

# Digital Pixel Based Event Vision Sensor with Simultaneous Event and Intensity Outputs

Tianye Wen, Jian Lu, Caihong Liu, and Xiaoguang Liu

In this letter, we report the design of a low-power event-based vision sensor (EVS) digital pixel in a 3D-stacked 45-nm backside illumination CMOS process. The design uses in-pixel analog-to-digital converters and memory to store digitized light intensities. By using a bit comparator to compare digital intensity values of the same pixel in the time domain, the high-fidelity event stream and intensity information from the same pixel can be generated simultaneously. The bit comparator can eliminate the false ON events in traditional DVS and achieve an equal processing rate for ON/OFF events. The power consumption of the proposed digital EVS pixel can be as low as 10 nW/pixel. The EVS sensor can achieve a maximum event rate of 260 Meps.

**Introduction:** Cmos image sensors (CIS) are widely used in a variety of applications, including consumer electronics, surveillance, industry, and scientific studies. Active pixel based CIS (APS) is the mainstream imaging technology due to its merits in high resolution and low temporal spatial noise. However, it suffers from issues such as insufficient dynamic range, relatively low frame rate, and data redundancy. The emergence of EVS[1] solves the drawbacks of traditional APS-type sensors. EVS generates less data, has higher temporal resolution, lower power consumption and wide dynamic range by asynchronously sampling and outputting only the relevant information on local pixel brightness changes.

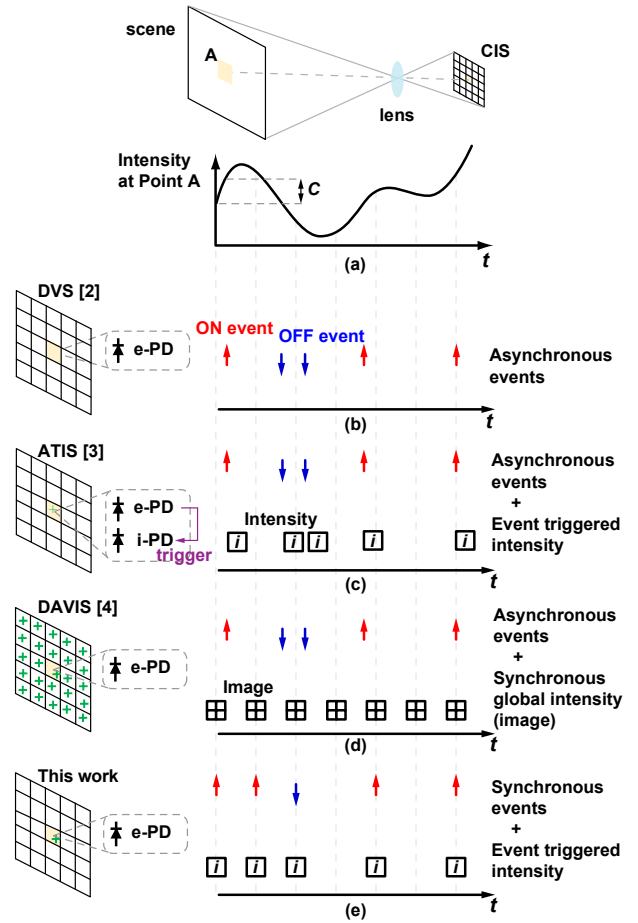
Dynamic vision sensors (DVS)[2] outputs only the event stream information of light intensity change above a certain threshold  $\Delta$ . With the continuous scaling of CMOS technology, the conventional sampling circuit of DVS is susceptible to false events caused by leakage current. As shown in Fig. 2 (a), Reset pmos disconnects after pixel reset, causing  $V_1$  to float until an event occurs in the pixel[5]. The leakage current ( $i_{leak}$ ) charges the  $V_1$  node. As a result,  $V_2$  keeps drifting downward. Even in the absence of any illumination changes, the pixel still periodically generates false ON events. Although algorithms can be used to filter out periodic ON events, they increase the load on post-processing and the overall power consumption of the system. In order to obtain a certain event threshold, there is a large difference between the bias currents of ON and OFF comparators. The low bias current  $i_{off}$  of the OFF comparator leads to a serious delay. DVS can solve spatial vision problems such as three-dimensional reconstruction, simultaneous localization and mapping and pose estimation. However, its application is limited to tasks involving object recognition and scene understanding due to its inherent inability to provide pixel intensity information.

