A 5 Million Frames Per Second 3D Stacked Image Sensor With In-Pixel Digital Storage

Laurent Millet, Margaux Vigier, Gilles Sicard, Nils Magotat, Fabrice Guellec, Wilfried Uhring

A 5 Million Frames Per Second 3D Stacked Image Sensor With In-Pixel Digital Storage

Laurent Millet, Margaux Vigier, Gilles Sicard,
Nils Margotat, Fabrice Guellec
Université Grenoble Alpes, CEA, LETI
MINATEC Campus, F-38054, Grenoble, France
laurent.millet@cea.fr

Wilfried Uhring
ICube Laboratory, University of Strasbourg and CNRS
Strasbourg, France
Wilfried.uhring@unistra.fr

***Abstract*—A CMOS burst image sensor reaching 5Mfps with 52 frames in-pixel digital memory has been designed and tested. It fully takes advantage of 3D stacked technology to implement a scalable architecture for 8-bits quantization and data storage at pixel level in CMOS technology. This imager also benefits from backside illumination (BSI) for improved fill factor and wide spectrum sensitivity. A demonstrator has been fabricated, embedding two types of 3D based pixel. In this paper we present the very first experimental test results of 3D stacked in-pixel digital burst image sensor. These results show advantages of using 3D technology to obtain a very high frame rate with both relaxed design conditions and readout timing constraint compared to conventional high speed burst image sensors.**

Keywords—burst imaging, digital storage, 3D stack, CMOS

I. Introduction

High speed imaging is a key technology to study high speed phenomenon such as micromechanics, explosives, plasma formation mechanisms or laser ablation. In a classic image sensor (IS), images are acquired with on the fly chip read-out. However, for high speed IS this read-out process is a bottleneck limiting the frame rate to about 10 kilo-frame-per-second (fps) for a 1 mega-pixel resolution as presented in [1]. To tackle this limitation, burst IS uses an on-chip memory to store the burst video sequence recorded during the acquisition. Images are then read-out at lower speed. While the frame rate of burst IS can exceed 10 M frames per second (Mfps), they suffer from the small embedded memory limiting the video length to a maximum of hundreds frames.

Burst image sensors are currently divided into two technological categories: CMOS and CCD. CCD imagers can achieve high frame rate with high in-pixel analog memory density [2], [3] and they also benefit from a very high fill factor when using BSI process. However this technology involves several drawbacks. First, they have a huge power consumption to supply the high gate control voltages. Secondly the analog memory is placed in the focal plane. While BSI configuration is much better for fill factor considerations, it exposes the memories to backside photo-current generation which can degrade the stored data. It requires thick substrate to cut wavelengths above 700nm [2], [3]. This in-pixel constraint also limits the memory to 100 frames. In addition the CCD technology impedes the use of embedded functions such as Analog to Digital Converter (ADC), or processing features which arise in smart CMOS solutions e.g. compressive sensing [4]. Up to now, high speed

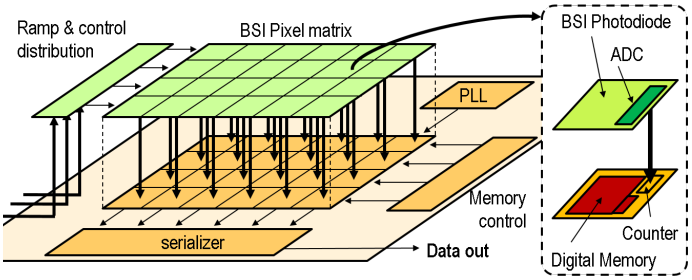

Fig. 1. Overview of our 3D stacked CMOS image sensor partitionning

image sensors using CMOS implementation hardly compete with CCD sensors. The CMOS local memorization is also performed thanks to very small capacitors despite large KTC noise. We distinguish two kinds of implementations regarding analog memories: either using memory banks outside the pixel array [5], [6], or using in-pixel memory [7]. In this last case, the main drawback is the very small pixel fill factor. The use of memory banks limits the scalability of the array size and faces power penalty due to signals propagation on the entire array through high capacitance lines at high speed. In both cases, there is also a strong limitation on the readout phase due to leakage on small capacitor memories. They require a fast readout and hence high working frequencies before data deterioration. Some works [4] circumvent this limitation by using time coded multi-capture, but are still limited to 15 frames memory.

