

# A high dynamic range linear vision sensor with event asynchronous and frame-based synchronous operation

Juan A. Leñero-Bardallo, Ricardo Carmona-Galán, and Ángel Rodríguez-Vázquez; Institute of Microelectronics of Seville, CSIC-Universidad de Sevilla, Sevilla 41004, Spain (e-mails: [juanle@imse-cnm.csic.es](mailto:juanle@imse-cnm.csic.es); [rcarmona@imse-cnm.csic.es](mailto:rcarmona@imse-cnm.csic.es); [angel@imse-cnm.csic.es](mailto:angel@imse-cnm.csic.es)).

## Abstract

We present a novel High-Dynamic-Range (HDR) image sensor with linear output. Photogenerated charge is continuously integrated at every pixel without saturating. Each time the photodiode voltage reaches a programmable threshold, the pixel resets and starts over integrating charge again. With an event-based approach, it is possible to count the number of times (if any) that a pixel has saturated during exposure. Pixel illumination is represented with a 20-bit word. The most significant 12b represent the number of times that a pixel has saturated during exposure. The least significant 8b are the result of an analog-to-digital conversion in the end of exposure. Thus, pixels provide linear outputs proportional to light intensity. A dynamic range of 120dB is expected. The maximum dynamic range that can be measured is limited by the maximum event rate that the chip peripheral circuitry can handle and by the space dedicated on memory to store the event information. Pixel pitch is 25 $\mu$ m. A prototype sensor with 128 x 96 pixels has been implemented in the AMS 180nm CMOS-HV technology. In this article, the pixel operation will be explained. Preliminary experimental results and snapshots will be also provided.

## Introduction

High-dynamic-range (HDR) image sensors are required whenever sensors have to operate in environments with uncontrolled illumination conditions. For instance, high dynamic range of operation is a must in automotive applications, and surveillance.

APS image sensor has limited dynamic range. For a fixed integration time, the lower photocurrent value that can be sensed is limited by the quantization noise or the read noise of the analog-to-digital-converter. The largest photocurrent value that can be measured is limited by the full well capacity. By increasing the integration time, it is possible to sense lower photocurrent values, but the highest photocurrent that can be gauged, will be lower. If the integration time is decreased, it will be possible to operate with higher illumination, but precision within lower illuminated areas will be lost. Therefore the dynamic range is inherently limited by sensor and cannot be extended adjusting the integration time. Typically APS image sensor DR is below 70dB [1].

For some applications, the intra-scene dynamic range, i.e. the ratio between the largest and the smallest illumination, is way beyond the ratio between the maximum and the minimum detectable signals with an APS pixel. In order to render high-dynamic-range images, there are several methods reported [1]-[2]. The most popular is to obtain dual or multiple captures of the same scene with different exposures [3]-[4]. These frames, captured with different integration periods, are processed and combined

following relatively complex post-processing algorithms to render the final HDR photograph [5]. This procedure requires data processing and a frame buffer for each different exposure time. Another approach is to use a compressive function to capture and codify high-dynamic-range images with a reduced number of bits [6]-[7]. Typically, these methods are based on the computation of a histogram of the scene to determine the pixel illumination distribution. Based on the histograms, the precision to codify each image region is assigned. The drawback is that these sensors cannot capture all the illuminated areas with the same accuracy. Event-based vision sensors usually offer high dynamic range operation, [8]. Unfortunately, their event-based description of the scene is not easily compatible with the most widely employed frame-based image processing tools and displays.

In this paper, we present an image sensor with linear response to illumination over a high dynamic range. It has a frame-based output combined with classic APS readout. Its pixels are never allowed to reach saturation during the integration time. Each time a pixel reaches a threshold, it is reset and an event is sent to be properly accounted for. In the end of the exposure time, pixel values are digitized with an analog-to-digital converter. The output is the combination of the ADC readings with the number of events associated to each pixel. With this we expect to obtain 20b digital words that represent pixel illumination values, what is equivalent to a 120dB linear dynamic range.

