

Current Starved Ring Oscillator Image Sensor

Devin Atkin, Orly Yadid-Pecht

Abstract—The continual demands for increasing resolution and dynamic range in complimentary metal-oxide semiconductor (CMOS) image sensors have resulted in exponential increases in the amount of data that need to be read out of an image sensor, and existing readouts cannot keep up with this demand. Interesting approaches such as sparse and burst readouts have been proposed and show promise, but at considerable trade-offs in other specifications. To this end, we have begun designing and evaluating various readout topologies centered around an attempt to parallelize the sensor readout. In this paper, we have designed, simulated, and started testing a light-controlled oscillator topology with dual column and row readouts. We expect the parallel readout structure to offer greater speed and alleviate the trade-off typical in this topology, where slow pixels present a major framerate bottleneck.

Keywords—CMOS image sensors, high-speed capture, wide dynamic range, light controlled oscillator.

I. INTRODUCTION

THERE is an ongoing push for increasing frame rate video cameras. The current industry standard for high-speed cameras, the Phantom branded cameras, can capture very high framerates, but only by sacrificing their resolution to the point of being unusable for many applications. For example, the Phantom v2640 (see [1] for the datasheet) can capture at rates up to 303,460 frames per second (fps); however, this can only be captured at a maximum resolution of 1792W x 8H. This extreme aspect ratio is a product of the sensor's readout method. It makes it challenging to utilize these cameras in many practical applications. If we want to target a more practical 640W x 480H (Video Graphics Array (VGA) resolution), the achievable framerate drops to 53,290 fps. This is because of the standard trade-off that exists between the image sensor's specifications of framerate, resolution, dynamic range, and noise. As the captured framerate increases, the other specifications tend to shrink such that the amount of data being captured remains roughly constant. This trade-off has remained throughout time and technologies, with older film-based technologies being limited by the sensitivity (ISO) of the film. A high ISO film allowed for higher captured frame rates, but also resulted in a substantial increase to the amount of film grain, the analog medium's equivalent to fixed pattern noise.

Most of the development regarding CMOS image sensors centers around modifying them to improve one of the four main steps in the sensor's operation cycle:

- 1) Reset: The voltage on the photodiode node is set to some known voltage. Most often, this would be the sensor's supply voltage.

- 2) Integration: Photocurrents are incredibly small. To detect them effectively, they will typically need to be integrated for some time, often by providing time for the photocurrent to drain the photodiode node.
- 3) Sample: Once the integration has reached a level by which the signal can be measured with sufficient accuracy, the integration is stopped so that the pixel can be read out.
- 4) Readout: The pixel signal needs to be read and converted into a usable format. Typically, the voltage on the photodiode is put through an analog-to-digital conversion and read out of the chip.

The majority of image sensors will operate with this overall loop; however, the specifics of its implementation vary infinitely. The two primary readout strategies are global and rolling shutter sensors. In rolling shutter sensors, pixels operate with a time offset from one another by row; this means that some pixels are reset, others are integrated, and others are read out at any given time. All pixels operate in unison in global shutters outside of the readout phase. Rolling shutters attempt to better utilize the readout hardware by having it active more of the time, but in doing so, they may introduce tearing as quick-moving objects move across the frame. This rubber-pencil-like effect can render the sensor useless in specific applications. Other topologies remove the reset or integration times, and can be read out at any time during the sensor operation [2]-[4]. Regardless, the readout will remain the primary bottleneck to higher framerates as the number of paths out of the chip are fundamentally finite. Substantial work in sparse and burst readout methods attempt to sidestep this issue. Sparse readout methods primarily fall into the category of address event representation (AER) image sensors [5]-[7]. These sensors only read out pixels where they have detected an event, thus substantially limiting the number of pixels that need to be read out. Some promising burst imaging sensors utilize in-situ memory to capture successive frames rapidly before reading them out of the sensor at a slower speed. The best of these sensors can achieve framerates ranging from 5 million to 1.25 billion during their burst recording [8], [9]. Some of these topologies have begun utilizing more advanced technologies to 3D stack the silicon so that digital memory and readout can be stored behind the photodiode array [10]. This technique helps maintain the pixel's fill factor when adding a large amount of digital logic to the sensor array.