A 3D stacked IC approach overcomes these restrictions by allowing a distribution of electronic functions on dedicated tiers [8]. Top tier photodiodes can benefit from imaging technology and backside illumination for a maximal sensitivity, while the bottom tier provide room for in-pixel memories without sacrifying the fill factor (Fig. 1). We designed a scalable 3D pixel using direct bonding stacking, which causes no dead area in the silicon [10] on the contrary of TSV solutions. We also take advantage of this topology to design an in-pixel ADC for embedded digital storage. Our Image Sensor benefits of a 40nm advanced technological node on bottom tier for high density capability. Using RAM-based memory could easily overcome the current hundreds of stored images and allows a video memory of thousand images [9].

This paper describes the architecture and electrical results of the very first 3D-stacked BSI burst image sensor with in-

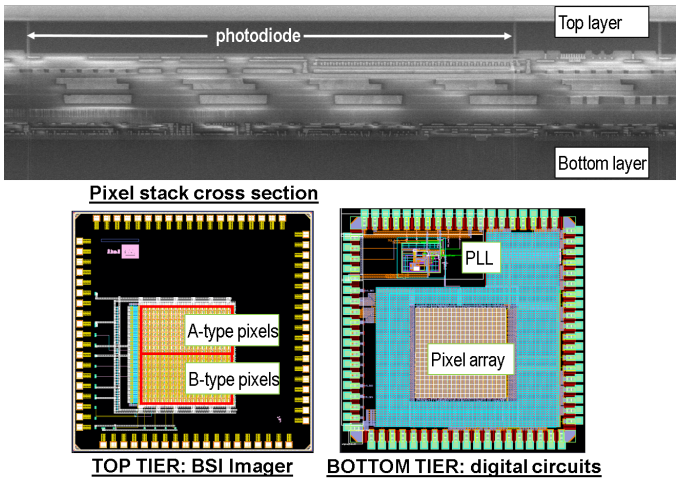

Fig. 2. Top and bottom layout view, with pixel cross section detail

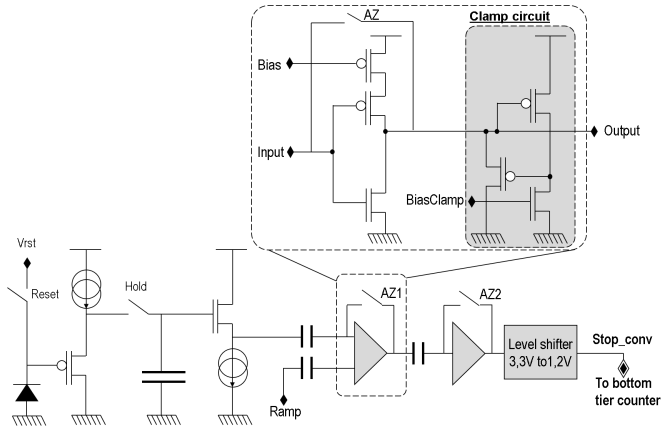

Fig. 3. Top tier readout and comparator circuits with clamp detail for current stabilization

focal-plane digital memory. The overall chip description is presented in section II. Top tier and bottom tier architecture are described in section III and IV, while experimental results are detailed in section V.