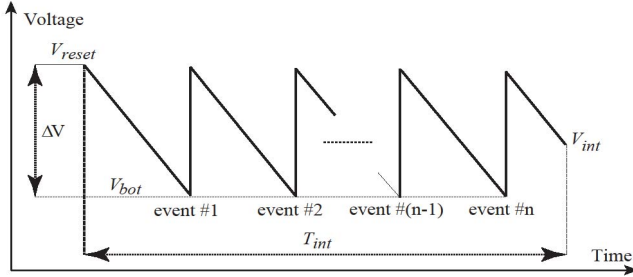
The proposed method does not require multiple exposures nor memory to store different frames. Furthermore, all the scene illumination is measured with the same precision and no post-processing to codify the pixel intensity is required.

The paper is structured as follows: Section II explains the principle of operation and the pixel schematics. Section III presents a system level description. Section IV displays preliminary experimental results. Finally, Section V draws conclusions.

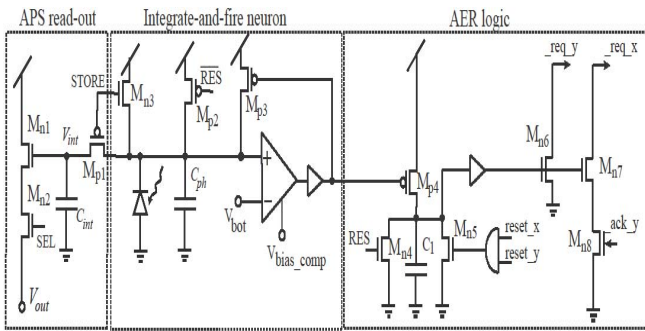
## Pixel's Concept

Figure 1 depicts the proposed new approach to extend the dynamic range. The idea is that the pixel is never allowed to overflow. If the voltage at the integration capacitance reaches a programmable voltage threshold ( $V_{bot}$ ), the pixel will reset itself immediately after and will continue integrating charge until the end of the integration period ( $T_{im}$ ). At this point, the final voltage at the integration capacitance ( $V_{im}$ ) will be digitized and stored on a memory. Knowing the number of times (if any) that each pixel has reached the threshold, it is possible to render HDR images. Note that with this approach high illuminated pixels will never saturate. As long as the number of events sent out can be stored on a memory, the operation will be correct. Thus, there is a trade-off between DR and the number of bits dedicated on memory to store the number of events spiked by each pixel.

The intra-scene dynamic range of a visual scene can be defined as the ratio between the highest and the lowest illumination values that can be sensed within the same frame. It is usually expressed in decibels:



**Figure 1** Principle of operation of the pixels. Every time that the voltage at the integration capacitance reaches the value  $V_{bot}$ , the pixel resets itself and sends out an event indicating this occurrence. Thereafter, it continues integrating charge. The final voltage  $V_{int}$  is digitized and stored on a memory. Combining the event information with the ADCs outputs, it is possible to render HDR images.



**Figure 2** Pixel's schematics. Every pixel has an APS readout, an integrate-and-fire-neuron that spikes with a frequency proportional to illumination, and asynchronous logic to readout the events outputs.

$$DR = 20 \log_{10} \left( \frac{I_{max}}{I_{min}} \right) \quad (1)$$

In our particular case, i. e. for a fixed sensing capacitance and for a particular noise floor,  $I_{min}$  is limited by the value of the integration time ( $T_{int}$ ). The higher  $T_{int}$  the smaller the photocurrent that still render a meaningful pixel voltage, and therefore the smaller  $I_{min}$ . The only limitation for the maximum light intensity that can be measured ( $I_{max}$ ) is either the number of bits dedicated to account for saturation events, or the maximum event rate that can be handled by the arbitration system.

