A Reconfigurable QVGA/Q3VGA Direct Time-of-Flight 3D Imaging System with On-chip Depth-map Computation in 45/40nm 3D-stacked BSI SPAD CMOS

David Stoppa[1*], Sargis Abovyan[2], Daniel Furrer[1], Radoslaw Gancarz[1],
Thomas Jessenig[2], Robert Kappel[2], Manfred Lueger[2], Christian Mautner[2], Ian Mills[1],
Daniele Perenzoni[1], Georg Roehrer[2], Pierre-Yves Taloud[1]
[1] ams OSRAM, Eggstrasse 91, 8803 Rueschlikon, Switzerland
[2]ams OSRAM, Tobelbader Strasse 30, 8141 Premstaetten, Austria
*Tel: +41795482964, e-mail: david.stoppa@ams.com

***Abstract* — We present a complete dToF system consisting of a RX camera- and TX laser- modules intended as a highly flexible development platform for next generation ams 3D-dToF imaging products. The 320x240-SPADs array dToF sensor has been fabricated in a 45/40nm 3D-stacked BSI CMOS process and features embedded CPUs and memory to enable full depth-map computation via combined HW/FW system architecture. Thanks to an optimized co-design of the VCSEL and its driver, the Class-I laser module is capable of generating 940-nm, 355-ps FWHM light pulses with peak optical power exceeding 35 W. Measurements from the full system prototype are provided featuring complete on-chip depth map computation. A maximum distance range up to 4.4m/2.0m for Q3VGA/QVGA under 50klux sunlight equivalent ambient light is demonstrated at 30/5fps.**

***Index Terms*—SPAD, Time-of-Flight, dToF, 3D-stacked, 3D imaging, range imaging.**

I. Introduction

THE advent of 3D-stacked BSI SPAD technologies opened the way to a new generation of advanced dToF sensors exploiting the intrinsic advantage offered by deep-submicron CMOS stacked beneath optimized sensor layer. Since the first demonstration of 3D-stacked BSI SPAD [1], it became evident how the key limiting factors of SPAD-based sensors realized in planar processes, namely poor Photon-Detection-Efficiency (PDE) and large pixel pitch for architectures embedding pixel-level processing, can be strongly mitigated. Since then, several implementations have been proposed including on the bottom CMOS wafer all the dToF essential processing blocks, i.e. SPAD front-end and compression, TDC for detected photons timestamping and histogram memory for recording the photons arrival time distribution, and with different level of pixel-to-processing sharing. In [2] 2x8x8 19.8-um pixels are shared using 45/65nm, [3] implemented a 4x4 9.2um SPAD macro-pixel based on 40/90nm, while [4] achieved 2x8x16 7-um SPADs sharing. All these advanced implementations successfully implemented full dToF histogram build-up within the focal-plane-array, however, requiring to read from each macropixel and to stream out of the sensor the large amount of data intrinsic within dToF time histogram. [5] presented a dToF sensor architecture targeting automotive LiDAR systems that exploits line-scanning to implement higher level of resource sharing and on-chip dToF histogram pre-processing enabling multi-echoes detection in a 90/40nm CMOS. In this contribution, we present a complete dToF system, consisting of a 320x240-SPADs array dToF receiver (RX) camera and a dToF-optimized laser module transmitter (TX). This combined HW/FW architecture enables very flexible implementation of different algorithms performing the key functions necessary for dToF depth maps generation, offering a development and testing platform for next generation ams 3D-dToF imaging products.

II. D-ToF Pixel Architecture

Figure 1 shows the block diagram of the dToF sensor macro-pixel. It consists of 4x4 SPADs implemented on the top 45-nm BSI SPAD wafer while the 2.5-V front-end electronics and 1-V processing logic are hosted on the bottom 40-nm CMOS wafer.

