5.4 A 1/4-inch 3.9Mpixel Low-Power Event-Driven Back-Illuminated Stacked CMOS Image Sensor

Oichi Kumagai[1], Atsumi Niwa[1], Katsuhiko Hanzawa[2], Hidetaka Kato[1], Shinichiro Futami[1], Toshio Ohyama[1], Tsutomu Imoto[1], Masahiko Nakamizo[1], Hirotaka Murakami[2], Tatsuki Nishino[1], Anas Bostamam[1], Takahiro Iinuma[1], Naoki Kuzuya[1], Kensuke Hatsukawa[3], Frederick Brady[2], William Bidermann[2], Toshifumi Wakano[1], Takashi Nagano[1], Hayato Wakabayashi[2], Yoshikazu Nitta[1]

[1]Sony Semiconductor Solutions, Atsugi, Japan
[2]Sony Electronics, San Jose, CA
[3]Sony LSI Design, Atsugi, Japan

Wireless products such as smart home-security cameras, intelligent agents, and virtual personal assistants, are evolving rapidly to satisfy our needs. Small size, extended battery life, transparent machine interfaces: all these are required of the camera system in these applications. These applications, in battery-limited environments, can profit from an event-driven approach for moving-object detection. This paper presents a 1/4-inch 3.9Mpixel low-power event-driven (ED) back-illuminated stacked CMOS image sensor (CIS) deployed with a pixel readout circuit that detects moving objects for each pixel under lighting conditions ranging from 1 to 64,000lux. Utilizing pixel summation in a shared floating diffusion (FD) for each pixel block, moving object detection is realized at 10 frames per second while consuming only 1.1mW, a 99% reduction in power from the same CIS at a full-resolution 60fps power of 95mW.

Figure 5.4.1 shows a block diagram of the low-power ED sensor. The sensor consists of a pixel array, row drivers, row decoders, single-slope column-parallel 10b ADCs, a DAC for slope generation, motion detection (MD)/optical detection (OPD) blocks, an image signal processor, SRAM for frame memory, a Mobile Industry Processor Interface (MIPI), and a CPU connected to the sensor control block. A phase-locked loop (PLL) is used to generate all internal clocks from a single master input clock running between 12 and 27MHz. The clock generation circuitry is programmable through an I2C interface. The sensor is operated by three power supplies: a 1.8V pixel supply, a 1.0V digital supply, and a 1.8V supply for both analog and low-power I/O. The pixel array consists of visible pixels and optically black pixels used in for on-chip image processing. The array has vertical signal lines (VSLs), used for readout of each 2×4 shared pixel unit. The analog pixel readout signal is converted to a digital signal using a column-parallel ADC. The MD and OPD blocks are used while in the sensing binning mode. In this mode, a very low resolution of 80 binned pixel units (1 output for each 16×5 pixel subarray) are read out on a linear scale. If a moving object is detected, the CPU generates an external interrupt signal and triggers the capture of a high-quality image using on-chip auto exposure (AE) with zero latency. The maximum output data rate is 1.6Gb/s per MIPI lane. The chip incorporates 2 MIPI lanes. The sensor includes various camera functions such as sensing binning mode, pixel binning mode, all pixel scan image mode and is capable of recording full-HD movies with high sensitivity at 60 fps.

Figure 5.4.2 shows the overall readout architecture in sensing binning mode. The sensor consists of an array of 2560(H) × 1536(V) pixels. Each pixel in the array is addressed through a horizontal word line and VSLs for each 2×4 shared pixel unit readout. Pixel readout is performed one row at a time through the pixel source follower (SF) amplifier. The entire row is then converted to digital output using the column comparators and counters. The sensor supports 2×1 analog binning (2-column binning, called SF binning mode), 2×2 analog binning (2-column/2-row binning, called SF/FD binning mode) and 160×154 analog/digital binning (8-column/154-row SF/FD binning and 20 horizontal digital binning, called sensing binning mode). Binning can be enabled when the sensing binning mode is enabled (SMEN = 1). In this mode, all pixels are read through vertical floating diffusion lines (VFLs) to avoid blind-spots [1]. A combination of short-exposure pixels and long-exposure pixels alternating every two lines is used to achieve 160×154 pixel analog/digital binning with high dynamic range [2]. To reduce power consumption, each column's ADC is controlled by column-enable signals (COLEN).