If we represent the illumination, and hence the photocurrent values, with binary words of  $N_{bits}$ , the dynamic range of the sensor is given by:

$$DR = 20 \log_{10} (2^{N_{bits}}) = 20 \log_{10} (2^{(N_s + N_b)}) \quad (2)$$

Where  $N_{bits}$  is the number of bits needed to encode all the illumination levels that can be measured.  $N_s$  is the number of bits dedicated to store the number of events, and  $N_b$  is the precision of the analog-to-digital conversion of the voltage ( $V_{int}$ ). Hence, to render 120dB-intra-scene-dynamic-range images, we need at least 20b. Conventional APS image sensors have a dynamic range usually below 70dB [1]. Analyzing the approach presented in Figure 1 to extend the dynamic range, the value of the measured photocurrent, and hence the detected brightness, will be proportional to:

$$I_{ph} \propto V_{reset} - V_{int} + (V_{reset} - V_{bot}) \cdot (\#events) \quad (3)$$

Figure 2 shows the pixels schematics. Every pixel has circuitry to implement APS readout. Furthermore, we have added an oscillator (also known as integrate-and-fire-neuron) inside each pixel. Its target is to pulse every time that the programmable voltage threshold ( $V_{bot}$ ) is reached. Note that the pulses reset the voltage at the integration capacitance and activate the asynchronous circuitry that handles the event communication. Knowing the number of times that a pixel has spiked and the value of  $V_{out}$ , it is possible to obtain a digital word which value is proportional to the pixel illumination.

The oscillator of Figure 2 generates pulses with a period that is approximately:

$$T = \frac{C \cdot (V_{reset} - V_{bot})}{I_{ph}} + T_d + T_{reset} \approx \frac{C \cdot \Delta V}{I_{ph}} \quad (4)$$

$T_{reset}=400ns$  is the amount of time required to reset the integration capacitance when  $\Delta V=4V$ .  $T_d=25ns$  is the controlled delay introduced to make the oscillator stable. For simplicity, its value and time required to reset the integration capacitance can be neglected in the forthcoming circuit analysis because there are much lower than the oscillation period under average illumination conditions.

For a given value of the integration period depicted in Figure 1 ( $T_{int}=1/FR$ ), the minimum detectable photocurrent is:

$$I_{ph_{min}} = \frac{C \cdot \Delta V \cdot FR}{2^{N_b}} \quad (5)$$

Therefore the dynamic range expressed as a function of the maximum photocurrent that can be measured ( $I_{ph_{max}}$ ) and the frame rate (FR) is:

$$DR = 20 \log_{10} \left( \frac{2^{N_b} I_{ph_{max}}}{FR \cdot C \cdot \Delta V} \right) \quad (6)$$

$I_{ph_{max}}$  is the maximum photocurrent that can be measured without saturating the AER (Address Event Representation) arbitration system [9]-[10]. As it was discussed in [10], one practical limitation of our approach is the maximum event rates

that the peripheral asynchronous circuitry can handle. Pixels events must be attended on time i.e. the handshaking cycle must be lower than the time interval between two consecutive events fired by the same pixel. If not, pixels digital output values will be corrupted. The arbitration peripheral circuitry that we are using can handle event rates up to  $MAX_{BR}=10$ Meps for pixels of different rows and up to 2Meps for pixels of different rows [9]. Let us define  $f_{max}$  as the maximum pixel average spiking frequency that the sensor can cope. If the sensor has  $M \times N$  pixels,  $f_{max}$  must satisfy:

$$M \cdot N \cdot f_{max} \leq MAX_{BR} \quad (7)$$

Combining Equations (6) and (7), we can express  $I_{ph_{max}}$  as a function of maximum event rate and the pixel parameters:

$$I_{ph_{max}} = f_{max} C \cdot \Delta V = \frac{MAX_{BR} \cdot C \cdot \Delta V}{M \cdot N \cdot \alpha} \quad (8)$$

$\alpha$  is a parameter that denotes the proportion of pixels exposed to  $I_{ph_{max}}$ . Substituting  $I_{ph_{max}}$  in Equation (6), it is possible to determine the system DR as a function of the FR and system parameters:

$$DR = 20 \log_{10} \left( \frac{2^{N_b} MAX_{BR}}{FR \cdot M \cdot N \cdot \alpha} \right) \quad (9)$$