As shown in Fig. 1, the Asynchronous Time-event Imaging Sensor (ATIS)[3] outputs both the event stream information and the changed pixel intensity information simultaneously by using a DVS based on events and an exposure measurement unit based on pulse width modulation. Intensity measurement is taken after an event is triggered, resulting in potential time mismatch and data loss in slow motion scenarios. ATIS uses two photodiodes per pixel, which impedes the miniaturization of the pixel area.

Dynamic and active pixel sensor (DAVIS)[4] provides both event information and intensity information of each pixel by combining DVS and APS modes of imaging. DAVIS still retains the drawbacks of traditional APS sensors. In high-speed motion scenarios, the accuracy of DAVIS can be affected by the prolonged data processing delay of APS.

This letter presents a digital pixel-based EVS architecture. Digital pixel (DP)[9] has the advantages of low power consumption, high speed and high dynamic range compared to analog pixel. Due to a larger number of circuit elements per pixel, DPs suffer from a lower fill factor (FF) compared to analog pixels. With the development of 3D stacking technologies, the FF of DP has significantly increased.

Shown in Fig. 3, the proposed EVS architecture achieves temporally-matched event stream information and intensity information read-out using a single photodiode. The design is carried out in a 3D stacked 45/55-nm CMOS CIS process. The top wafers contains the photodiode array, and the bottom logic wafer includes an logarithmic front-end, an in-pixel ADC and memory to convert and store the light intensity information at each pixel. A digital bit comparator compares the light intensity of the current moment with that of the previous moment to determine whether an event should be generated. The threshold for event detection is stored in the digital logics.



**Fig. 1** The architecture and output data types of various EVS. (a) Realistic scene. The absence of any symbol in the realistic scene means there is no change in light intensity. (b) DVS[2] (c) ATIS[3] (d) DAVIS[4] (e) This work. Nothing in the pixel means no output data.

The digital nature of the event generation prevents data loss in low-speed motion scenes. If an event is generated, the sensor also outputs the light intensity at the pixel, facilitating easier image processing later on.

**Analog front-end:** Fig. 4 shows the analog front-end circuits of the proposed pixel design. The photodiode current is converted to a photovoltage by an NMOS transistor ( $M_2$ ) working in the sub-threshold region[7]. The common source amplifier composed of  $M_1$  and  $M_3$  forms a negative feedback loop to clamp the voltage at point A at a fixed voltage, thus speeding up the AFE circuit.

In this work, SS-ADC is used within the pixel circuit to convert the photovoltage to a digital value at the pixel level. SS-ADC is known for compact size, low power consumption, and relatively good linearity [8]. As shown in Fig. 4, the comparator flips when the ramp voltage  $V_{ramp}$  cross the pixel voltage  $V_{pix}$ . The seven-bit Gray code  $gc[6:0]$  corresponding to the photovoltage is stored in the current latch when the enable signal is triggered.

**Digital comparison circuit:** As shown in Fig. 4, the output data  $d_{out1}[6:0]$  from the current registers represents the photovoltage at the current time. The photovoltage data  $d_{out2}[6:0]$  at the previous time is stored in the reference registers. The bit selection signal  $rdbit[6:0]$  selects one bit from  $d_{out1}[6:0]$  and  $d_{out2}[6:0]$  for comparison at each clock cycle. Specifically, the  $i$ -th bit of  $d_{out1}[i]$  and  $d_{out2}[i]$  are compared when selected. If the comparison output  $diff=1$ , it indicates that  $d_{out1}[i]$  is different from  $d_{out2}[i]$ , i.e. either  $d_{out1}[i] = 1, d_{out2}[i] = 0$  or  $d_{out1}[i] = 0, d_{out2}[i] = 1$ . The odd-even parity check is then performed by XOR1 to determine which of the two cases is true. Here, the initial value of  $parity$  is 0. When  $parity = 0$  and  $d_{out1}[i] = 1, ab = 1$ . The  $diff=1$  and  $ab = 1$  signals jointly determine that  $d_{out1}[i] = 1$  and  $d_{out2}[i] = 0$ .  $M_e$  and  $M_d$  are conducting when  $diff = 1$  and  $ab = 1$  are satisfied, and now  $gt = 0$ . Eight SRAM latches (SR) complete the tracking and transmission of the different signals under the control of two different phases  $phi1$  and  $phi2$ .

