From: C. Bamji et al., "1Mpixel 65nm BSI 320MHz Demodulated TOF Image Sensor with 3.5µm Global Shutter Pixels and Analog Binning," *ISSCC Deg. Tech. Papers*, pp. 94-95, Feb 2018.
IEEE Explore Link: https://ieeexplore.ieee.org/document/8310200/

5.8 1Mpixel 65nm BSI 320MHz Demodulated TOF Image Sensor with 3.5µm Global Shutter Pixels and Analog Binning

Cyrus S Bamji, Swati Mehta, Barry Thompson, Tamer Elkhatib, Stefan Wurster, Onur Akkaya, Andrew Payne, John Godbaz, Mike Fenton, Vijay Rajasekaran, Larry Prather, Satya Nagaraja, Vishali Mogallapu, Dane Snow, Rich McCauley, Mustansir Mukadam, Iskender Agi, Shaun McCarthy, Zhanping Xu, Travis Perry, William Qian, Vei-Han Chan, Prabhu Adepu, Gazi Ali, Muneeb Ahmed, Aditya Mukherjee, Sheethal Nayak, Dave Gampell, Sunil Acharya, Lou Kordus, Pat O'Connor

Microsoft, Mountain View, CA

The quest for accurate, high resolution, low power consumption and small footprint 3D depth cameras has driven a rapid improvement in Continuous Wave (CW) Time-of-Flight (ToF) technology. Commercially available 3D image acquisition techniques include Stereo Vision, Structured Light, and ToF. CW ToF imaging systems offer excellent mechanical robustness, no baseline requirement, high effective depth image resolution, low computational cost, and simultaneous IR ambient light invariant intensity capture (Active Brightness). In a CW ToF camera, light from an amplitude modulated light source is backscattered by objects in the camera's field of view, and the phase delay of the amplitude envelope is measured between the emitted and reflected light. This phase difference is translated into a distance value for each pixel in the imaging array.

Considerable effort has been applied to improve the spatial resolution, accuracy, and operating range of CW ToF cameras while lowering power consumption [1-4]. Uncertainty, range and power consumption are improved by: increasing Modulation Contrast (MC), Quantum Efficiency (QE) and Modulation Frequency, and reducing: read noise and analog-to-digital conversions; optical stack height is reduced by smaller pixels. This paper presents a state-of-the-art ToF image sensor fabricated in a TSMC 65nm 1P8M backside illumination (BSI) CMOS technology. This 1024×1024 pixel ToF global shutter image sensor achieves a high MC of 87% at 200MHz. The small pixel size of 3.5µm×3.5µm is competitive with commercial global shutter RGB sensors and facilitates a small optical stack height in mobile devices. A DLL controlled clock distribution system spreads the column clock peak currents across the columns thus reducing overall chip peak currents allowing a high 320MHz modulation frequency. The readout circuitry is fully differential and supports per- pixel adaptive gain selection, analog Correlated Double Sampling (CDS), analog pixel binning modes (e.g. 2×2) and a 9 or 10 bit Analog-to-Digital Converter (ADC). These chip specifications set a new benchmark for CW ToF performance.

A new type of differential ToF pixel photodetector operates in a *quasi-digital* demodulation mode. In this scheme two polysilicon gates compete to collect photo-charges and the gate with a higher bias voltage captures almost all of the photo-charges. The gates also create a strong drift field allowing fast charge collection resulting in a high photodetector MC of 87% at 200MHz and 78% at 320MHz. QE is 44% at 860nm with a standard TSMC BSI process. Lower detector gate capacitance and voltage swing result in less than ½ the power consumption per unit area of our previous work [2]. Fig. 1 shows a simplified cut through the photodetector structure and the collected photo current when polysilicon gate PGB is at a higher potential than polysilicon gate PGA. Simulation shows almost all charges are collected by gate PGB. Pixel effective fill factor is close to 100% due to the use of BSI technology and an additional optimized micro-lens structure.

Fig. 2 presents the differential pixel schematic and timing diagram. The schematic includes two in-pixel memory storage elements which store collected photo charges as minority carriers suitable for analog CDS. The pixel layout has centroid symmetry, minimizing offsets and noise. *Global reset* clears charges from the gates and the memory elements. During *integration*, modulation gates PGA and PGB are driven with complementary column clocks and collected photo charges accumulate into in-pixel memories A and B.

A DLL-based clock driver system generates uniformly-time-spaced pixel column clocks to the sensor array avoiding the large peak current transients often generated by balanced clock trees. Each delay line element incorporates a feed forward component crossing from the A domain to the B domain, speeding up the delay performance. This achieves a guaranteed column-to-column time delay of 9.5ps while creating two delay paths (Clock A and Clock B that drive pixel gates PGA and PGB) with guaranteed orthogonality at speeds beyond 320MHz. As a result the clocking system provides enforced symmetric zero crossings with Process Voltage and Temperature (PVT) independent fast delay times and interpolated A to B to A spacing of 9.5ps.