II. PROPOSED TOPOLOGY

Fig. 1 shows the proposed pixel topology. The topology consists of a current starved ring oscillator where the pixel

Devin Atkin and Orly Yadid-Pecht are with Electrical and Computer Engineering, University of Calgary, I2Sense Laboratory, Calgary, Canada (e-mail: dmatkin@ucalgary.ca, orly.yadid-pecht@ucalgary.ca).

photocurrent sets the frequency of operation. This topology is similar in concept to previously presented light controlled oscillators [3], [4], [11]; however, other implementations have operated by counting pulses generated by the sensor. This is inherently slow and may become effectively unreadable under low light conditions. The topologies have two primary differences. First, pixels in our work act as a reset for a pulse width counter instead of being used to count the frequency. This difference means that instead of measuring the number of pulses to determine the light level, we need to only measure the length of time for a single pulse to occur which can then be interpreted directly. We have also introduced a secondary readout path which is not present in the other light controlled oscillators. This strategy allows us to read slow pixels without waiting, as one readout can handle fast pixels while the other can handle slow ones through appropriate software control. This design was completed in the AMS 350nm optical process, which includes an NWELL photodiode; this necessitates the anode of the reverse biased diode to be connected directly to the ground to prevent substrate shorting. In a deep NWELL process, this may be simplified by removing the additional current mirror.

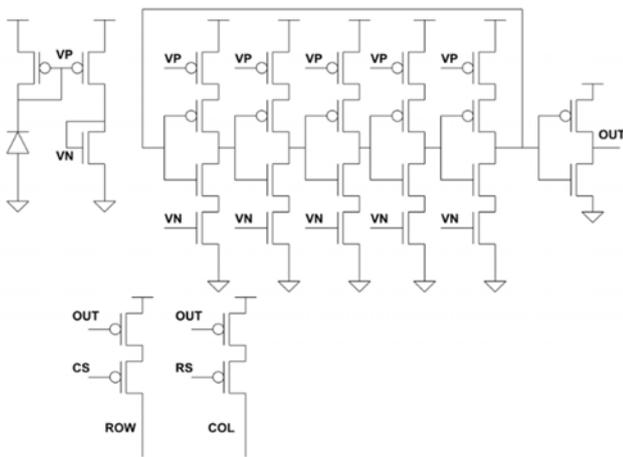


Fig. 1 Current Starved Ring Oscillator Pixel

One considerable advantage of this topology is the adjustability of the readout to serve different purposes. While the initial prototype transistor sizing was kept very simple, a lot of opportunities exist to either adjust transistor sizing or make simple additions to optimize the performance further, thus opening an avenue for development. Two potential development avenues would be creating a ratio between the current mirror transistors to adjust the ring oscillator curve to run faster, or adding a settable bias voltage to allow the pixels to be calibrated. These were dismissed for the initial prototype to keep the design as simple as possible.

AER sensors where pixels generate events saying which pixel has been triggered may also operate using row and column select outputs [5], [12]. Pulses simultaneously on two buses can determine which pixel generated a specific output at any given time. Those pulses can then determine the light level hitting any given pixel. We are not employing this address-based

representation because it loses the frame-based readout, which is useful and expected by most final applications. This readout requires some form of order arbitration circuitry to determine which pixels generated the event as pulses may overlap. At higher speeds, the propagation delays can become quite tricky to decide accurately on the order of pixel outputs. Although some of the AER sensors, which include arbitration systems, will also have requests and acknowledges passed between the pixel and the driver circuitry to help with this problem [6], this is a double-edged sword that prevents fast pixels blocking slower pixel's ability to read out at the cost of increasing the design complexity. In addition, this method results in increased silicon area required and operation complexity which can further contribute to the negative side of this readout method. Instead, we utilize our secondary output to account for low-light pixels by allowing us to address multiple pixels simultaneously; slower pixels can be read through one bus, while faster pixels can be read through the other.

A. Pulse Width Counter

The second element of the design is the pulse width counter. This element was somewhat of a novelty for our lab as our previous Wide Dynamic Range (WDR) works have all relied on an external analog to digital converter to produce output values. As this design creates a pulse frequency modulated signal, we needed to adapt to utilize this topology effectively and thus included a pulse width counter internal to the sensor. This component consists of an incrementing register fed by an external clock, reset by the pixel output, and latched to an output register. This setup allows the sensor to simultaneously count the time it takes for a given pixel to pulse for an entire row or column. The level of degradation between light levels is determined by the speed of the input clock and is limited primarily by the length of the included counting register.