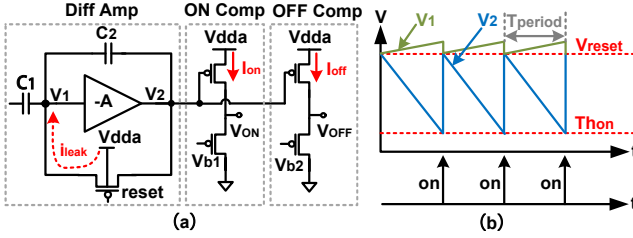


Fig. 2 (a) Architecture of traditional DVS[5]. (b) The timing diagram of periodic false event generation induced by voltage drift.

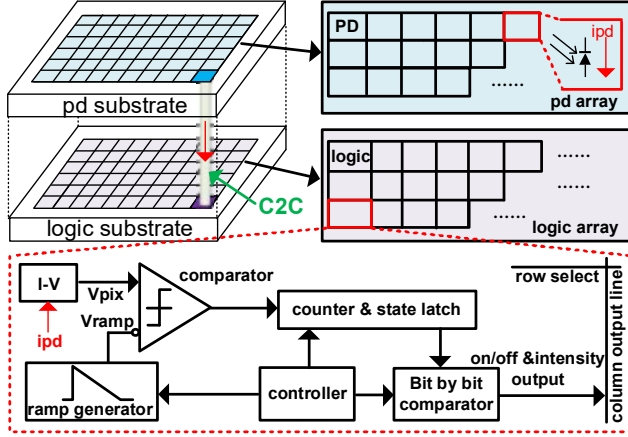


Fig. 3. General architecture of the proposed digital EVS pixel.

The MOS transistors  $M_a$  to  $M_g$  act as switches. The output signal *greater* is asserted when *different*=1 and *great*=1. Completing the comparison and transfer of signals  $d_{out1}[i]$  and  $d_{out2}[i]$  requires two cycles of  $\phi_{i1}$  and  $\phi_{i2}$ . Events will not be triggered when  $d_{out1}[0]$  and  $d_{out2}[0]$  are different because there is only one cycle of  $\phi_{i1}$  and  $\phi_{i2}$ . The threshold for events is determined by the number of SRs in this design. To lower the event trigger threshold, an increase in the corresponding number of SRs is required. After the occurrence of the event, the result of performing the "AND" operation between *evt* and *up\_data* is used as the enable signal for the reference register. The data from the current register is then updated into the reference register.

**Output logic:** The output signals *greater* and *event* of the SR latches control  $M_4$ ,  $M_5$ , and  $M_6$ . The bit controller can enable or disable each pixel depending on the value of *cfg\_din*. When the *row\_sel* and *pixel\_enal* are both equal to 1, the *evt\_bl* is at a low voltage. If *evt*, *grter*, *row\_sel*, and *pixel\_enable* are all equal to 1,  $M_5$ – $M_8$  are turned on and the *on\_bl* is at low voltage. The *on\_bl* and *evt\_bl* passes through an OR gate to obtain the ON signal. The inverse of *on\_bl* and *evt\_bl* passes through another OR gate to obtain the OFF signal. The inverse of the *event* (*evt*) serves as the enable signals of the latch, and the  $d_{out2}[6:0]$  output is generated. The  $d_{out}[6:0]$  is the intensity value of the pixel at this moment.

**Operation of control logic:** Fig. 5 shows the timing diagram of a single pixel performing the A/D conversion and comparison circuit, with each signal corresponding to the nodes labeled in Fig. 4. The whole timing diagram can be divided into two stages: A/D conversion period and digital pixel comparison period. The signals *clr*, *rdbit[6:0]*,  $\phi_{i1}$ ,  $\phi_{i2}$  and *up\_data* only contribute to digital pixel comparison period. The *clr* is transmitted to each SRAM latch for signal clearing. Bit selector *rdbit[6:0]* starts from MSB *rdbit[6]*. There is a  $\phi_{i1}$  and  $\phi_{i2}$  pulse signal during each bit *rdbit[i]* comparison period to implement data tracking. The *up\_data* sequence emitted by pluse generator is operated with the *event* to update the data of the reference latch. *unrd* and *rst* are always at a low level during the comparison period of *rdbit[6:0]*. *unrd* rises to a high level to prevent data leakage after the comparison is completed.