There are two new functions on-chip: an MD function used to detect a change of object position relative to its surroundings, and an auto exposure (AE) function executed 1 frame before normal mode operation. In the ED system of this sensor, MD and AE are used in sensing binning mode. Figure 5.4.3 shows the MD and AE processes in this sensor. There are two event-driven methods employed: a frame difference method, and a background difference method. The MD function is used to minimize power and trigger a full resolution capture. In the MD process, moving objects are detected and the luminance difference between 2 frames of images are used as the detection trigger if the luminance difference exceeds a certain threshold. If a moving object is detected, the sensor generates an interrupt signal to wake up an application processor (AP). The AE is used to determine the appropriate exposure parameters: shutter exposure time and gain. The value of OPD is used to calculate current illuminance level (evaluation value); then the required exposure value to achieve brightness target called AE scale is calculated before the AE scale value is input to AE diagram. The appropriate value of shutter exposure time and gain is obtained by scaling the AE value from the minimum gain point in the AE diagram. The calculated exposure time and gain value are applied to the next normal readout frame continuously after moving object is detected.

As shown in Fig. 5.4.4, the sensing binning mode and normal mode are successfully captured using the fabricated ED sensor. The first image captures the entire image with low resolution and high contrast. With high accuracy, a moving object is detected in sensing binning mode and is followed by a high-quality image in normal mode.

A performance summary and comparison to recently published devices are shown Fig. 5.4.5. The chip consumes 95mW of total power at a frame rate of 60fps in normal mode achieving an FOM1 of 0.7e-·nJ for the ADC. Furthermore, the ADC demonstrates an FOM2 of 1.5e-·nJ/DRU. The chip characteristics are summarized in Fig. 5.4.6. The top chip uses a 90nm 1Al-4Cu CMOS process with specialized add-on steps to enable backside illumination. The bottom chip uses a 40nm 1Al-6Cu logic process technology. The total number of effective pixels is 3.9MPixels (2560 (H) × 1536 (V)), excluding the optically black pixels. The pixel pitch is 1.5µm and incorporates an on-chip micro-lens. A saturation signal of 7800e- at 60°C, a sensitivity of 8033 e-/lx·s, dynamic range of 67dB at 10b in normal mode and dynamic range of 96dB in sensing binning mode are all measured. An RMS random noise of 1.8e- and a conversion gain of 55.8µV/e- are confirmed. The die micrograph is shown in Fig. 5.4.7. The chip size is 4.959mm (H) x 4.401mm (V).

In conclusion, this sensor has an intelligent sensing function for real-time moving object detection within predefined areas. This low-power ED sensor facilitates event recording, which significantly reduces power consumption and data bandwidth of camera systems while in low-power sensing mode. These features enhance the device usability and satisfies the new demands of a low-resolution always-on sensing device that is also capable of high-quality imaging.

Acknowledgments:
The authors would like to thank the members of Sony Semiconductor Manufacturing, Sony Electronics Inc., Sony LSI Design Inc. and Sony Semiconductor Solutions Co. for their support of this work.