$$f_{min} = \frac{1}{t_{max}} \quad (1)$$

$$f_{max} = \frac{1}{t_{min}} \quad (2)$$

$$t_{dist} = \frac{1}{c_{HS}} \quad (3)$$

$$t_{mdist} = (t_{dist} * 2^{reg}) \quad (4)$$

Equations (1)-(4) are the equations that determine the sensor's performance. Equation (1) represents the pixel's dark current frequency (the frequency when the pixel is not exposed to light), this should be relatively large; if it is small, the required register size to measure it becomes excessive. Equation (2) is the maximum frequency output of the pixels in each scene; as shown in Fig. 2, it can be extremely large relative to the dark current frequency, providing the sensor with a great inherent dynamic range. Equation (3) is the shortest distinguishable time frame for a given input clock and can be adjusted in software as needed. Finally, (4) is the maximum distinguishable time, a function of the register length and the speed of the input clock. The register length is inherently set at

the design time and places constraints on the discernable contrast between light levels.

$$\text{bits obtainable} = \log_2((t_{max} - t_{min}) / t_{dist}) \quad (5)$$

Equation (5) is the maximum number of effective bits obtainable for a given clock speed if the counter register does

not overflow. If the counter does overflow, this indicates that the t_{dist} is too small given the counter register size, and the input clock needs to be reduced to accommodate the scene dynamic range. However, doing this will limit the available dynamic range as the effective bit count will be less than the output bit count size.