**Simulation results:** Fig. 6 shows the simulated (post-layout) internal and output signals of the digital pixel for varying  $i_{pd}$  values. Here,  $C = 5\%$  (equivalent to 2 mV in the voltage domain) is the threshold for triggering an event. When the two signals  $V_{pix}$  and  $V_{ramp}$  cross each other, *cmp\_out*

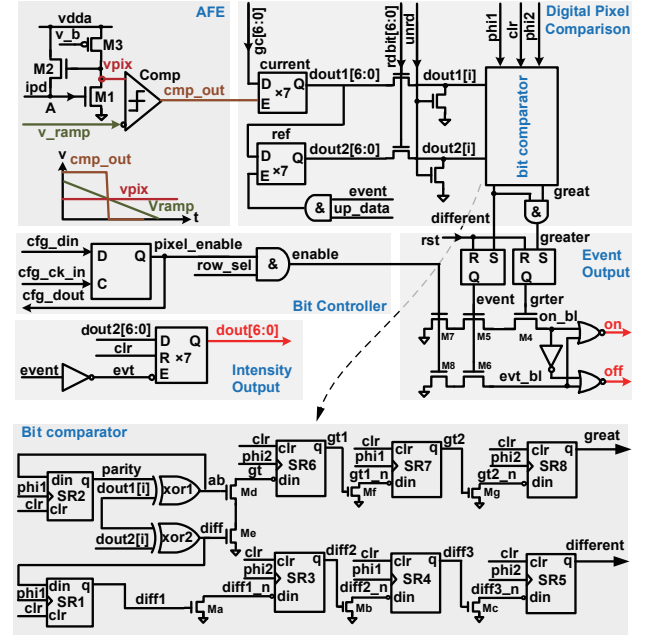


Fig. 4. Schematic of the proposed digital EVS pixel.

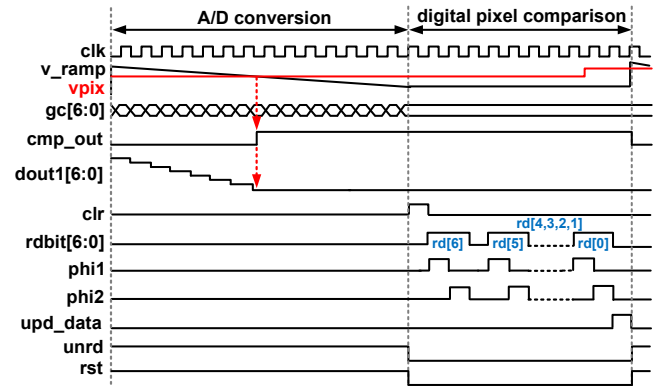


Fig. 5. Pixel timing diagram and operation.

switches from high (1.2 V) to low (0 V). The current light intensity value  $gc[6:0]$  is written into the pixel memory after the A/D conversion. The stored intensity value from the previous conversion is read out from the pixel memory and compared with the current intensity value to determine whether an event should be generated. For example, at time  $t_2$ , the pixel senses a  $V_{pix}$  of 610 mV which is one threshold above the  $V_{pix}$  of 608 mV at time  $t_1$ . The pixel generates an ON event and output the corresponding intensity value of 28. When there is no event output, e.g. at time  $t_5$ , no intensity output is generated.

Fig. 7 shows the simulated (post-layout) AFE output voltage  $V_{pix}$  with respect to the input photocurrent  $i_{pd}$ , the AFE exhibits a dynamic range of 120 dB. Fig. 8 shows the top and bottom layout of the BSI 3D-stacked digital EVS pixel.

Fig. 8 shows the top and bottom layout of the BSI 3D-stacked digital EVS pixel. Table 1 shows the main performance results of the EVS presented in this letter compared with other advanced EVS.

**Conclusion:** This letter presents a digital pixel based EVS sensor that can output both ON/OFF events and digitized light intensity simultaneously. The EVS incorporates in-pixel SS-ADC, memory and bit comparators. It eliminates spurious toggle events in traditional distributed switches, enabling equal processing rates for on and off events. The power consumption is as low as 10 nW/pixel. We believe that our developed EVS holds great prospects for applications in artificial intelligence imaging, such as object vision and scene understanding.