II. Overall description

As said, the 3D stacked implementation is very suitable for mixing analog and digital circuits with dedicated technological nodes. Our prototype uses an 90nm image sensor layer over a 40nm digital layer connected by hybrid bonding process. The selected partitioning is illustrated in Fig. 2. The top tier embeds an array of 20×20 BSI pixels. Each top tier pixel contains a photodiode with its associated readout structure and the analog frontend of a ramp ADC, while the bottom tier pixel contains the counter and digital memory to store 52 images. The bottom chip embeds the control circuits, power distribution, interface controls, and a PLL to drive the counters. The IO ring is also located on bottom tier involving that control signals come from the same die for both top and bottom chips. The top tier is supplied by the bottom chip through 3D Cu-Cu contacts.

III. Top tier pixel structure

The 3D partitioning relaxes area constraints on readout and ADC circuits. We designed a scalable pixel with a total area of $50\mu m \times 50\mu m$.

A. Photodiode

The 3T BSI photodiode dimensions are $40\mu m \times 50\mu m$ and reach a fill factor of 80%. It is surrounded by a deep trench isolation to avoid inter-pixel interferences and to protect the photosensitive elements from disturbances of the nearby active CMOS circuits (leakage and electron harvesting).

Due to large photodiode dimensions we used multiple collecting contacts. Two different photodiode designs were implemented: A-type version with 4 contacts and B-type version with 15 contacts. These two designs bring different integrating capacitances values (made by junction photodiode and parasitics). A-type shows a lower value (25fF) than the B-type (50fF), but the B-type may exhibit a better time response during charge collection.

B. Pixel readout structure Analog to Digital Converter

The chosen frontend readout structure is a 3T pixel with a sample and hold stage, selected for its fast time response. A CTIA could be implemented for better voltage control on sensing node and signal gain thanks to feedback capacitances. Yet, since the photodiode capacitance is expected to be low, the source follower has been considered as better option.

A sample and hold circuit is used to stop integration and ensure a clean signal for the ADC. The acquired signal is then quantized while next frame is already integrating on photodiode capacitance.

C. Analog to Digital Converter

The ADC topology is a ramp converter. The comparator is designed in the top tier (Fig. 3) while the counter circuit is in the bottom layer. The comparator is a double stage amplifier followed by a level shifter for voltage compatibility with digital layer. It performs a correlated double sampling on pixel signal, and embeds a double offset cancellation circuits through AZ1 and AZ2 stages. Note that a comparator usually produces kickback noise on reference voltage lines and power supply, magnified by the matrix dimensions. This is an issue at our operating frequencies, since the references take time to come back to their initial level. To tackle this problem we implemented a clamp circuit to drift current even after the comparator triggering, which ensure a constant power consumption over the matrix array and brings a significant kickback noise reduction.

There are two 3D contacts per pixel. The first is the comparator output signal called “stop_conv”, from the top to bottom. It freezes the counter value. The second one comes from the bottom chip. It brings the 1.2V power supply to the top tier for the level shifter.

D. Top tier pixel overview

The layout implementation of an A-type pixel and the readout circuits is shown in Fig. 4. A metal 1 and 2 shield is drawn over the photodiode, and the BSI implementation allows to use free space for fringe capacitances. One of the 3D connections hangs over the neighbor pixel.

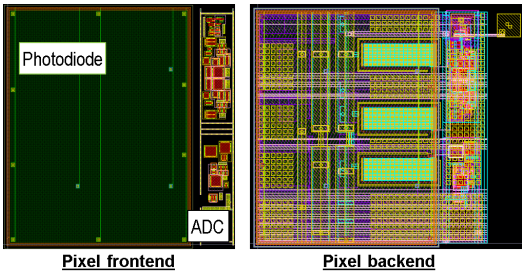

Fig. 4. Layout implementation of the top tier BSI pixel (A-type)

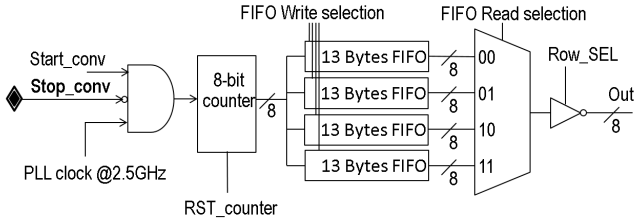

Fig. 5. Bottom tier pixel architecture

IV. Bottom tier pixel structure

The bottom tier pixel embeds counters for the ramp ADC, a bank of 52 Bytes memory, and top/bottom synchronization circuits.