Reference:
[1] G. Kim, et al., "A 467nW CMOS Visual Motion Sensor with Temporal Averaging and Pixel Aggregation," *ISSCC Dig. Tech. Papers*, pp. 480-481, Feb. 2013.
[2] S.K. Nayar, et al., "High Dynamic Range Imaging: Spatially Varying Pixel Exposures," *IEEE Conf. Computer Vision and Pattern Recognition*, vol. 1, pp. 472-479, June 2000.
[3] Y. Chae, et al., "A 2.1Mpixel 120frame/s CMOS Image Sensor with Column-Parallel ΔΣ ADC Architecture," *ISSCC Dig. Tech. Papers*, pp. 394-395, Feb. 2010.
[4] A. Suzuki, et al., "A 1/1.7-inch 20Mpixel Back-Illuminated Stacked CMOS Image Sensor for New Imaging Applications," *ISSCC Dig. Tech. Papers*, pp. 110-111, Feb. 2015.
[5] T. Arai, et al., "A 1.1µm 33Mpixel 240fps 3D-Stacked CMOS Image Sensor with 3-Stage Cyclic-Based Analog to-Digital Converters," *ISSCC Dig. Tech. Papers*, pp. 126-127, Feb. 2016.
[6] M. Kobayashi, et al., "A 1.8erms Temporal Noise Over 110dB Dynamic Range 3.4µm Pixel Pitch Global Shutter CMOS Image Sensor with Dual-Gain Amplifiers, SS-ADC and Multiple-Accumulation Shutter," *ISSCC Dig. Tech. Papers*, pp. 74-75, Feb. 2017.

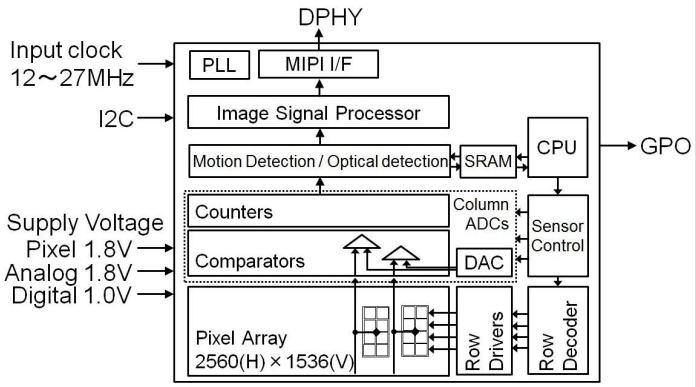


Figure 5.4.1: Block diagram.

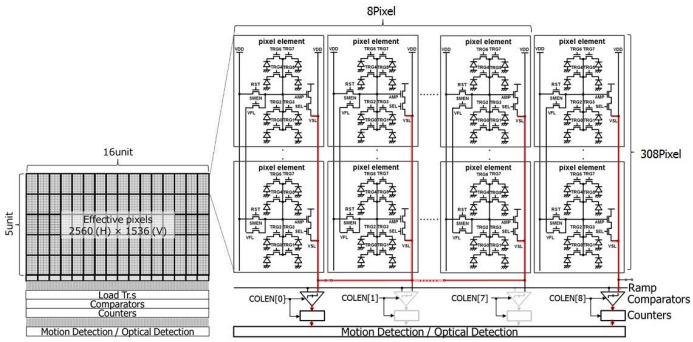


Figure 5.4.2: Readout architecture for sensing binning mode.

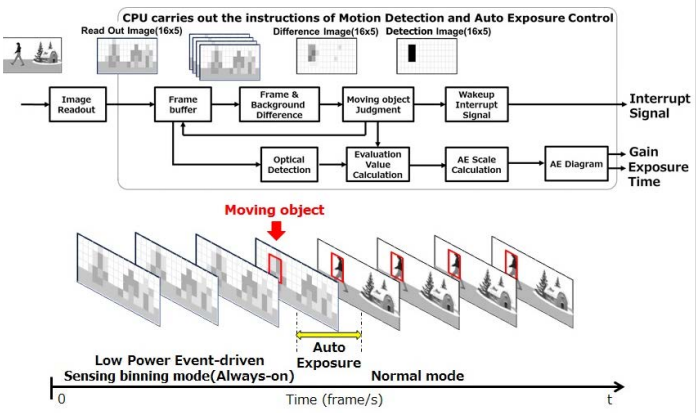


Figure 5.4.3: Mode transition from sensing binning mode to normal mode with auto exposure.