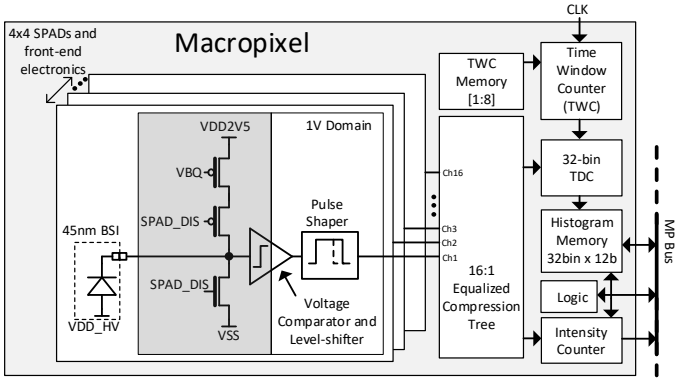

Figure 1: Simplified block diagram of the dToF sensor macro-pixel.

Each SPAD can be individually enabled/disabled and the configured SPAD mask is stored within dedicated register within each macro-pixel. The 16-SPAD outputs are connected

via an equalized 16:1 compression tree logic to an ultra-low DNL 32-bin TDC with programmable resolution between 200-300ps and an intensity counter. The desired ToF range is covered via 8/16 consecutive time windows (Time Window Counter), with neighbor windows partially time-overlapped to guarantee continuous mapping of the whole time/distance range. A 32bin x 12b histogram memory and an intensity counter store the detected photons arrival time (time histogram) and total number of photons respectively. The interaction between each MP and the processing logic located outside the pixel-array is governed by dedicated MP logic and column-shared bus.

III. BSI SPAD

Figure 2 shows a schematic cross-section of the developed 12.5-um pitch BSI SPAD device based on [6], while a summary of the main parameters is summarized in *Table 1*. The device structure is formed by an n+ in p-well diode with an enhancement layer and a virtual guard ring. It was designed targeting low breakdown voltage VBD<17V and optimal trade-off between high PDE>4.5%@940nm and low jitter tail FW@10%<310ps to maximize the overall dToF system performance. Deep trench isolation between the SPADs minimizes the optical cross talk.

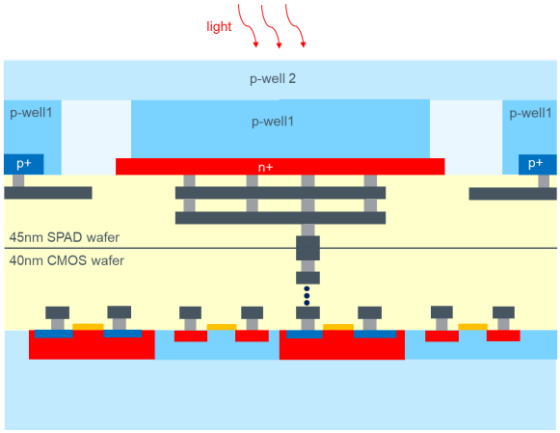

Figure 2: Schematic cross-section view of the 3D-stacked SPAD device and CMOS front-end.

IV. System Architecture

Figure 3 shows the overall dToF system IC block diagram consisting of three semiconductors components: the dToF sensor chip, the VCSEL driver ASIC and a 940-nm Vertical Cavity Surface Emitting Laser (VCSEL). The dToF sensor features 80x60 macro-pixels for a total of 320x240 SPAD pixels. The sensor can operate at different resolutions, from 80x60-pixel full parallel processing up to 320x240-pixel by enabling 1:16 SPAD multiplexing or any intermediate configurations. The MP information (time histogram, intensity counter) is processed via dedicated Processing Cluster Units (PCU) featuring dedicated CPU+RAM digital logic. Dedicated MP logic allows implementing different operating modes enabled via FW and proper interactions between MPs and respective PCUs. This combined hardware/firmware (HW/FW) architecture enables very flexible implementation of different algorithms, performing the following key functions necessary for dToF depth maps generation: i) run-time monitoring of photons arrival time histogram, ii) target peak detection based on programmable SNR threshold, iii) histogram decimation to reduce the amount of data transfer between MPs and PCUs, and iv) histogram-to-distance conversion via sub-bin interpolation. More details on histogram processing and sub-bin interpolation are available in [7]. A top-level master CPU/RAM handles the communication with PCUs and generates final complete depth+intensity images streamed out via CSI-2 MIPI interface. I2C interfaces are used for system/sensor configurations and host interrupts. Non-volatile flash memory is also available to store FW and calibration data.

