27.2 A 467nW CMOS Visual Motion Sensor with Temporal Averaging and Pixel Aggregation

Gyouho Kim, Mahmood Barangi, Zhiyoong Foo, Nathaniel Pinckney, Suyoung Bang, David Blaauw, Dennis Sylvester

University of Michigan, Ann Arbor, MI

Visual monitoring with CMOS image sensors opens up a variety of new applications for wireless sensor nodes, ranging from military surveillance to *in vivo* molecular imaging. In particular, the ability to detect motion can enable more intelligent power management through on-demand duty cycling and reduced data-retention requirements. Conventional imager designs focus on achieving higher resolution, frame rate [1], or dynamic range [2], resulting in power consumption levels that are unsuitable for battery-powered wireless sensor nodes [3].

Several in-pixel motion-detection (MD) designs have been proposed [4,5], in which the previous pixel value is stored on an in-pixel capacitor until the end of the next integration cycle for immediate frame differencing. This avoids the need for high-power DSP. However, these designs implement MD in all pixels and still consume mWs of power. In addition, the in-pixel schemes are limited to frame differencing of two consecutive frames, reducing sensitivity to slow-moving objects compared to more sophisticated DSP approaches that operate on multiple frames.

To capitalize on the low-power aspect of in-pixel frame differencing without compromising sensitivity to slow-moving objects, we propose *temporal averaging* (TA), where an additional integration time is used for certain pixels in the array. The intuition behind this is that slow objects make negligible differences at high frame rates but can be detected at slow frame rates. We show that interleaving two integration times in one array increases the range of detectable motion by 10× (Fig. 27.2.1, top right). Secondly, to further reduce power consumption and limit the impact of MD on pixel fill factor, only a subset of the pixels is instrumented with MD (3 of 64 in our design). As a result, power consumption is reduced by 20×, but so-called blindspots are created due to the presence of inactive pixels, which allow small objects to escape detection (Fig. 27.2.1, bottom). To address this we propose *pixel aggregation* (PA), where multiple pixels are combined and operate as one to increase coverage by 6× with no power penalty. Finally, we achieve significant power savings by biasing analog tail currents at subthreshold, operating digital components in the near-threshold region, and clock-gating high-speed blocks. Combining these techniques we demonstrate an imager with in-pixel motion detection showing high sensitivity to low-speed motion and a power consumption of only 467nW, marking a 400× reduction over prior art (at same fps and normalized resolution, see comparison table in Fig. 27.2.6) and making continuous motion detection practical for low-power wireless sensor nodes.

The proposed sensor array consists of 128×128 pixels, with groups of 8×8 pixels forming an MD cluster (Fig. 27.2.2). To minimize the area overhead of in-pixel motion detection, the MD circuitry is distributed within the cluster across its 64 pixels, resulting in a pixel fill factor of 38%. Within each MD cluster, two TA cells and one PA cell are placed in an interleaved fashion, resulting in an overall 32×16 TA array and 16×16 PA array. TA is implemented by increasing the integration capacitance by 3× and extending the integration period in the frame controller. PA is implemented by charge-sharing photodiodes at the circuit level.

There are four types of pixels in the array: base, TA-SHA, PA-SHA, and PA-COMM (Fig. 27.2.2). Figure 27.2.3 shows the pixel and column peripheral schematics. The base pixel uses a conventional 3-T structure with reset device M0, source follower input device M1, and column line access device M2. A 15.6μm^2 psub/n+ parasitic diode is used as the photodiode. The base pixel is used only for regular imaging, and its spare layout area is shared for capacitance distribution. The TA detection cell consists of a TA-SHA pixel, an explicit integration capacitor, C_{TAVG}, and C_{HOLD}. C_{TAVG} is required to adjust the integration capacity for longer exposure time. The PA detection cell consists of a PA-SHA pixel, PA-COMM pixels, and C_{HOLD}.