A. Analog to Digital Conversion

Digital pixel comprises an 8 bit counter, driven by a clock up to 2,56GHz generated by a peripheral PLL circuit. This frequency ensures an 8bits conversion at 5Mfps. When the conversion begins, the counter value is incremented at each clock rise and stops as soon as the top tier comparator has switched. The digital value is then automatically stored in the selected 13 bytes FIFO out of the 4 available per pixel (Fig. 5). This implementation enables continuous recording while no event occurs. This is convenient if the imager is waiting for an specific event, since the acquisition can be stopped with a dedicated external trigger. Data are then read-out through a low frequency serial link. Each pixel memory can be selected through "Row_SEL" gate. There is no time constraint here since digital memories do not suffer from leakage issues like in CCD or CMOS analog storage solutions.

B. Signal control and synchronization

Bottom tier handles synchronization signals. Since the frame capture global shutter based, the 20×20 pixels of the array work in parallel. Hence, 3D drive signals from bottom to top tier are sent outside the focal plane for line-broadcast communication, as well as ramp distribution. Start conversion signals are sent to top and bottom tier outside pixel array for conversion synchronization of both layers, as depicted in Fig. 6.

C. Embedded Digital Memory

For this first prototype digital memories are realized with D Flip-Flops for a total memory area of $45\mu m \times 45\mu m$. This memory topology allows 52 pictures of 8 bits to be sequentially written. A much larger memory depth could be achieved by using a high density SRAM instead of D-Flip Flops. It would involve an important increase of storable pictures. Typically it could allow around 775 frames in 28nm technology and 387 frames in 40nm technology for the same pixel pitch pitch while including the SRAM control circuit. Implementing SRAM will involve partitioning issues (one cut for several pixels) and local pixel addressing considerations. For instance this SRAM partitionning has already been investigated in the context of high speed imaging [8].

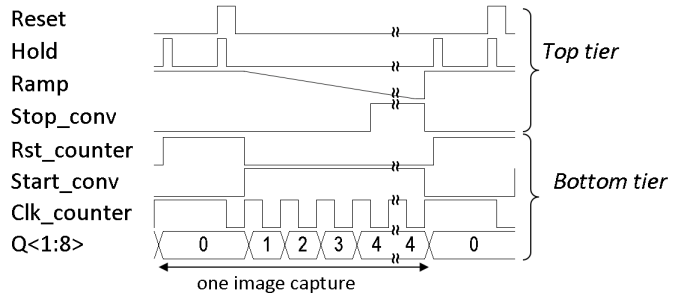

Fig. 6. Digital pixel control time diagram

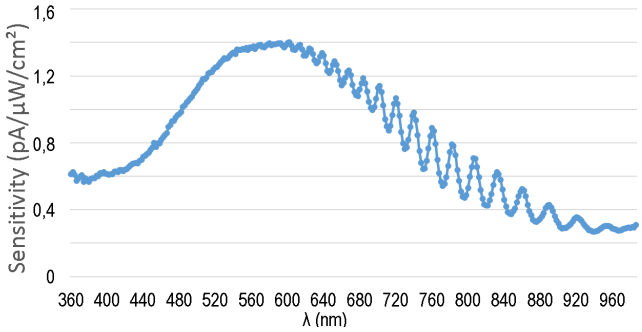

Fig. 7. Spectral sensitivity of the A-type pixel

V. Electrical Measurements

To evaluate our demonstrator performances, two kind of measure analysis have been performed. First photodiodes have been characterized thanks to a test structure allowing access to the internal nodes. Secondly we also demonstrated standard full frame electrical acquisition.