Note that the DR does not depend on the value of  $\Delta V$  and the integration capacitance. The equation is valid as long as the number of bits on memory dedicated to store the illumination values is high enough, i.e.  $DR \leq 20 \cdot \log_{10} (2^{N_b+N_s})$ . Usually, in scenes with large intra-dynamic range, the percentage of pixels exposed to the maximum illumination is below 10%. Figure 3 displays the relation between the expected dynamic range for a given frame rate and different values of the  $\alpha$  parameter. It can be seen that with our approach is possible to achieve a dynamic range higher than 105dB for video rates. Reducing the event rate, it is even possible to render images with an intra-scene dynamic range beyond 130dB. In real application scenarios, it is challenging to find natural scenes with an intra-scene dynamic range higher than 120dB [1]. To test of our system, we assigned  $N_s=16$ bits. Thus, according to Equation (2), if  $N_b=8$ bits, the maximum theoretic intra-scene dynamic range that can be measured is close to 145dB.

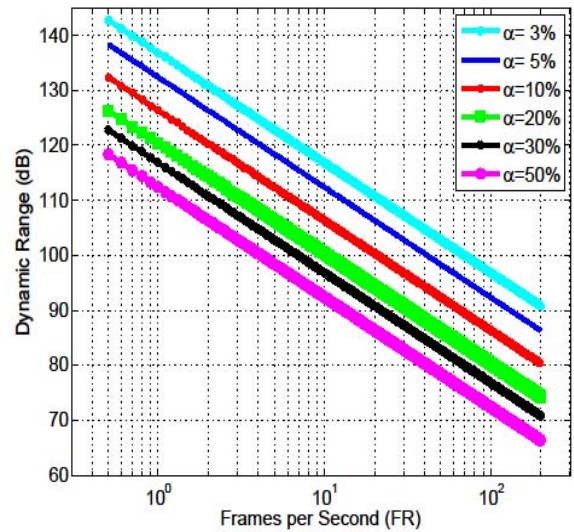


Figure 3 Expected dynamic range versus frame rate for different values of the  $\alpha$  parameter

## System Level Description

Figure 4 displays the system level block diagram. In the middle, there is the pixel array. On the periphery we have placed the AER and APS readouts. Both operate in parallel generating two data flows. One is related to the event outputs and the other one corresponds to the analog-to-digital-conversion of the voltage at the integration capacitance ( $V_{int}$ ), in the end of the integration period. On the top and right, there is the AER readout circuitry. It is made of arbiters, asynchronous generic logic, and decoders. First, row petitions are attended. Then column petitions are arbitrated.

A detailed description of the AER logic and its implementation can be found in [9].

On the left there is some synchronous logic to generate the typical control signals of the APS pixel (SEL, RES, and STORE). We use global shutter for the APS readout. The signals RES, and STORE are selected globally. The signals SEL are activated sequentially, row by row. Finally, on the bottom, there is the circuitry for the APS readout. There are 128 column-parallel ramp ADCs to digitize the output voltages of the columns. The converters digital outputs are latched to a SRAM memory where they are stored until they are readout. For this purpose, there is a block that reads and sends out the chip each value stored on memory.

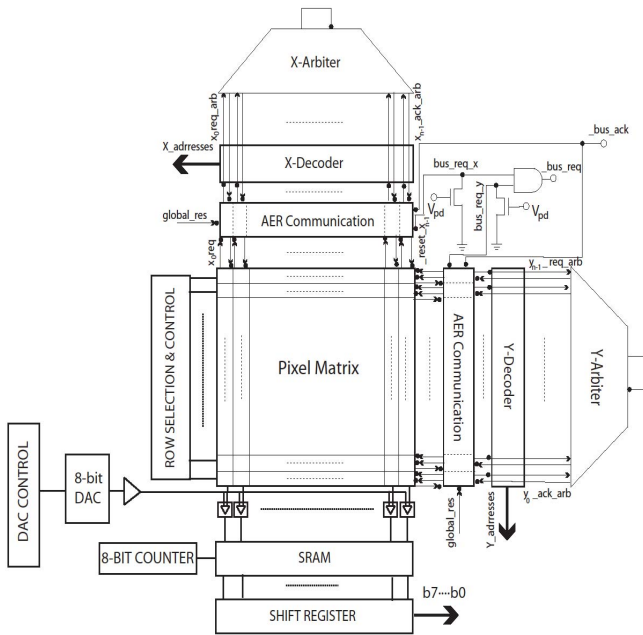


Figure 4 System's block diagram.