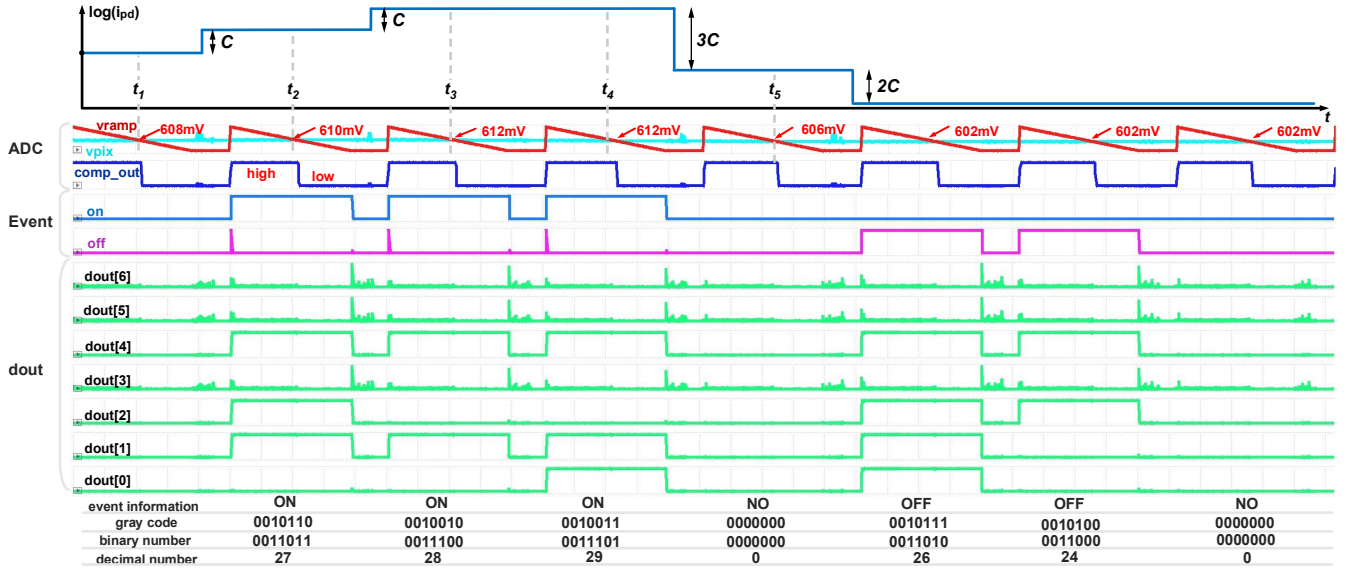


Fig. 6. The output data of a single pixel with respect to varying input current  $i_{pd}$ .

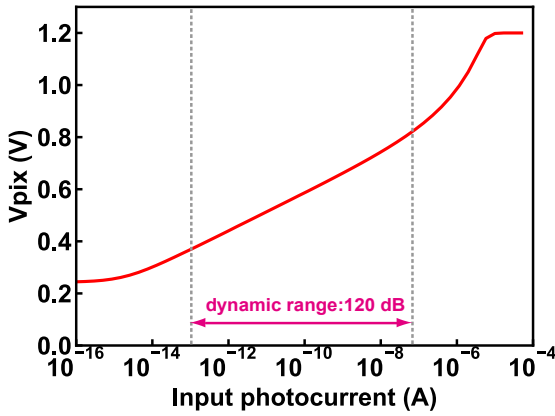


Fig. 7 Simulated AFE output voltage  $V_{pix}$  with respect to the input photocurrent, showing a  $> 120$  dB dynamic range.

Table 1: Comparison with the state-of-the-art

Metrics	This Work	2019[2]	2017[5]	2015[6]
CMOS process	45/55-nm	65-nm	90-nm	180-nm
	BSI 3D	FSI	BSI	FSI
$V_{DD}$ [V]	1.2	1.2	1.2/1.8	1.8
Pixel size [ $\mu m^2$ ]	$15 \times 15$	$10 \times 10$	$8 \times 8.5$	$31.2 \times 31.2$
Chip size [ $mm^2$ ]	$4 \times 4$	$2 \times 2$	$9 \times 9$	$3.2 \times 1.6$
Dynamic range [dB]	120	-	80	130
Fill factor [%]	100	20	-	10.3
Power/pixel [nW]	10	18	160	400
Event threshold [%]	5	-	9	1
Max event rate [Meps]	260	180	300	-

**Acknowledgment:** This work is supported by the project grant from the Natural Science Foundation of Shenzhen City, No.JCYJ20220818100408018.