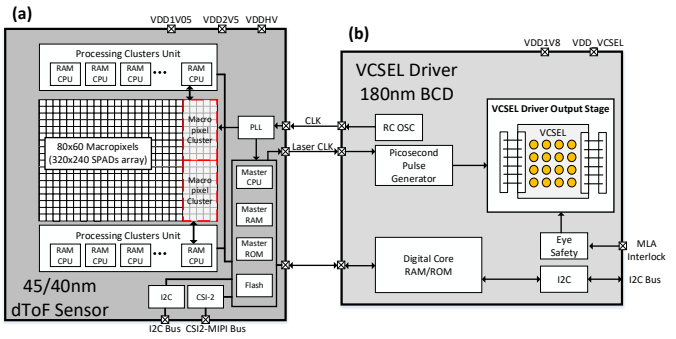

Figure 3: Block diagram of RX/TX ICs and relative interface: a) 45/40-nm BSI SPAD dToF sensor, b) 180-nm BCD VCSEL driver.

The dToF sensor is connected to a companion VCSEL driver ASIC developed on a 180-nm BCD process to enable high-voltage output driving stage. The main interfaces between the two ICs are: reference system clock (CLK), laser trigger (Laser_CLK) and digital interface used for driver registers configuration. Due to the severe requirements on target laser pulse width (<400ps) and high peak optical power (>35W) the VCSEL is stacked directly on top of the driver adopting an optimized layout to minimize parasitic interconnections inductance and fulfill electromagnetic compatibility. The driver implements several eye-safety features that automatically shut down the emission: i) electrical continuity monitor to verify the integrity of Micro-Lens-Array (MLA) optical diffuser, and ii) laser frequency and duty cycle within maximum-safety-range check. Finally, driver temperature monitoring is also implemented both for calibration and for device protection purposes.

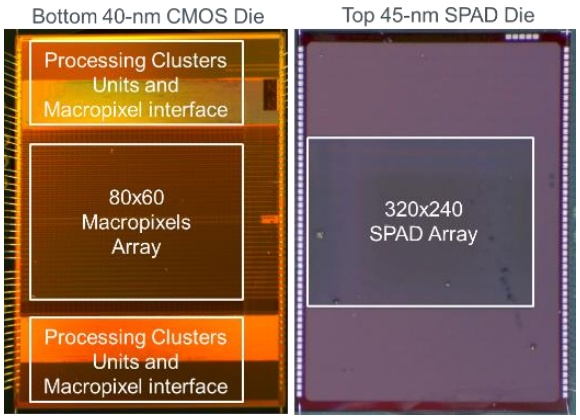

Figure 4: Top and Bottom die micrographs.

V. Measurement Results

The dToF sensor has been realized in a 45/40nm BSI CMOS 3D-stacked technology. Figure 4 shows the chip top- and bottom-micrographs with the main blocks highlighted. To enable robust detection of target peak versus noise peaks generated under strong ambient light condition, we have developed a TDC featuring very low DNL/INL distortions.
The TDC DNL/INL characterization as extracted from five macro- pixels (MPs) for a 210-ps LSB setting over 32 bins x 8 windows removing the first and last bins is shown in Figure 5.

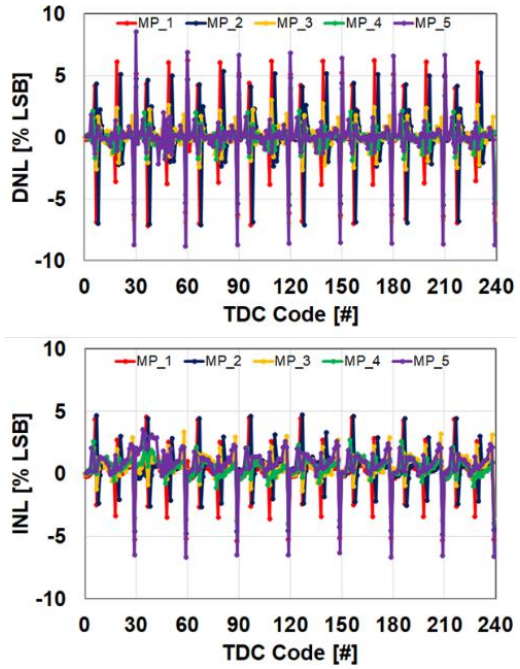

Figure 5: TDC DNL and INL extracted over five MPs.