Figure 5 depicts the column parallel ADCs implementation. We have implemented a DAC with a resistive divider to generate the reference ramp voltages. Different DAC output voltages are selected sequentially with shift registers. It is possible to control the highest the voltage operation range ( $V_{top}-V_{bot}$ ). The initial value of  $V_{DAC}$  can be tuned with the  $V_{standby}$  signal. Each pixel output voltage ( $V_{pix}$ ) is connected to a comparator that compares it to the DAC output. When the voltage  $V_{pix}$  reaches the voltage  $V_{DAC}$ , a digital signal (EON) is activated to indicate the end of the pixel conversion. Then the digital code that corresponds to each DAC output is stored on the SRAM memory shown in Figure 4. The voltage  $V_{DAC}$  is driven to all the comparators (128 in total) with an analog buffer designed for that purpose. The conversion time of an entire row is approximately  $35\mu s$ .

Figure 6 shows in details the digital blocks that have been implemented on chip to readout the APS outputs. When the signals End of Conversion (EOC) are active, the counter output is latched at the SRAM memory registers that corresponds to the column that has activated the signal EOC. Once the digital-to-analog-conversion has finished, the data stored on the SRAM ( $N_{cols} \times N_b$  bits) are latched to the READ BUFFER block. Every cycle of the signal READ\_BUFFER\_CLK, its outputs are sent out of the chip. Note that with this approach, pipeline operation is possible. While one row pixel outputs are digitized, the previous row output values can be sent out of chip.

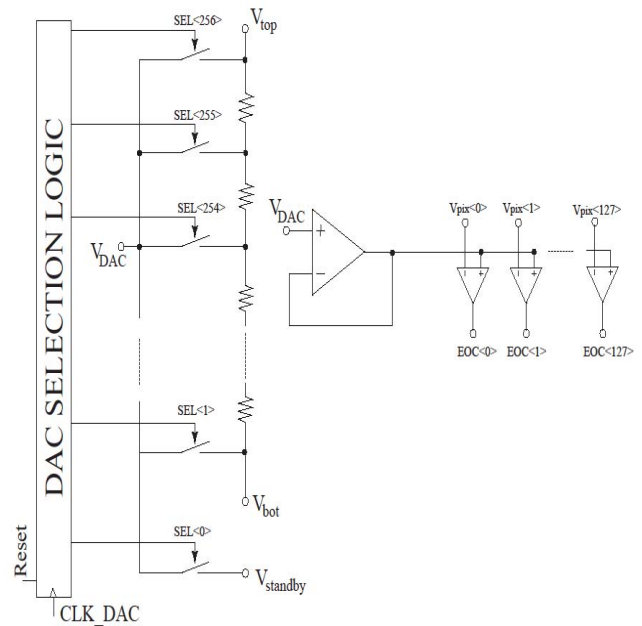


Figure 5 Detail of the column parallel ramp ADCs implementation.

Figure 7 shows the chip communication with an external Opal Kelly XC7K160T board. It contains a Kintex 7 FPGA that handles the event and the APS output data flows. Both of them are stored on an external DDR3 SRAM memory. The most  $N_s$  significant bits correspond to the event information. The less significant bits ( $N_b$ ) come directly from the analog-to-digital-conversion of the pixel voltages. The usage of an external memory to store the pixel illumination levels gives a lot of versatility for the chip test. The number of bits on memory dedicated to store the event information ( $N_s$ ) can be configured, depending on the dynamic range requirements. The FPGA internally contains two FSMs to readout the two data flows. One of them implements the AER communication with the arbitration system on chip. The second one is dedicated to read sequentially the stored ADCs outputs.