Tianye Wen, Jian Lu, Caihong Liu, and Xiaoguang Liu (*Southern University of Science and Technology (SUSTech)*)

E-mail: liuxg@sustech.edu.cn

## References

- 1 T. Serrano-Gotarredona and B. Linares-Barranco, "A  $128 \times 128$  1.5% Contrast Sensitivity 0.9% FPN 3  $\mu s$  Latency 4 mW Asynchronous Frame-Free Dynamic Vision Sensor Using Transimpedance Preamplifiers," in *IEEE Journal of Solid-State Circuits*, vol. 48, no. 3, pp. 827-838, March 2013, doi: 10.1109/JSSC.2012.2230553.

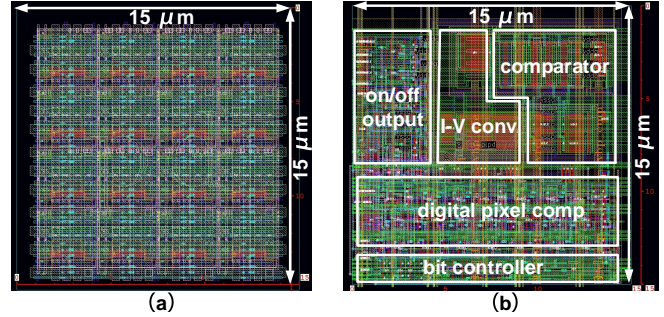


Fig. 8 Layout of the pixel: (a) top layout (photodiode), (b) bottom layer (AFE & digital logic).

- 2 C. Li, L. Longinotti, F. Corradi and T. Delbruck, "A 132 by 104  $10 \mu m$  Pixel 250  $\mu$  1kefps Dynamic Vision Sensor with Pixel-Parallel Noise and Spatial Redundancy Suppression," 2019 Symposium on VLSI Circuits, Kyoto, Japan, 2019, pp. C216-C217, doi: 10.23919/VLSIC.2019.8778050.
- 3 C. Posch, D. Matolin and R. Wohlgenannt, "A QVGA 143dB dynamic range asynchronous address-event PWM dynamic image sensor with lossless pixel-level video compression," 2010 IEEE International Solid-State Circuits Conference - (ISSCC), San Francisco, CA, USA, 2010, pp. 400-401, doi: 10.1109/ISSCC.2010.5433973.
- 4 C. Brandli, R. Berner, M. Yang, S. -C. Liu and T. Delbruck, "A  $240 \times 180$  130 dB 3  $\mu s$  Latency Global Shutter Spatiotemporal Vision Sensor," in *IEEE Journal of Solid-State Circuits*, vol. 49, no. 10, pp. 2333-2341, Oct. 2014, doi: 10.1109/JSSC.2014.2342715.
- 5 B. Son et al., "4.1 A  $640 \times 480$  dynamic vision sensor with a  $9 \mu m$  pixel and 300Meps address-event representation," 2017 IEEE International Solid-State Circuits Conference (ISSCC), San Francisco, CA, USA, 2017, pp. 66-67, doi: 10.1109/ISSCC.2017.7870263.
- 6 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.
- 7 P. Lichtsteiner, T. Delbruck and J. Kramer, "Improved ON/OFF temporally differentiating address-event imager," *Proceedings of the 2004 11th IEEE International Conference on Electronics, Circuits and Systems*, 2004. ICECS 2004., Tel Aviv, Israel, 2004, pp. 211-214, doi: 10.1109/ICECS.2004.1399652.
- 8 M. -W. Seo et al., "2.45 e-RMS Low-Random-Noise, 598.5 mW Low-Power, and 1.2 kfps High-Speed 2-Mp Global Shutter CMOS Image Sensor With Pixel-Level ADC and Memory," in *IEEE Journal of Solid-State Circuits*, vol. 57, no. 4, pp. 1125-1137, April 2022, doi: 10.1109/JSSC.2022.3142436.
- 9 El Gamal A, Yang D X D, Fowler B A. Pixel-level processing: why, what, and how?[C]//Sensors, Cameras, and Applications for Digital Photography. SPIE, 1999, 3650: 2-13.