After integration, global shutter functionality is enabled during the *anti-blooming* period to reduce shot noise from ambient light. Shutter efficiency is about 99.8%. During a *row readout*, kT/C noise is cancelled by first resetting the floating diffusions FDA, FDB and reading out a corresponding kT/C and offset value. Then the values stored in the in-pixel memories A, B are dumped into floating diffusions FDA, FDB and the final signal value is read out through bitlines A, B. The just read kT/C and offset value is then subtracted in the analog domain from this final value by the column readout circuitry and a single analog-to-digital conversion is performed on the analog subtracted value.

A block diagram of the readout circuit is shown in Fig. 3. The column multiplexer performs both pixel selection and electrical pixel binning functions. Six 1.2V, 4µA gain selection comparators simultaneously check both input signal polarities versus three decision thresholds to select one of four amplifier (AMP) gains. The column amplifier gain is programmed with adjustable input and feedback capacitors to allow gains from 0.25X- 24X. Analog CDS operation requires the amplifier to subtract the sampled pixel reset voltage from the pixel voltage before amplification while rejecting common-mode differences and cancelling amplifier offset. As shown in Fig. 3, during $\phi 1$ and $\phi 2$, the pixel reset and amplifier offset are sampled on C3 and C5 and the image data and amplifier offset are sampled on C4 and C6. During $\phi 3$ the charge on C3-C6 is moved to feedback capacitors C1 and C2.

Each AMP spans two pixel columns and two ADCs sample and convert the processed data from the AMP. Linearity of the 2GHz counter, 9 or 10-bit single-slope ADCs is improved with distributed feedback source followers in a global ramp generator. Differential read noise is 3e- (un-binned) in 10 bit mode. Each of the 512 readout circuits including pixel binning multiplexer, gain selection comparators, programmable gain amplifier, and two ADC's are implemented on a 7µm pitch with a height of about 1.2 mm and provide 3.4GS/s 9bit or 1.7GS/s 10bit digital data.

Fig. 4 shows the overall system accuracy and uncertainty at various sunlight equivalent ambient levels up to 25kLux as a function of distance. Fig. 5 compares the system in high performance and low power operating modes with our previous work [2] and prior art. Fig. 6 shows a 1024×1024 image of 24 people over 40cm-7m at 120^0×120^0 Field of View (FOV). The images are from a single frame of data for: Active Brightness, color coded depth (mm), depth point cloud and a depth point cloud zoomed in on a ping pong ball at 2m which still shows a good spherical depth shape. A die photo with labeled sub-blocks is shown in Fig. 7. We believe the presented ToF chip provides the performance, footprint and power envelope for demanding ToF scenarios including mobile devices.

References:

[1] S-M. Han, et al., "A 413×240-Pixel Sub-Centimeter Resolution Time-of-Flight CMOS Image Sensor with In-Pixel Background Canceling Using Lateral-Electric-Field Charge Modulators," *ISSCC Deg. Tech. Papers*, pp. 130-131, Feb 2014.

[2] A. Payne et al., "A 512×424 CMOS 3D time-of-flight image sensor with multi-frequency photo-demodulation up to 130MHz and 2GS/s ADC," *ISSCC Deg. Tech. Papers*, pp. 134-135, Feb 2014.

[3] Y. Kato et al., "320x240 back-illuminated 10µm CAPD pixels for high speed modulation Time-of-Flight CMOS image sensor," *2017 Symposium on VLSI Circuits*, pp. C288 - C289, June 2017.

[4] S. Kim, et al., "Time of Flight Image Sensor with 7µm Pixel and 640x480 Resolution," *2013 Symposium on VLSI Technology*, pp. T146 – T147, June 2013.