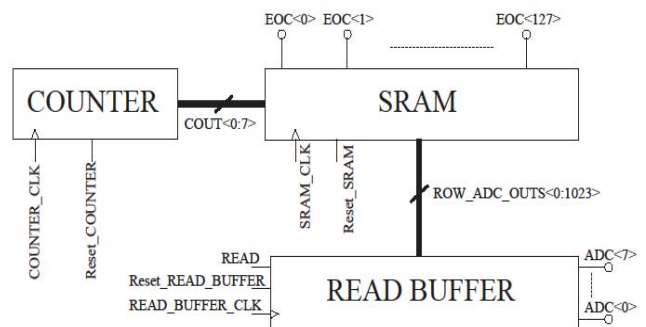
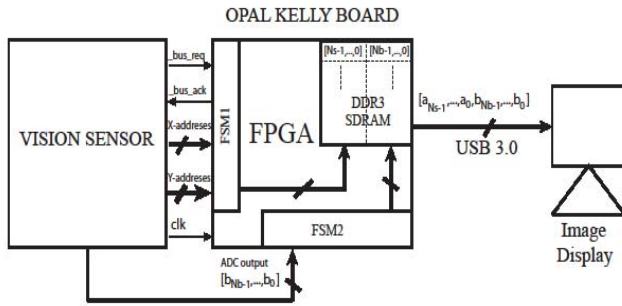
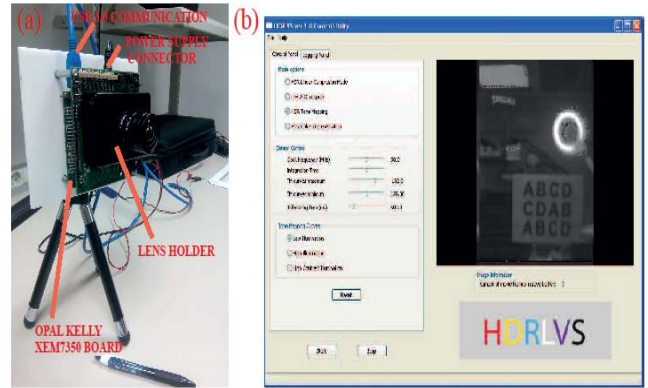


Figure 6 Detail of the digital circuitry dedicated to readout the ADCs outputs and to send their values out of chip.

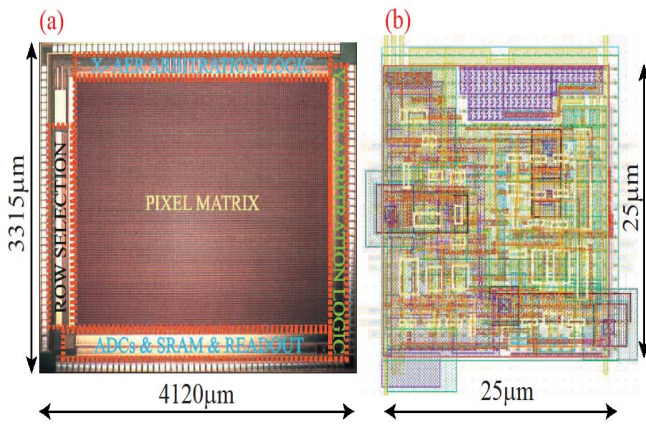




**Figure 7** Chip communication with the external Opal Kelly XC7K160T board. The signals involved in the APS and event readout are shown. Both data flows are stored on a DDR3 SRAM memory. Finally, output data is send through an USB 3.0 port to a PC.