TA-SHA and PA-SHA pixels include M3-5 and C_{HOLD} to retain the previous frame's pixel value. Subthreshold leakage through M3 is the primary leakage source for V_{HOLD}; hence SMP is pulled low to –200mV to super-cutoff M3. Simulation shows a maximum leakage-induced droop of 5mV for 200ms (<1% of signal range). C_{HOLD} for TA unit is 3× larger than in the PA unit, in accordance with the integration period ratio. All explicit capacitors are distributed in the cluster, with a unit capacitance value of 25fF. Out of 61 (64 - 3 SHA pixels) available shared slots, 24×2 are used for TA C_{HOLD}, 3×2 for TA C_{TAVG}, and 7 for PA C_{HOLD}. M8-9 connect PA photodiodes to the cluster's charge sharing network, V_{CSN}. Up to 4×4 PA-COMM pixels can be selectively aggregated with PA-SHA per cluster. All devices in the array, including capacitors, are thick-oxide I/O devices to minimize gate and subthreshold leakage.

Column readout uses the n-type source follower M1, whose output is sampled by M12 onto C_{SMP} when COL_EN is high. For columns with MD units (3 per 8), additional column peripheral circuitry including M14-15 is added. During MD mode, the previous pixel value on C_{HOLD} is buffered through a p-type source follower M4, and the current pixel value is buffered twice through M1 and M15 to provide the same common mode. The resulting two analog signals, V_{PREV} and V_{CUR}, feed into the MD comparator to determine the presence of motion. The only mismatch that must be considered between V_{PREV} and V_{CUR} arises from process variation between M4 and M15, and is addressed with an offset-cancellation scheme (Fig. 27.2.4). A 9b single-slope ADC is implemented per column to capture regular images, and is only used during imaging mode.

The timing diagram for offset cancellation scheme and MD comparison are shown in Fig. 27.2.4. When one integration period is complete, the MD controller and 250kHz clock are enabled. The source followers of Row [i] are enabled by EN[i] and MD_EN, after which the difference between V_{PREV} and V_{CUR} is sampled onto C_1 by ϕ1. When SMP[i] goes high, the previous pixel value is overwritten by the current pixel value, and V_{CUR}-V_{PREV} now represents the V_{th} mismatch between M4 and M15, which is sampled onto C_2. During ϕ3 the MD comparison occurs, with C_1 and C_2 in series subtracting out the V_{th} mismatch. Coupling capacitors C_{C1-2} and pulses P_{1-3} are used to set the motion threshold and latch MD output, as shown in Fig. 27.2.4 (bottom right). The MD output (*Motion*) triggers if $|V_{CUR}-V_{PREV}|$ is greater than the coupled voltage from C_{C1-2}. After marching through 16 rows, ϕ5 cuts off static power through the MD comparator and the 250kHz clock is disabled until the subsequent integration finishes. TA and PA units are separately controlled and can operate simultaneously with different frame rates since the column readout structure is independent with its own peripherals.

The proposed design is fabricated in a logic 130nm 8M1P CMOS technology. Figure 27.2.5 shows measured results. In MD mode, the sensor consumes 467nW at 5fps with both TA and PA enabled. In imaging mode, the chip consumes 16μW at 6.4fps. Experiments show that TA cells are effective for motions slower than 70 pixels/s, boosting the detection level by up to 42%. PA cells capture moving objects smaller than 2 pixels wide, providing nearly complete visual coverage despite the use of sub-arrays for detection.

References:

[1] Y. Tochigi, *et al.*, "A Global-Shutter CMOS Image Sensor with Readout Speed of 1Tpixel/s Burst and 780Mpixel/s Continuous," *ISSCC Dig. Tech. Papers*, pp. 382-383, Feb. 2012.

[2] D. Stoppa, *et al.*, "A 120-dB Dynamic Range CMOS Image Sensor With Programmable Power Responsivity," *IEEE J. Solid-State Circuits*, vol. 42, no. 7, pp. 1555-1563, July 2007.

[3] Y. Lee, *et al.*, "A Modular 1mm^3 Die-Stacked Sensing Platform with Optical Communication and Multi-modal Energy Harvesting," *ISSCC Dig. Tech. Papers*, pp. 402-403, Feb. 2012.

[4] Y. M. Chi, *et al.*, "CMOS Camera with In-Pixel Temporal Change Detection and ADC," *IEEE J. Solid-State Circuits*, vol. 42, no. 10, pp. 2187-2196, Oct. 2007.

