. The X-axis is the intensity change in percentage between two consecutive frames, while the Y axis shows the percentage of ON and OFF events from Monte Carlo simulations for different user defined threshold voltages, namely, 7, 17, and 23 mV, which correspond to percentage changes in the light intensity of 0.9, 2.6 and 3%. An ideal scenario is that of a sudden jump from 0% to 100% of events for a given threshold, shown with continuous lines in Fig. 6a and Fig. 6b. Mismatch and temporal noise in actual circuits cause to have a minimum threshold to generate events.

We have run Monte Carlo simulations to emulate a whole array of global shutter HDR 4T-APS pixels for event generation. Every pixel in the plots of Fig. 6a and Fig. 6b is a percentage of 300 nominal Monte Carlo simulations, that we have thought of as percentage of pixels in an array yielding events. The percentage % of intensity change per frame along the X-axis is calculated with the formula $(P1 - P2)/FSO_{Si}$, where $P1$ and $P2$ mean the light input

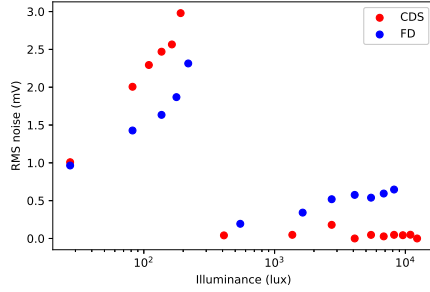


(a) Intensity change sensitivity of signal $S1$, i.e., conventional operation mode.

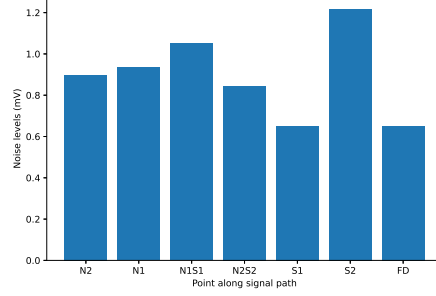
(b) Intensity change sensitivity of signal $S2$, i.e., HDR mode.

Figure 6: False event simulations due to mismatch effects in our HDR 4T-APS solution.

power of two consecutive frames, and FSO_{S_i} is the full-scale output of either signal $S1$ or $S2$. The results expressed as percentage of events in Fig. 6a and Fig. 6b are similar to one another, but it should be taken into account that the power of signals $S1$ and $S2$ differ. The light input power is kept constant during the first frame ($P1$) at 5 pW for $S1$ signal (low illumination) and 200 pW for $S2$ signal (high illumination), while it is subject to increasing and decreasing variations during the second frame to generate the percentage of illumination changes along the X-axis. The input light is simulated to be a green monochromatic light with wavelength $\lambda = 555$ nm, making the conversion from photometric to radiometric units more straightforward. The integration time has been set to 1 ms. Finally, Fig. 6a and 6b show that the 100% of correct cases is only met with around of 5% of intensity change for $S1$ (conventional mode) and $S2$ (HDR mode) signals, as can be seen by the red points. Temporal noise adds more inaccuracies. As apparent, the application dictates the user-defined threshold voltage $V_{THevent}$.



(a) RMS of total noise (circuit and photon shot noise).



(b) RMS of temporal noise along the signal path.

Figure 7: Noise analysis of our HDR 4T-APS pixel for event generation.

4.2. Temporal Noise

The temporal noise has to be added to the spatial noise to determine the noise floor in our implementation in order to calculate the dynamic range of our HDR 4T-APS pixel for frame differencing.

4.2.1. Circuit Noise

The effect of thermal noise in our implementation has been obtained by averaging 10,000 nominal transient noise simulations. $N1$ and $N1S1$ ($N1 + S1$ in Fig. 3) are noise samples taken on the FD node of our circuit (Fig. 2) at their corresponding time instants of the timing diagram of Fig. 3, while $N2$ and $N2S2$ ($N2 + S2$ in Fig. 3) are taken on the $FD + CS$ node. The subtraction of $N1S1$ and $N2S2$ from $N1$, $N2$ runs the CDS operation.

4.2.2. Photon Shot Noise

Photon shot noise has also been added for a given input light, with the number of photons as $N_p = \frac{PT_e}{E_f}$, where P is the light input power with a given wavelength, in our case $\lambda = 555$ nm for an easy conversion from radiometric

to photometric units, T_e is the integration time, and E_f is the energy of a single photon. From the number of photons the Poisson distribution has been derived and added to our CAD simulator.