**Figure 9** (a) Experimental setup. (b) Custom interface for sensor debugging and data representation.



**Figure 8** (a) Chip microphotograph. (b) Pixel layout.

## Experimental Results

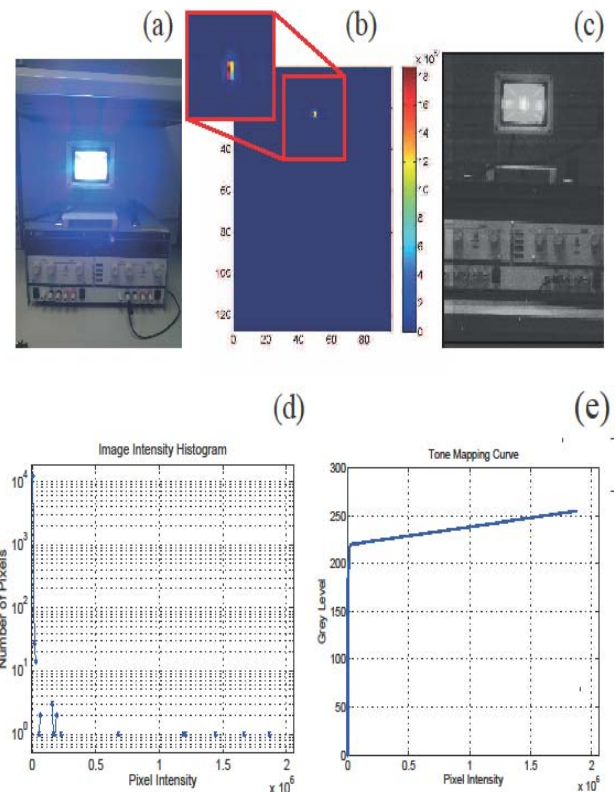
### Chip Microphotograph and Layout

Figure 8 displays a chip microphotograph and the pixels layout. Chip dimensions are  $4120\mu\text{m} \times 3315\mu\text{m}$ . Pixels size is  $25\mu\text{m} \times 25\mu\text{m}$  with a fill factor of 10%. Table I summarizes the main sensor features.

### Experimental Setup and Interface

Figure 9 (a) displays the experimental setup. We designed a custom PCB and a lens holder to test the sensor. The lens holder was printed with a 3D-printer. The Opal Kelly XC7K160T board is attached to the PCB with a FPGA Mezzanine Card (FMC).

In order to debug the sensor and display real-time images, a custom interface was programmed in C++ language. Figure 9 (b) shows the interface displaying real-time images. Several sliders to control parameters like the integration time and the analog circuit biasing have been added.



**Figure 10** Samples Images. (a) Image taken with a BQ Aquaris E 4.5 smartphone working in HDR mode. (b) Snapshot of the same visual scene taken with our sensor. Intensity levels are encoded with a thermal code. (c) Snapshot of the same visual scene. Grey levels are encoded processing the sensor outputs with a tone mapping algorithm. (d) Image histograms. (e) Tone mapping curve used to encode grey levels of image (c).

### Sample Images

Figure 10 (a-e) shows some example data. Figure 10 (a) displays a HDR image of a natural scene taken with a BQ Aquaris E 4.5 smartphone operating in HDR mode. A snapshot of the same visual scene was taken with our sensor. In Figure 9 (b) the intensity levels encoded with the sensor have been represented using a thermal code. The intra-scene dynamic range of the scene was 124dB. We set a value of  $T_{int}=110ms$  to capture it. With a thermal code it is possible to assign up to 24 bits to the intensity levels representation. This representation is lineal and preserves all the scene illumination values measured with the camera. However, visually it is difficult to appreciate the image details because the intensity levels differ a lot and span within a very large linear scale. Figure 10 (c) displays the same visual scene after processing the sensor's outputs with a tone mapping algorithm [6]. Tone mapping algorithms map the input intensity levels into a grey scale with 256 levels that can be represented with a conventional display. There is a non-linear relation between the input intensity levels and grey levels assigned to each pixel. The user can define curves that adapt to the image histograms to achieve more precision encoding grey levels with the illumination values that are more frequent in the visual scene. Figure 10 (d) displays the image histograms. Figure 10 (e) shows the tone mapping curve used the render the image of Figure 10 (c).

**Table I: Sensor Features**

Technology	AMS 0.18 $\mu$ m HV
Power Supply	1.8V (digital) /5V (analog)
Chip Dimensions	4120 $\mu$ m $\times$ 3315 $\mu$ m
Pixel Size	25 $\mu$ m $\times$ 25 $\mu$ m
Number of Pixels	128 $\times$ 96
Pixel Complexity	34 Transistors + 2 Capacitors
Fill Factor	10%
Dynamic Range	120dB@3fps, 105dB@30fps
Power Consumption	58.6mW @100frames/s, 200keps
Max. event rate	10Meps (different rows), 2Meps (same row)

### Conclusions

A new sensor with high dynamic operation and linear response has been presented. The sensor pixels reset themselves and never overflow during the integration time. Events are sent out

every time that a pixel resets itself. The sensor can achieve high dynamic operation detecting images with intra-scene dynamic range beyond 130dB. The maximum high dynamic range is limited by the maximum event rate that the arbitration system can handle and by the number of bits dedicated to store the number of events associated to each pixel. Power consumption is 58.6mW under average operation conditions. Preliminary experimental results and sample images are provided.