[5] P. Lichtsteiner, *et al.*, "A 128×128 120 dB 15 μs Latency Asynchronous Temporal Contrast Vision Sensor," *IEEE J. Solid-State Circuits*, vol. 43, no. 2, pp. 566-576, Feb. 2008.

[6] S. Hanson and D. Sylvester, "A 0.45-0.7 V Sub-Microwatt CMOS Image Sensor for Ultra-Low Power Applications," *IEEE Symp. VLSI Circuits* , pp.176-177, June 2009.

 978-1-4673-4516-3/13/$31.00

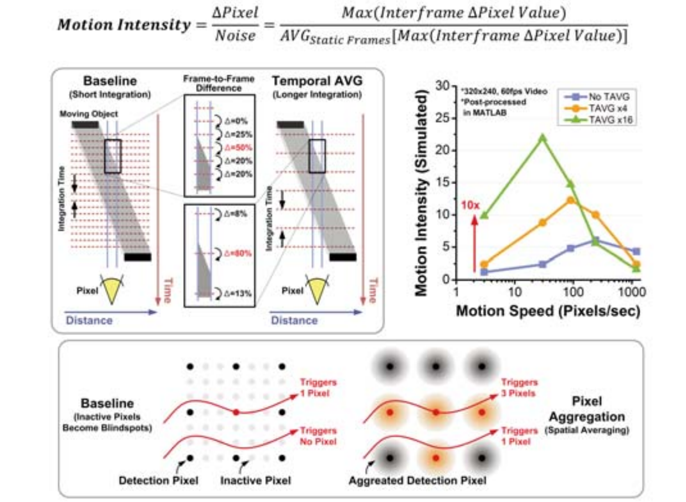

Figure 27.2.1: Temporal averaging (TA) increases sensitivity to slow motions (conceptual diagram at top-left, simulation results at right), while pixel aggregation (PA) reduce blindspots (bottom).

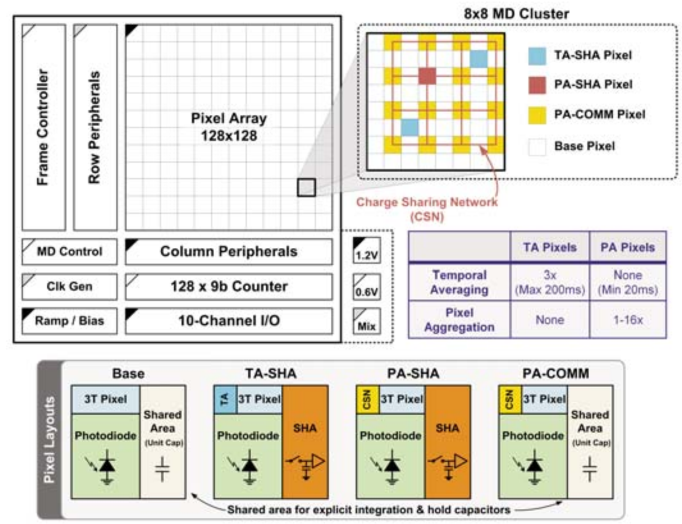

Figure 27.2.2: System block diagram showing pixel placements within a motion detection (MD) cluster.

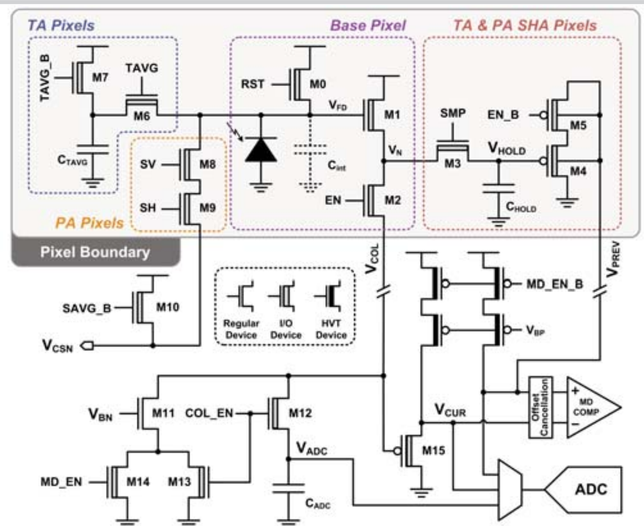

Figure 27.2.3: Pixel and column schematics. Different pixels have different add-ons to the base 3-T pixel.