The ICs have been integrated into a dual module flex-PCB prototype (Figure 6) in order to extract all the key system metrics based on realistic system assembly. The camera module consists of a 74.5-deg field-of-view (FOV) diagonal, F#1.4 objective and a 940-nm pass-band filter to mitigate ambient light noise. The laser module exploits an MLA to shape the VCSEL emitted beam into a uniform light beam matched with the rectangular-shape camera FOV.

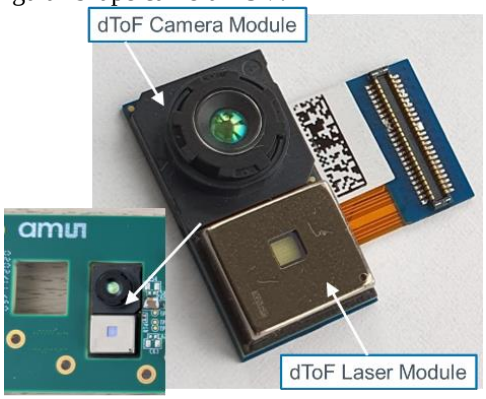

Figure 6: ams 3D dToF RX/TX System Prototype.

The TX-module performance has been extracted using combined power-meter and Time-Correlated-Single-Photon-Counting (TCSPC) setups to extract the average transmitted optical power and laser pulse time shape. The laser module is capable of generating pulses from 250ps to ~1ns FWHM. Figure 7 shows the typical TX-module optical power over time for a nominal FWHM=350ps setting. The peak optical power is extrapolated from the measured average optical power and laser pulse shape, being higher than 35W for FWHM~400ps. Despite the high value for the peak optical power, the system is fully Class-I.

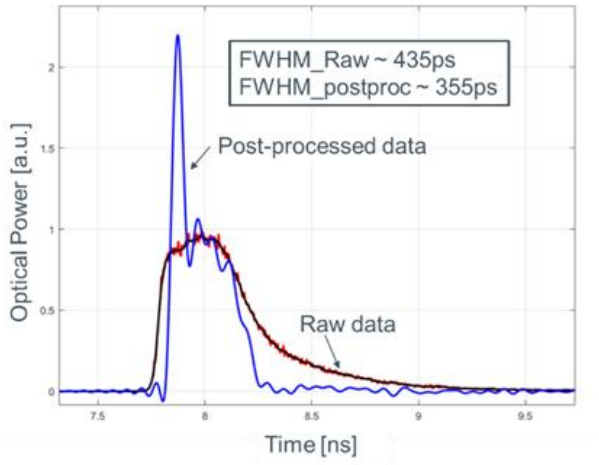

Figure 7: Single-pulse TX laser module optical power versus time acquired from TCSPC system (raw) and post-processed data obtained via de-convolution of TCSPC IRF.

Functional tests using automated stage flat targets featuring different reflectivity of 18% and 92% scanned over 500mm-6000mm distance range, and for different levels of ambient light, have been performed to extract the overall system performance. Depth accuracy, precision, and number of non-detect pixels as extracted over a 5x5 MPs central region-of-interest for Q3VGA and QVGA resolutions are shown in Figure 8 and Figure 9 respectively. The vertical dotted lines mark the maximum distance range at which the percentage of non-detects is <1% (as for QVGA 50klux plot) or at the end of the last useful time window compatible with the target frame rate (all other plots).