4.2.3. Total Noise

Fig. 7a shows root mean squared (RMS) values of photon shot and circuit noise on the output node (V_S) of the subtraction circuit of Fig. 2, with and without CDS for a sudden transition from low illumination, where photon shot noise dominates, and the global shutter HDR 4T-APS works in the normal region with signal $S1$, to high illumination, where circuit noise prevails, and the pixel works within the HDR extension, with a lower conversion gain with signal $S2$. Circuit simulations do not account for parasitic light sensitivity with effects of leakage currents caused by the reset transistor of the subtraction unit [16]. Nevertheless, we do not expect a large impact because the global shutter HDR 4T-APS works with reset between two consecutive frames, where integration times are usually in the order of ms, and, thus, short for the leakage currents to have a significant impact [4]. Fig. 7b shows the temporal noise along the data path of our HDR 4T-APS for event generation. The noise floor is estimated to be 0.5 mV_{rms} .

4.3. Dynamic Range

The estimation of dynamic range (DR) is given by 3. $E_{v_{sat2}}$ is the upper limit of $S2$ and after this limit signal $N1$ is saturated resulting in nonlinearity and false threshold $V_{THS1/S2}$ detection. $E_{v_{floornoise}}$ is the lowest illuminance that can be detected.

$$DR = 20\log\left(\frac{E_{v_{sat2}}}{E_{v_{floornoise}}}\right) \quad (3)$$

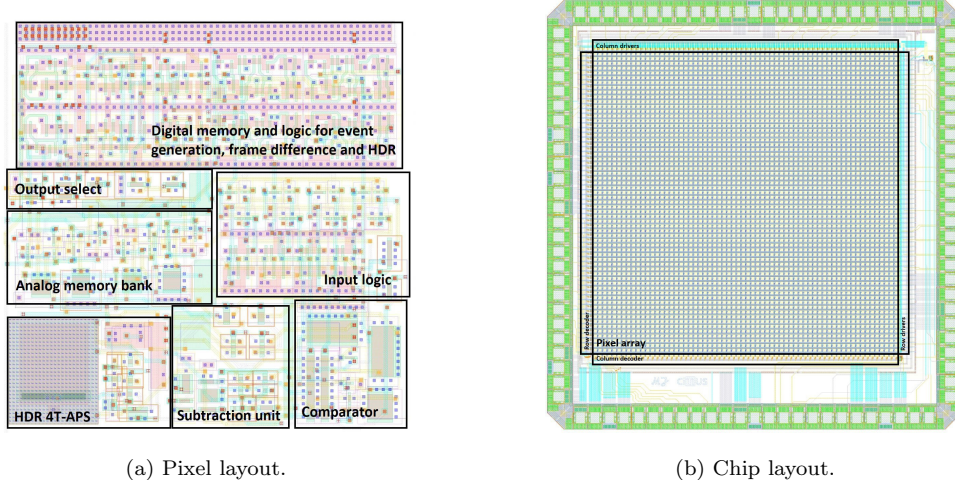


Figure 8: Pixel and chip layouts of our HDR 4T-APS for event generation.

The HDR extension rises the dynamic range from 53 dB up to 85 dB, which is lower than that of the original HDR pixel with overflow capacitor [13], which achieves 100 dB. This is due to two factors, the higher number of active elements in our circuit to generate events, which increases the noise floor, and the lower power supply voltage, 3.3 V vs 5 V in [13].

4.4. Chip Data and Comparison with Prior Art

We have laid down a 64×64 pixel array in 180 nm CMOS technology. Our pixel is able to run at 1000 events frames per second (efps), measured as post-layout simulations. Our pixel pitch is $32.3 \mu\text{m}$. The photodiode size is $5.4 \times 5.4 \mu\text{m}^2$. The layout of the pixel with all its individual blocks labeled can be seen in Fig. 8a. The layout of the complete chip is displayed on Fig. 8b. The area of the chip is $2.81 \times 2.90 \text{ mm}^2$. The array is surrounded by row decoders on the left, row drivers on the right, column drivers on the top and select circuits on the bottom side of the array.

Table 1: Chip Data and Comparison with Prior Art.

	This Work	ISSCC 2020 ^a [6]	ISCAS 2020[8]	VLSI 2019[9]	ISSCC 2020 ^b [7]
Process	180 nm	180 nm	65 nm	65 nm	90 nm
Resolution	64×64	64×64	1280×960	132×104	1280×720
Pixel Size	$32.305 \times 32.305 \mu m^2$	$15 \times 15 \mu m^2$	$4.95 \times 4.95 \mu m^2$	$10 \mu m^2 \times 10 \mu m^2$	$4.86 \times 4.86 \mu m^2$
Fill Factor	2.8%	21%	22%	20%	>77%
Supply	3.3 V/1.8 V analog 1.8 V digital	0.8V	2.8 V/1.8 V analog 1.0 V digital	1.2 V	2.5/1.1 V
Max event rate/ Frame rate	1000 efps	510 fps	1.3 Geps	180 Meps	1066 Meps
Power Consumption	500-550 nW	18.1 nW	122 nW	357 nW	35 nW
Dynamic range	85 dB	64.2 dB	-	-	>124 dB
Readout	Sequential	Sequential	Sequential	Sequential	Asynchronous

Post-layout simulations show a higher power consumption in the range of 500-550 nW per pixel depending on the incident light power. Our excess power consumption comes mainly from the in-pixel HDR algorithm. For instance, the peak power consumption happens when power input generates $S1$ signal that is similar to the $V_{THS1/S2}$. Compared to prior art- Table 1- our design suffers from low fill-factor and an overhead pixel pitch caused mainly by the in-pixel HDR algorithm to deal with signals $S1$ and $S2$.