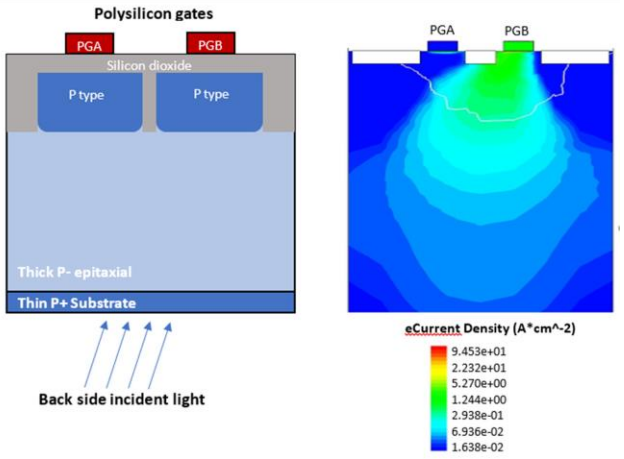

Figure 5.8.1: Pixel structure and collected photo-charges

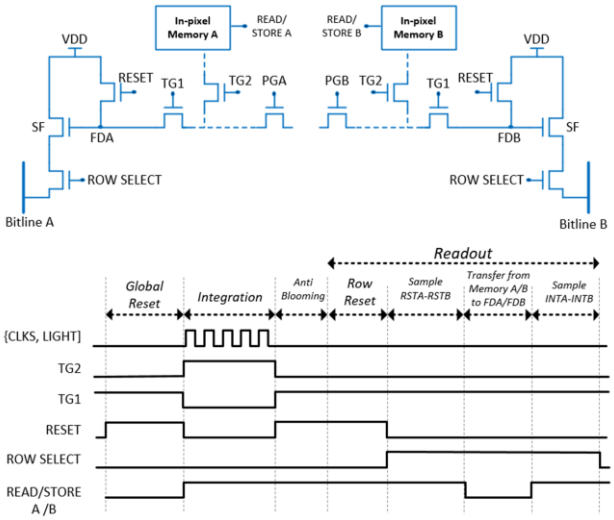

Figure 5.8.2: Pixel schematic and timing diagram

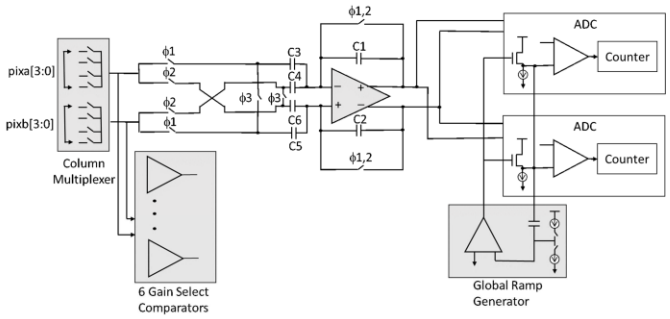

Figure 5.8.3: Schematic of Amplifier (AMP) and Analog-to-Digital Converter (ADC)

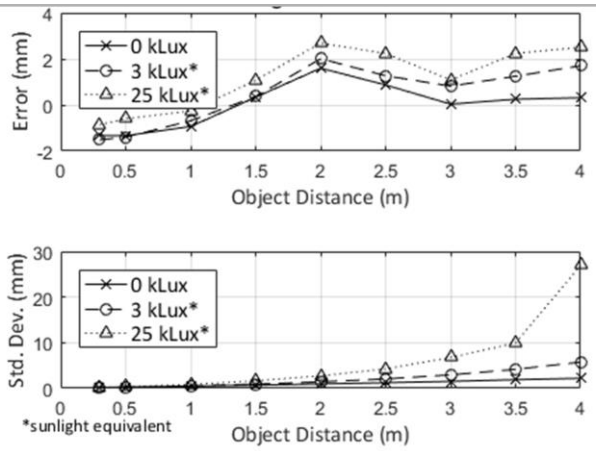

Figure 5.8.4: Measured system accuracy and uncertainty for 20% reflective target at varied ambient light levels in center region of field of view

	This work		Kinect v2 [2]	Sony 2017 [3]
Process Technology	TSMC 65nm BSI 1P8M		TSMC 0.13µm 1P5M	90nm BSI
Pixel Pitch	3.5µm × 3.5µm		10µm x 10µm	10 µm x 10 µm
Pixel Array	1024 × 1024 (active pixels)		512 x 424	320 x 240
Chip Size	5.4mm × 9.8mm		8.2mm x 14.2mm	
System Dynamic Range	>2500 = 68dB		>2500 = 68dB	
Modulation Contrast	87% @860nm @200MHz 78% @860nm @320MHz		68% @860nm @50MHz	85% @ 850nm @100MHz (thin epi)
Modulation Frequency	10 to 320MHz		10 to 130MHz	
System Average Modulation Frequency	200MHz		80MHz	100MHz
Responsivity	0.305A/W @860nm		0.144A/W	0.34A/W
FOV (H x V)	120° x 120°		70° x 60°	--
Depth Uncertainty	< 0.2% of range (<3kLux)		<0.5% of range	<1% of range
F#	1.2		1.07	--
Frame Rate	30 fps (>60 fps w/reduced resolution)		30 fps max 60 fps	
ADC resolution	Selectable 9 bits or 10 bits		10 bits	
ADC speed	3.4GS/s @ 9b 1.7GS/s @ 10b		2GS/s	--
Effective Fill Factor	~100%		60%	>80%
Reflectivity	15% to 100%		15% to ~95%	--
Operating Mode	Low Power	High Performance		
Operating Distance	0.2 to 1.8m	0.4 to 4.2m	0.8 to 4.2m	
Readout Resolution	512x512 (analog binning)	1024x1024		
Readout Noise	12e- (210µV)	3e- (210µV)	(320µV)	87e-
Chip Power	150mW	650mW	2.1W	