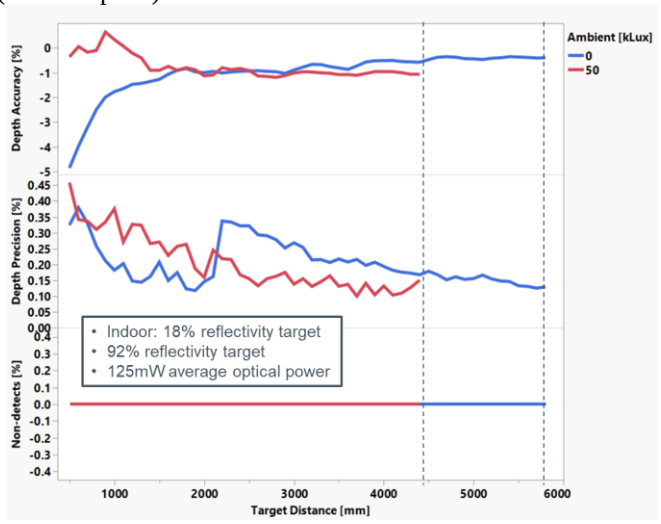

Figure 8: Q3VGA flat-target distance performance at 30fps.

As an example, the maximum detectable distance for the Q3VGA resolution case shown in Figure 8 is 4.4m/5.8m for 50klux/indoor respectively, however, this is not limited by an increase of the percentage of non-detects (which is still zero). The bottleneck in this case is the available overall exposure time, in fact detecting a target beyond these distances would require adding an extra time window, however, this is not possible without decreasing the frame rate below 30fps.

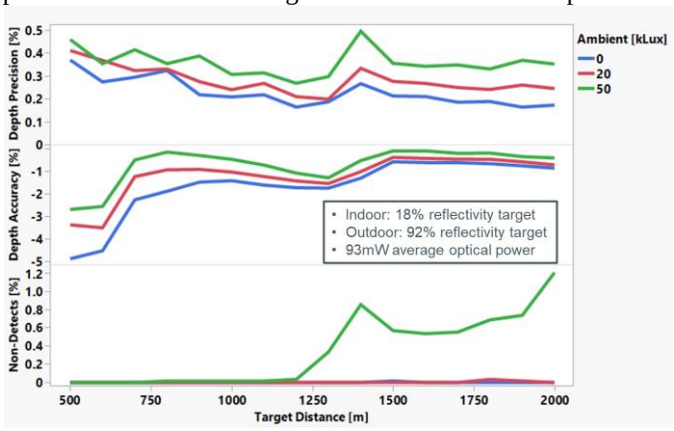

Figure 9: QVGA flat-target distance performance at 5fps.

The effective average optical power used for these measurements changes with exposure time and frame rate and it is reported in the figure insets for each operating modes.
All depth data and images are raw depth data fully processed on chip in real-time without any further external post-processing. On-chip TDC bin digital calibration and one-shot offset calibration is applied.
Finally, Figure 10 shows some examples of real complex scenes depth maps acquisitions. Depth images are single frame snapshots from 5-fps raw data as computed on-chip, while high-resolution RGB pictures are from off-the-shelf camera provided for reference.

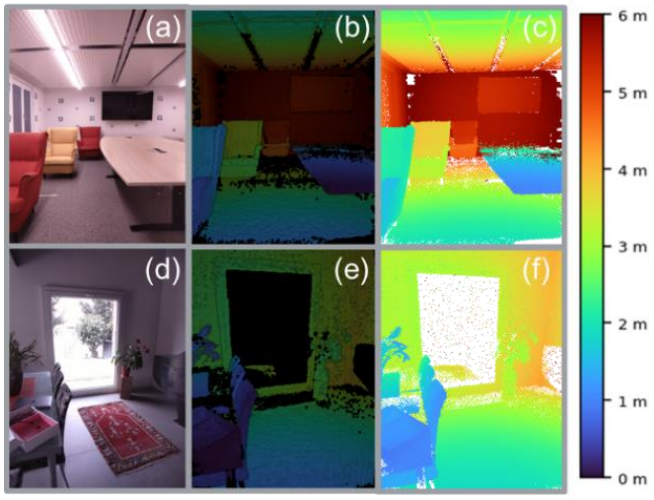

Figure 10: RGB high resolution scene-images (a,d); QVGA 3D points cloud (b,e) and 3D color maps (c,f).