5. Outlook and Conclusion

This paper has delved into the two main CMOS options to run on-chip dynamic information from a scene, namely, global shutter pixels and dynamic vision sensors. We provide post-layout simulations of a global shutter 4T-APS with HDR extension for event generation by means of a local algorithm on an in-pixel overflow capacitor in order to cut the HDR edge of dynamic vision sensors on global shutter pixels. Our HDR mode extends the dynamic

range from 53 to 85 dB, which although can be improved, it is still far from the state-of-the-art value of 124 dB of dynamic vision sensors. Additional circuits in global shutter pixels for HDR extension and event generation hamper fill-factor and increase noise floor. Power consumption of global shutter pixels for event generation and dynamic vision sensors are in the same range. The edge of global shutter pixels on dynamic vision sensors comes from meaningless background activity noise due to the reset between frames, which avoids the use of dedicated circuits to this end or additional processing stages.

6. Acknowledgements

This work has received funding from the projects: European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 860370, PID2021-128009OB-C32; and the European Union (European Regional Development Fund): 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.

References

- [1] X. Zhong *et al.*, “A Fully Dynamic Multi-Mode CMOS Vision Sensor With Mixed-Signal Cooperative Motion Sensing and Object Segmentation for Adaptive Edge Computing,” *IEEE Journal of Solid-State Circuits*, vol. 55, no. 6, pp. 1684–1697, 2020.
- [2] J. Choi *et al.*, “Always-On CMOS Image Sensor for Mobile and Wearable

- Devices,” *IEEE Journal of Solid-State Circuits*, vol. 51, no. 1, pp. 130–140, 2016.
- [3] I. Park *et al.*, “A 640×640 Fully Dynamic CMOS Image Sensor for Always-On Operation,” *IEEE Journal of Solid-State Circuits*, vol. 55, no. 4, pp. 898–907, 2020.
- [4] D. García-Lesta *et al.*, “CMOS vision sensor for background subtraction,” in *IEEE International Symposium on Circuits and Systems*, 2020.
- [5] M. Benetti *et al.*, “A Low-Power Vision System With Adaptive Background Subtraction and Image Segmentation for Unusual Event Detection,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 65, no. 11, pp. 3842–3853, 2018.
- [6] 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.
- [7] T. Finateu *et al.*, “5.10 A 1280×720 Back-Illuminated Stacked Temporal Contrast Event-Based Vision Sensor with $4.86\mu\text{m}$ 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.
- [8] Y. Suh *et al.*, “A 1280×960 Dynamic Vision Sensor with a $4.95\text{-}\mu\text{m}$ Pixel Pitch and Motion Artifact Minimization,” in *2020 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2020, pp. 1–5.

- [9] C. Li *et al.*, “A 132 by 104 10 μ m-Pixel 250 μ W 1kefps Dynamic Vision Sensor with Pixel-Parallel Noise and Spatial Redundancy Suppression,” in *2019 Symposium on VLSI Circuits*, 2019, pp. C216–C217.
- [10] P. Lichtsteiner, C. Posch, and T. Delbruck, “A 128 \times 128 120 dB 15 μ s Latency Asynchronous Temporal Contrast Vision Sensor,” *IEEE Journal of Solid-State Circuits*, vol. 43, no. 2, pp. 566–576, 2008.
- [11] 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.
- [12] 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.
- [13] 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.
- [14] M. Suárez *et al.*, “Low-Power CMOS Vision Sensor for Gaussian Pyramid Extraction,” *IEEE Journal of Solid-State Circuits*, vol. 52, no. 2, pp. 483–495, 2017.
- [15] D. García-Lesta *et al.*, “In-pixel Analog Memories for a Pixel-Based Background Subtraction Algorithm on CMOS Vision Sensors,” *International Journal of Circuit Theory and Applications*, pp. 1631–1647, 2018.

- [16] Y. Nozaki and T. Delbruck, “Temperature and Parasitic Photocurrent Effects in Dynamic Vision Sensors,” *IEEE Transactions on Electron Devices*, vol. 64, no. 8, pp. 3239–3245, 2017.