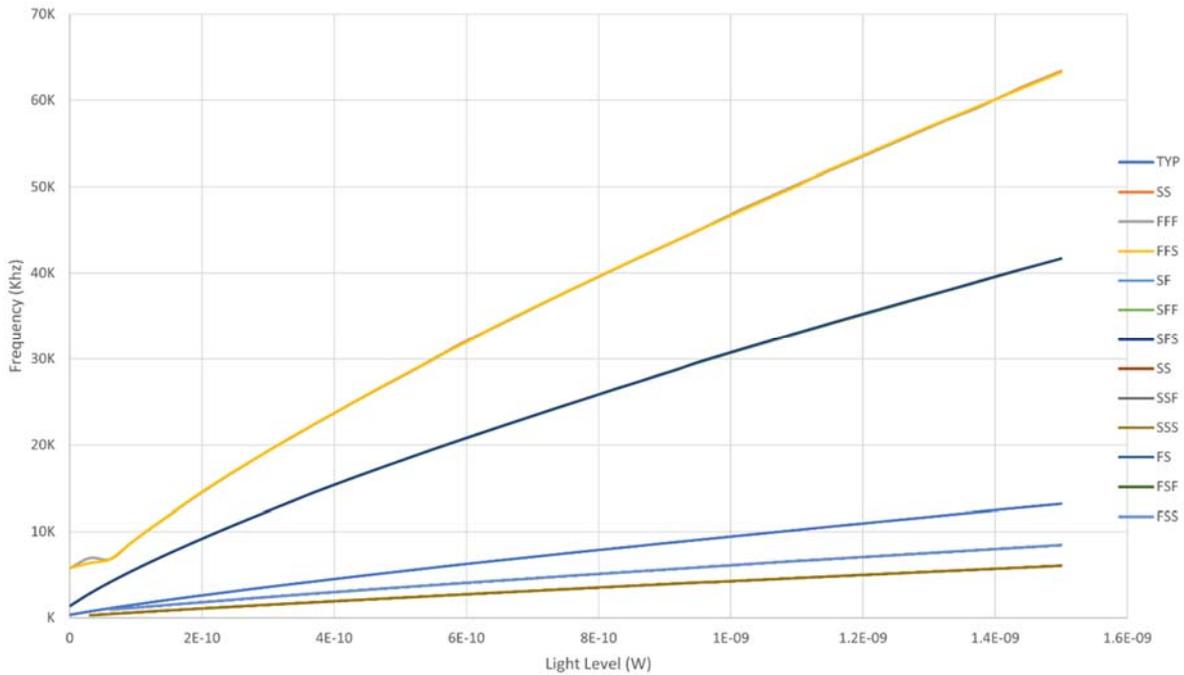


Fig. 2 Single Pixel Corner Simulation Output

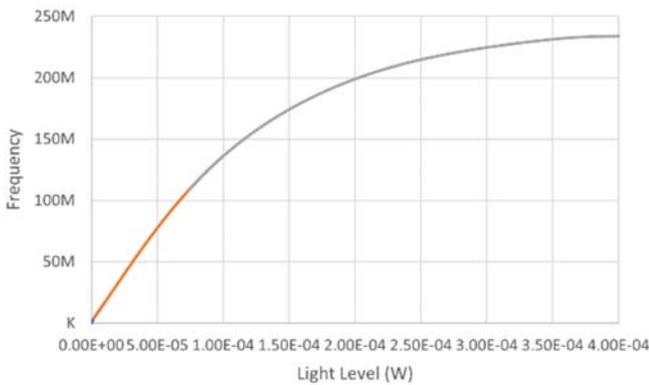


Fig. 3 Single Pixel Simulated Output

III. SIMULATION RESULTS

The output was simulated across different light levels on all provided silicon corners to confirm that the pixel would function as expected. The photodiode simulation input is a voltage level that represents the number of watts on the photodiode area at the optimum illumination wavelength of 850 nm. Based on this, the input power was swept between 0 W and 1.5 nW to determine the frequency behavior. Additional simulations were run up to 300 μW with the typical corner to predict the expected range of frequencies fully. The pixel did

not saturate above this; however, the frequency response flattened substantially with minimal frequency change above 224.6 MHz regardless of increases to the light input. Given a maximum operating frequency of 224.6 MHz at the 300 μW input power and the dark current operating frequency of 290.5 Hz we can estimate a dynamic range of roughly 271 dB using (6).

Verilog A was used to simulate varying light levels (0 W – 1.5 nW) on individual pixels in a 64 x 64 pixel array, which was modeled in Verilog A to approximate the pixel behavior. The output from this simulation was fed into MATLAB to verify the array's performance. This simulation allowed an image to be recovered using the outputs and then simulated recovered outputs based on the digital output chain. Fig. 4 shows this output with the ideal recovery based on the simulation data. Using a $f_{max} = 12.97 \text{ KHz}$, $f_{min} = 290.5 \text{ Hz}$, and a $C_{HS} = 13 \text{ KHz}$, if we set our maximum distinguished time to the minimum time frequency such that, $t_{mdist} = t_{max}$ we can use (4) to calculate the needed output register size of 6 bits. If we accommodate all corners, we can redo the calculations with $f_{max} = 41.32 \text{ KHz}$, $f_{min} = 116.2 \text{ Hz}$, and a $C_{HS} = 42 \text{ KHz}$. Using the same equation, we come to 9 bits required to capture the entire frequency range. The number of bits that the output can effectively show is 7, according to (5). Simulations and the actual results of the fabricated chip show

that additional bits are required to get the needed contrast. This mistake occurs because of how the natural light levels vary compared to the calculations. More minor changes in light levels need to be distinguished to produce the appropriate contrast in natural scenes.

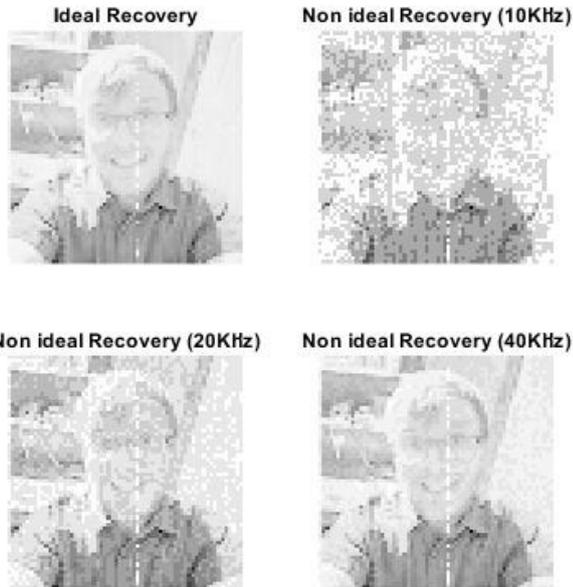


Fig. 4 Simulated Image Recovery

IV. MEASURED OUTPUTS

The chip described above was manufactured using the AMS 350 nm Opto-Process and has undergone preliminary testing to determine the overall viability of the topology. During the testing, several minor silicon errors were discovered with the digital output, which significantly complicated the output reading and limited the output image's performance. However, pictures from the prototype chip show promise for future chip developments despite these issues. To frame the following discussion, the readout operation is as follows:

$$DR \approx 20 \log \left(\frac{f_{max}}{f_{min}} \right) \quad (6)$$

- 1) Reset the chip to clear the column and row select registers and the output registers. This step was intended only to be required at the beginning of operation after startup. However, in the prototype sensor, this step is necessary after each row is read out to prevent pixels from ghosting between rows. This error is due to a mistake present in the digital logic and will be corrected in future silicon.
- 2) Load the column and row select registers. Loading these two registers sets the X and Y of the column and row readouts. The other coordinate is selected with the output 640:10 output mux.
- 3) Start clocking the input at a set frequency. This step may be done at the beginning of operation or adjusted on the fly to try and maximize the scene contrast. This clock is fed into the pulse width counter to determine the frequency of