Figure 5.8.5: Comparison Table

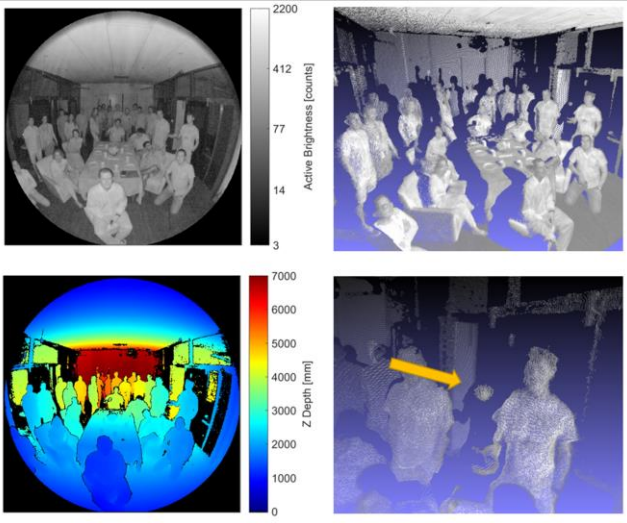

Figure 5.8.6: 1024×1024 Image: Active brightness, depth [mm], depth point cloud, depth point cloud zoomed in on bouncing ping pong ball

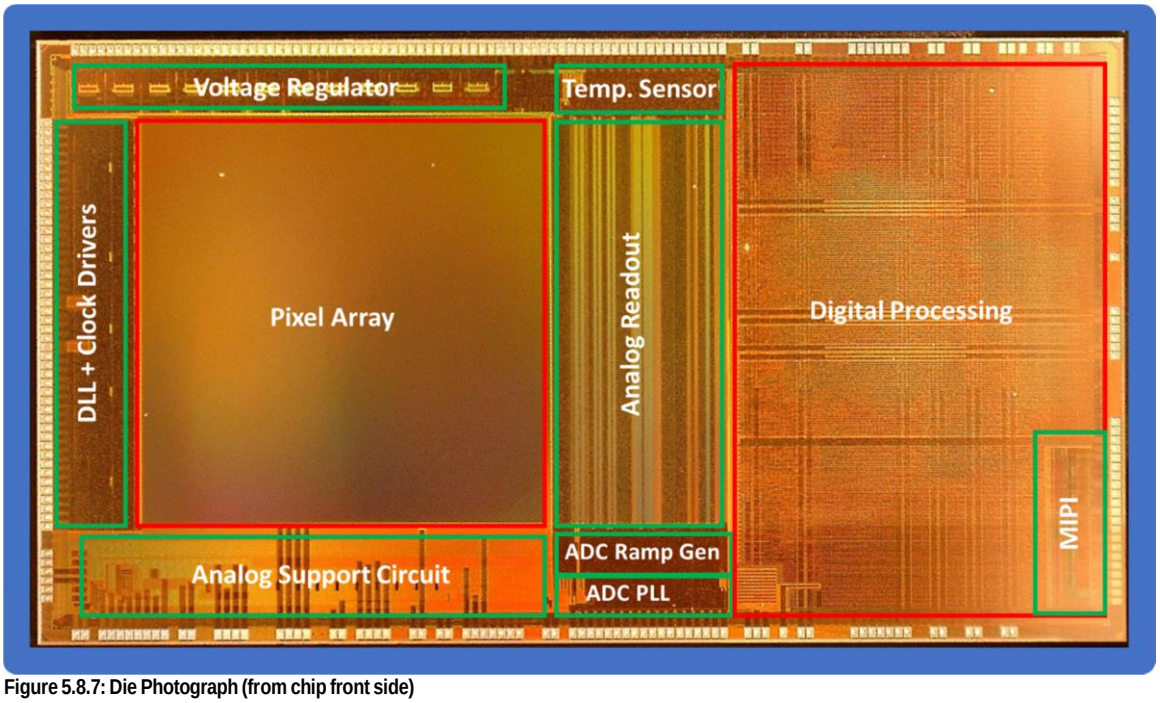

Figure 5.8.7: Die Photograph (from chip front side)