### Acknowledgements

This research work has been partially supported by ONR project N00014-14-1-0355, project MONDEGO (TEC2012-38921-C02-02) MINECO (European Region Development Fund, ERDF/FEDER), and project SMART CIS-3D (P12-TIC-2338), European Region Development Fund, ERDF/FEDER.

We are really grateful to Miguel A. Lagos-Florido for the PCB and the lens holder designs.

### References

- [1] A. Darmont, "Methods to extend the dynamic range of snapshot active pixel sensors," in Proceedings of SPIE Vol. 6816, 681603 (2008), February 2008.
- [2] A. Spivak, A. Belenky, A. Fish, and O. Yadid-Pecht, "Wide dynamic range CMOS image sensors —comparative performance analysis," IEEE Transactions on Electron Devices, vol. 56, no. 2, pp. 2446–2461, November 2009.
- [3] M. Mase, S. Kawahito, M. Sasaki, and S. Wakamori, "A wide dynamic range cmos image sensor with multiple exposure-time signal outputs and 12-bit column-parallel cyclic A/D converters," IEEE J. Solid-State Circuits, vol. 40, no. 12, pp. 2787–2795, December 2005.
- [4] O. Yadid-Pecht and E. R. Fossum, "Wide intrascene dynamic range CMOS APS using dual sampling," IEEE Trans. Electron Devices, vol. 44, no. 10, pp. 1721–1723, October 1997.
- [5] P. E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," in University of California at Berkeley, 1997, SIGGRAPH97 Conference, August 1997.
- [6] L. Meylan, "Tone mapping for high dynamic range images," Ph.D. dissertation, Ecole Polytechnique Federale de Lausanne, Switzerland, 2006.
- [7] S. Vargas-Sierra, G. Liñán-Cembrano, and A. Rodríguez-Vázquez, "A 151dB high dynamic range CMOS image sensor chip architecture with tone mapping compression embedded in-pixel," IEEE Sensors Journal, pp. 1721–1723, July 2014, DOI: 10.1109/JSEN.2014.2340875.
- [8] 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," IEEE Journal of Solid State Circuits, vol. 46, no. 1, pp. 259–275, January 2010.
- [9] P. Häfliger, "A spike based learning rule and its implementation in analog hardware," Ph.D. dissertation, ETH Zürich, Switzerland, 2000, <http://www.ifi.uio.no/~hafliger>.
- [10] Juan. A. Leñero-Bardallo, R. Carmona-Galán, and A. Rodríguez-Vázquez, "A high dynamic range image sensor with linear response based on asynchronous event detection," in 22nd European conference on circuit theory and design, ECCTD 2015, August 2015, pp. 1–4. DOI: 10.1109/ECCTD.2015.7300079

## **Author Biography**

*Juan A. Leñero-Bardallo received the M.Sc. degree in telecommunications engineering and the Ph.D. degree in microelectronics from the University of Seville, Seville, Spain, in 2005 and 2010, respectively.*

*From January 2006 to January 2010, he was working toward the Ph.D. degree at the Institute of Microelectronics of Seville, sponsored by a national grant. In 2008, he was a Visiting Scholar at the University of*

*Oslo, Oslo, Norway, for two months. From September 2010 to March 2010, he worked as a Postdoctoral Associate at Yale University, New Haven, CT, USA. From March 2010 to August 2013, he was a Postdoctoral Associate at the University of Oslo. Since August 2013, he is a postdoctoral associate at the University of Seville, Spain. His main research interests include address event representation vision systems, frame-based vision sensors, smart sensors, wireless vision sensor networks, signal processing, and very large scale integration emulators of biological systems.*