- 4) Adjust the output mux to read out each pixel in each row/column successively. While one coordinate of the pixel being read out is set by the Column and Row registers, the other coordinate is set by the 640:10 mux connected to the output of each column. A more complicated readout algorithm should account for the dual row/column outputs to include both slow and fast outputs; however, for preliminary testing, the second output was treated primarily as a backup.
- 5) Read output values. Interpret these values as the varying light levels. The output value is the number of times the input clock has cycled since the last time the pixel has reset itself. The output register was too short and frequently overflows multiple times before a reset.

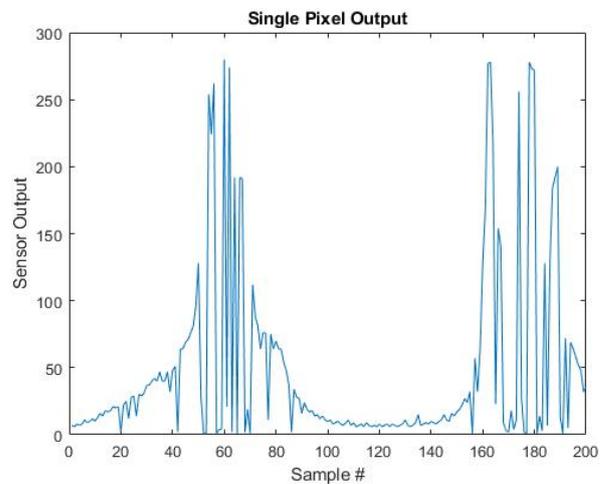


Fig. 5 Single Pixel Output

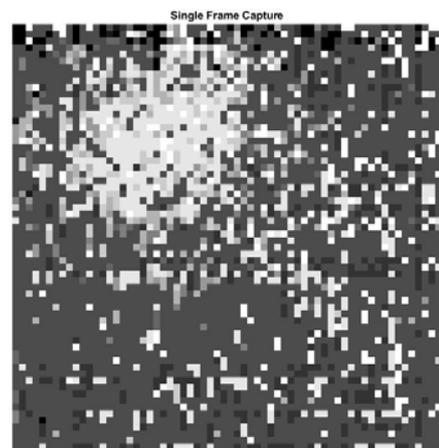


Fig. 6 Image of Lamp Light

There are a series of improvements beyond the most basic output structure. Given this basic setup, it can take quite some time to read slower pixels, with them often not resetting within the time taken to read any given row. Improvements to the design would include either increasing the dark current

frequency to greater than the maximum allowable pixel time or adding additional circuitry to adjust the curve for a low light mode.

Fig. 5 shows a single pixel output over multiple samples given a constant 10 KHz input clock, with a lamp being waved in front of the sensor. The graph shows that the readout value changes depending on the light level and how the pixel frequency compares to the input clock frequency. This test revealed the first design flaw in the fabricated silicon; the output does not have a value updated signal to indicate a change. This design oversight means that if the pixel output is constant, the change in value cannot be adequately seen. Typically, the output fluctuates by 1-bit depending on how the reset lines up with the clock frequency; however, making multiple reads of a single pixel to more accurately determine its frequency is challenging. A resettable new value flag would solve this issue quickly. This test also revealed that our initial light estimates were inadequate to gain sufficient contrast; the 10-bit bus selected resulted in an overly noisy picture with insufficient bit depth in the dynamic range. Future revisions of this chip will therefore include a substantially larger pulse width counter to increase the degree to which differing light levels can be distinguished.

Fig. 6 shows an image captured of a full frame of the sensor. The output from the sensor was fed into MATLAB using a prototype board and then assembled to give the captured image. The image is very low contrast in large part due to insufficient register sizes making differing light levels challenging to distinguish. For context, the lighter-colored blob in the frame is a lamp held above the sensor.