A. Photodiode caracteristics

Photodiode characterization measures were performed on optical bench with a monochromator. The extracted spectral sensitivity is shown in Fig. 7. Since we are not limited by CCD constraints which needs to use thick oxide to prevent large wavelength from corrupting the analog memories, our measured diagram exhibit a wide spectrum range from 360nm to 900nm. The visible oscillations are known as etalon effect and they appear when a beam of monochromatic light causes constructive and destructive reflection interferences when reaching two parallel interfaces.

B. Dynamic captures: pulse response

We performed several transient measures to evaluate the IS response to fast moving scenes. On the first example shown

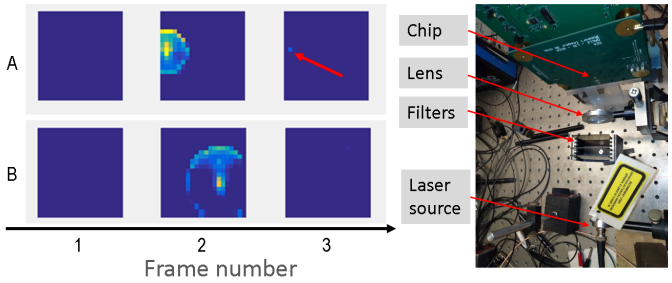

Fig. 8. Two 3 frames sequences to evaluate beam persistance on A-type and B-type pixels (left), and the test bench (right)

TABLE I. TECHNICAL SPECIFICATIONS SUMMARY

Scalable Pixel specifications		Readout characteristics	
Technology	4M1P 90nm 7M1P 40nm	Read Noise	0.59 mV rms
Pixel size	$50\mu m \times 50\mu m$	Dynamic range	1 V
Fill factor	80%	Convertion gain	7.3 $\mu V/e-$
Framerate	5 Mfps	full well capacity	137k e-
Memory depth (frames)	52 (8 bits images) (387 in 40nm SRAM)* (775 in 28nm SRAM)*	Chip Considerations	
		Pixel count	20×20
		chip size	$2.7 \times 2.7mm$

* estimated

in Fig. 8 the imager is illuminated with a stream of laser pulses adjusted at 660nm wavelength. It emits a 10pJ, 100 ps Full Width at Half Maximum (FWHM) long pulse of light that illuminates the whole sensor sensitive area [11]. The acquisition mode is full frame and the ADC dynamic is set to 8 bits with a 5 Mfps frame rate.

On the first acquisition (stream A) a single pulse focused on the upper side of the imager is sent during the frame 2 to test the A-type pixel response. The laser spot is surrounded by a diffraction circle. The high illuminated central spot causes a small residual remanence illustrated by a 20 codes dot visible on the frame 3 (red arrow). This effect is not visible when the B-type pixel is illuminated on the lower side of the imager (stream B). This means that the A-type pixel has a small amount of residual drifting charges due to insufficient collecting contacts, which is not the case for the B-type photodiode. The circuit exhibits a pixel aperture ratio of 73.7%, measured thanks to the FWHM light pulse.

C. Dynamic captures: sine response

This second example evaluates the response to a half-sine wave exposition. 5 different acquisition streams where made at 5Mfps. Each capture is slightly shifted by 300ns from the previous one, and is presented on Fig. 9. Several pictures of the first capture stream are displayed for assessment. For a better estimation of the captured signals, a graph shows the median value of each photodiode type (A or B) for each picture. We clearly see that the B-type photodiode is less sensitive than the A-type variation. This is due to the higher junction capacitance, hence a higher full well capacity. However in both cases each sine is perfectly visible and we demonstrated that the process is realiable and replicable.