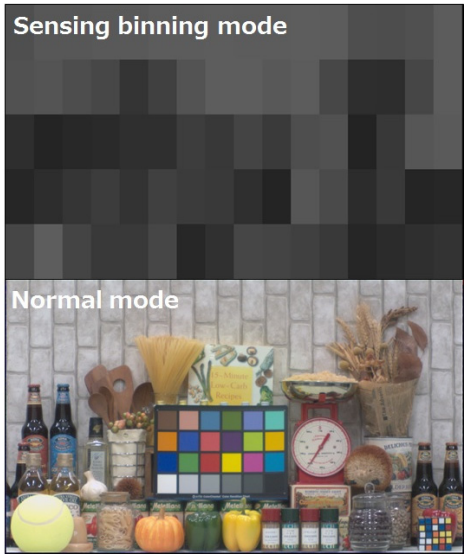


Figure 5.4.4: Captured images.

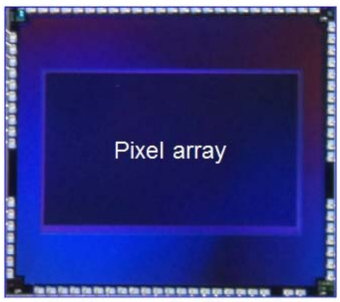
	ISSCC 2010 [3]	ISSCC 2015 [4]	ISSCC 2016 [5]	ISSCC 2017 [6]	This work
ADC resolution [bit]	12.5	12	12	10	10
Frame rate [fps]	120	30	240	120	60
# of V pixels	1212	3934	4320	2054	1536
# of H pixels	1696	5256	7680	2592	2560
Random noise [e⁻rms]	2.4	1.3	3.6	1.8	1.8
Power [mW]	180	532	3000	450	95
FoM1 [e⁻·nJ] for ADC	1.8	1.1	1.4	1.3	**0.7**
FoM2 [e⁻·nJ/DRU] for ADC	0.4	3.3	3.4	0.3	1.5

FoM1=Power[W]xNoise[e^-]x10^9/FPS[s^{-1}]xNum. of Eff. Pixels
FoM2=Power[W]xNoise[e^-]x10^{12}/FPS[s^{-1}]xNum. of Eff. Pixels x DRU
DRU=Saturation signal[e^-]/Noise[e^-]/Gain[times]

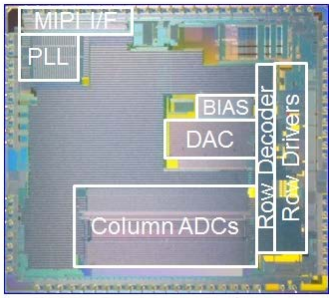
Figure 5.4.5: Performance comparison.

Fabrication Process		90nm 1AL4Cu CIS/40nm 1AL6Cu Logic
Number of effective pixels		2560 (H) × 1536 (V) 3.9 M pixels
Image size		Diagonal 4.48 mm (1/4-type)
Pixel size		1.5 µm (H) × 1.5 µm (V)
Supply voltage		1.8V / 1.8V / 1.0V
Frame rate	All-pixel scan	10 bits 60 fps
	Full HD	10 bits 60 fps
	Sensing 16x5	10 bits 10 fps
Power consumption	All-pixel scan	95 mW at 10 bits 60fps 57 mW at 10 bits 30fps
	Full HD	71 mW at 10 bits 60 fps 41 mW at 10 bits 30 fps
	Sensing 16x5	1.1 mW at 8 bits 10 fps
Saturation signal		7800 e- at 60ºC
Sensitivity(Typical value F5.6)		8033 e-/lx·sec(Green Pixel)
RMS random noise		1.8 e-rms (analog gain:18dB)
Dynamic range	All-pixel scan	67 dB at 10 bits
	Sensing 16x5	96 dB (1~64klx)
Conversion Gain		55.8µV/e-
Input clock frequency		6~27 MHz
Image output format		Bayer RAW
Outputs		MIPI(CSI-2) 2 lane DPHY 1.6 Gbps/lane

Figure 5.4.6: Chip characteristics.



A micrograph of the chip



A micrograph of bottom part

Figure 5.4.7: Die micrograph.

978-1-5090-4940-0/18/$31.00 ©2018 IEEE