VI. Conclusions

A complete dToF system, consisting of a 320x240-SPADs array dToF receiver (RX) camera and a dToF-optimized laser module transmitter (TX) is presented. The dToF sensor architecture, based on multiple CPUs processing, allows run-time interaction with the macro-pixel array, enabling advanced processing functions that can be flexibly configured via FW programming and allows on-chip full depth map computation and calibration. This architecture represents an ideal platform to explore and validate different dToF operating modes and algorithms, enabling shorter time-to-market development. The full system has been integrated within a compact form-factor package, and validated through functional tests to extract the main system performance. Table 1 shows a summary of the key dToF system design and performance parameters.

Table 1: Performance summary.

	Parameter	Unit	Value	Comment
General	dToF Sensor Technology	-	45nm BSI SPAD stacked on 40nm Flash CMOS	
	Laser Driver Technology	-	180nm BCD CMOS	
	Resolution (HxV)	-	320x240 SPADs array	
	Parallel processing channels	-	80x60 macropixels	
	FoV diagonal	deg	74.5	
	F# RX Optics	-	1.4	
SPAD	SPAD pitch	µm	~12.5	
	Breakdown voltage	V	17.0	
	Median DCR	cps	<270	@75C, Vex=2V
	PDE@940nm	%	>4.5	
	Jitter, FWHM@940nm	ps	145	
	Jitter, FW10%M@940nm	ps	310	
TDC	TDC depth	bit	32 x 8	32 bit per time-window, 8 time-windows
	TDC resolution	ps	~200 - 300	Programmable
	DNL / INL	% LSB	±9 / ±6	Expressed as % of TDC LSB
Laser TX	Wavelength	nm	940	Typical @ RT
	Laser pulse width, FWHM	ps	~360	After de-convolution of IRF of testing equipment
	Optical peak power	W	>35	
System Perf.	Distance range QVGA indoor / @50klux	m	2.0	@<1% non-detect pixels, 92% refl. target, @5fps
	Distance range Q3VGA indoor / @50klux	m	5.8 / 4.4	@<1% non-detect pixels, 18% reflectivity target for indoor, 92% target for 50klux, @30fps
	Total System power consumption	W	<1.5	@30fps, depth map fully processed on-chip
	Average optical power	mW	93/125	QVGA/Q3VGA

Acknowledgments

We would like to thank A. Biber, Q. Zhang, N. Rossi and G. Colzi for their support on functional tests and data processing, and F. Moulin for complex scenes images acquisition.

References

[1] T. Al Abbas, et al., "Backside illuminated SPAD image sensor with 7.83µm pitch in 3D-stacked CMOS technology," 2016 IEEE International Electron Devices Meeting (IEDM), 2016, pp. 8.1.1-8.1.4.
[2] A. R. Ximenes, et al., "A 256×256 45/65nm 3D-stacked SPAD-based direct TOF image sensor for LiDAR applications with optical polar modulation for up to 18.6dB interference suppression," *Proc of ISSCC* 2018, pp. 96-98,
[3] R. K. Henderson et al., "A 256x256 40nm/90nm CMOS 3D-stacked 120dB Dynamic Range Reconfigurable Time Resolved SPAD Imager", Proc. of ISSCC 2019.
[4] P. Padmanabhan et al., "A 256×128 3D-Stacked (45nm) SPAD FLASH LiDAR with 7-Level Coincidence Detection and Progressive Gating for 100m Range and 10klux Background Light", Proc. of ISSCC 2021.
[5] O. Kumagai et al., "A 189x600 Back-Illuminated Stacked SPAD Direct Time-of-Flight Depth Sensor for Automotive LiDAR Systems", Proc. of ISSCC 2021.
[6] G. Roehrer et al., "3D-Stacked SPAD in 40/45nm BSI Technology", Proc. of ISSW'20.
[7] P. Padmanabhan et al., "Depth precision in dToF sensors for AR applications", P26 to be presented at IISW'21.