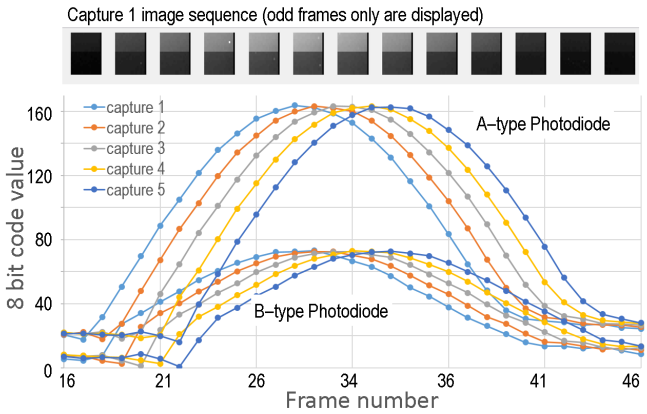

Fig. 9. Median value of 5 sinewave captures with 300ns shift @5Mfps

VI. CONCLUSION

In this paper with presented electrical results of the very first burst image sensor exploiting 3D stacked technology. Our scalable pixel with digital memories allows robust data storage, and then suppress the spectrum limitations inherent to CCD technologies. With 52 images video memory at 8bit quantization, our prototype paves the way to high density high resolution CMOS burst imagers. A list of technical specifications of our imager is summarized in TABLE I.

ACKNOWLEDGMENT

The authors acknowledge the support of ST Microelectronics for technological contribution, and Emilie Huss for advices on photodiode design.

REFERENCES

[1] G. Meynants, G. Lepage, J. Bogaerts, G. Vanhorebeek, and X. Wang, *Limitations to the frame rate of high speed image sensors*, in Proc.Int. Image Sensor Workshop, Jun. 2009, pp. 153-156.

[2] Nguyen, H.D, *Design an event detection system for 16-Mfps extremely-high-sensitivity video camera*, in ICCE, july 2016, pp. 181-185.

[3] Toshiki Arai, et al., *A 252-V/luxs, 16.7-Million-Frames-Per-Second 312-kpixel Back-Side-Illuminated Ultrahigh-Speed Charge-Coupled Device*, IEEE transaction on electron devices, oct. 2013, pp. 3450-3458.

[4] Futa Mochizuki, et al., *Single-Shot 200Mfps 53-Aperture Compressive CMOS Imager*, ISSCC, feb. 2015, pp. 116-117.

[5] Manabu Suzuki, et al., *10Mfps 960 Frames Video Capturing Using a UHS Global Shutter CMOS Image Sensor with High Density Analog Memories*, in international image sensor workshop, 2017, pp. 308-311.

[6] R. Kuroda, Y. Tochigi et al, *A 20Mfps Global Shutter CMOS Image Sensor with Improved Light Sensitivity and Power Consumption Performances*, ITE Transactions on MTA, 2016, Vol.4, No.2, pp. 149-154

[7] L. Wu, D. San Segundo Bello, P. Coppejans, et al., *In-Pixel Storage Techniques for CMOS Burst-Mode Ultra-High-Speed Imagers*, in International Image Sensors Workshop, may 2017, pp. 312-315.

[8] W. Uhring, L. Millet, et al., *A Scalable Architecture for Multi Millions Frames per Second CMOS Sensor With Digital Storage*, unpublished.

[9] R. Bonnard, F. Guellec, J. Segura, A. Dupret, W. Uhring, *New 3D-integrated burst image sensor architectures with in-situ A/D conversion*, DASIP 2013, Cagliari, Italy, pp. 215-222, IEEE

[10] L. Benaissa, et al., *Next Generation Image Sensor via Direct Hybrid Bonding*, EPTC IEEE, 17th, dec. 2015, DOI: 10.1109/EPTC.2015.7412373

[11] W. Uhring, V. Zint, J. Bartringer, *A low-cost high-repetition-rate picosecond laser diode pulse generator*, Proc of SPIE, Vol. 5453, pp 583-590, doi:10.1117/12.545038.