V. CONCLUSION

In this paper, we have presented an image sensor for achieving a combination of high frame rate and high dynamic range. This initial prototype has a few minor silicon bugs discussed above, which will be fixed in a future revision of the chip. Overall, this topology has potential to improve, which will be explored in future works.

Future work will include producing a version with an increased output register to solve the contrast problem and yield more easily recognizable images. The second improvement for future iterations is the digital readout chain, which will include a more standardized interface capable of higher speed with additional flags to simplify readout. In addition, this change will allow a more thorough simulation-based test before fabrication. Finally, we will add other controls to handle the available dynamic range effectively and additional output signals to help coordinate the dual readouts cleanly. The long-term goal for this design is to integrate earlier WDR work undertaken by the I2Sense laboratory to achieve a high-speed readout alongside an easily processable frame-based image sensor.

ACKNOWLEDGMENT

The completion of this research was made possible thanks to the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Alberta Informatics Circle of

Research Excellence (iCORE)/Alberta Innovates - Technology Futures (AITF).

REFERENCES

- [1] Vision Research and A. M. A. Division, "Phantom v2640 Datasheet and Product Page." <https://www.phantomhighspeed.com/products/cameras/ultrahighspeed/v2640> (accessed Dec. 18, 2019).
- [2] N. Ricquier and B. Diericks, "Pixel structure with logarithmic response for intelligent and flexible imager architectures," IEEE Conference Publication, 1992. <https://ieeexplore.ieee.org/document/5435374> (accessed Dec. 19, 2019).
- [3] F. Taghibakhsh and K. S. Karim, "Light controlled oscillators; pixel architecture for large area linear digital imaging using amorphous silicon," Canadian Conference on Electrical and Computer Engineering, pp. 1279–1281, 2008, doi: 10.1109/CCECE.2008.4564745.
- [4] W. Yang, "Wide-dynamic-range, low-power photosensor array," in Digest of Technical Papers - IEEE International Solid-State Circuits Conference, 1994, pp. 230–231. doi: 10.1109/isscc.1994.344657.
- [5] H. Zhu, M. Zhang, Y. Suo, T. D. Tran, and J. van der Spiegel, "Design of a Digital Address-Event Triggered Compressive Acquisition Image Sensor," IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 63, no. 2, pp. 191–199, Feb. 2016, doi: 10.1109/TCSI.2015.2512719.
- [6] E. Culurciello, R. Etienne-Cummings, and K. Boahen, "Arbitrated address event representation digital image sensor," Dig Tech Pap IEEE Int Solid State Circuits Conf, pp. 92–93, 2001, doi: 10.1109/ISSCC.2001.912560.
- [7] 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," Dig Tech Pap IEEE Int Solid State Circuits Conf, vol. 53, pp. 400–401, 2010, doi: 10.1109/ISSCC.2010.5433973.
- [8] M. M. El-Desouki, O. Marinov, M. J. Deen, and Q. Fang, "CMOS active-pixel sensor with in-situ memory for ultrahigh-speed imaging," IEEE Sens J, vol. 11, no. 6, pp. 1375–1379, 2011, doi: 10.1109/JSEN.2010.2089447.
- [9] L. Millet et al., "A 5 Million Frames per Second 3D Stacked Image Sensor with In-Pixel Digital Storage," in ESSCIRC 2018 - IEEE 44th European Solid State Circuits Conference, Oct. 2018, pp. 330–333. doi: 10.1109/ESSCIRC.2018.8494287.
- [10] P. Vivet et al., "Advanced 3D Technologies and Architectures for 3D Smart Image Sensors," Proceedings of the 2019 Design, Automation and Test in Europe Conference and Exhibition, DATE 2019, pp. 674–679, May 2019, doi: 10.23919/DATE.2019.8714886.
- [11] O. de Gaetano Ariel and D. F. Martin, "Light controlled oscillator for image sensing," 2018 2nd Conference on PhD Research in Microelectronics and Electronics Latin America, PRIME-LA 2018, pp. 1–3, May 2018, doi: 10.1109/PRIME-LA.2018.8370385.
- [12] A. M. Haas, S. L. Williams, M. H. Cohen, and P. A. Abshire, "Dark address event representation imager," Midwest Symposium on Circuits and Systems, vol. 2005, pp. 388–391, 2005, doi: 10.1109/MWSCAS.2005.1594119.