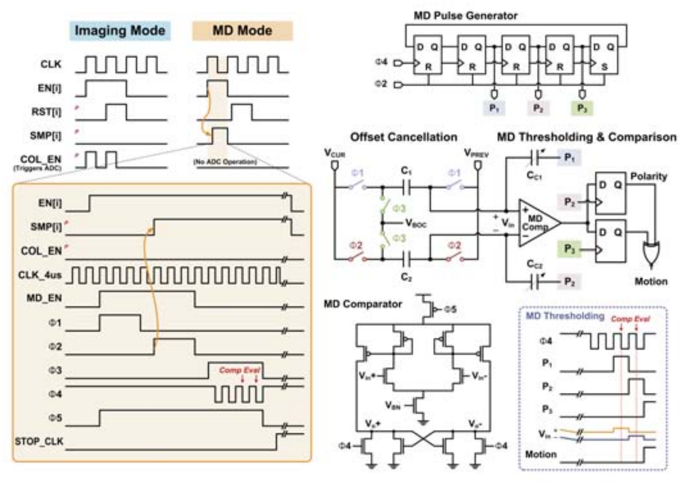

Figure 27.2.4: Timing diagram of readout scheme and schematics of offset cancellation and MD thresholding circuit. The example scenario shows motion being detected (bottom right).

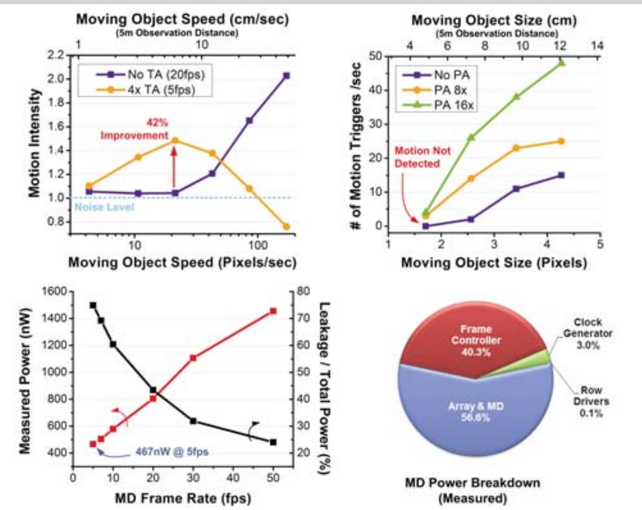

Figure 27.2.5: Measurement results. TA pixels are shown to be effective for slow motion (top left). With PA turned off, objects smaller than 7cm at 5m away can escape detection entirely (top right).

	[4]	[5]	[6]	This Work
Technology	0.5μm CMOS 3M2P	0.35μm CMOS 4M2P	0.13μm CMOS 8M1P	0.13μm CMOS 8M1P
Array Resolution	90 x 90	128 x 128	128 x 128	128 x 128
MD Resolution	90 x 90	128 x 128	Imager Only	48 x 16 (Up to 64x72 pixels)[3]
Pixel Size	25 x 25 μm²	40 x 40μm²	5 x 5 μm²	6.4 x 6.4 μm²
Fill Factor	17%	9.4%	32%	38%
FPN	0.5%	2.1%	6.6%	2.3%
Dynamic Range	51dB	120dB	23.4dB	38.5dB
Imaging Power	4.2mW @ 30fps	23mW	1.2μW @8.5fps	16μW @6.4fps 29μW @19fps
Imaging Energy/fr[1]	280μJ/frame	Frame Rate Variable	140nJ/frame	1.5μJ/frame
MD Power[2]	398μW @ 30fps	1.1mW	Imager Only	467nW @5fps 1.1μW @30fps
Supply Voltage	3V	3.3V	0.5V	1.2V / 0.6V
Area	9mm²	37.8mm²	1.1mm² (Core)	2.4 mm²(Core)

[1] Scaled for 128x128 imaging, [2] Scaled for 48x16 MD, [3] With pixel aggregation

Figure 27.2.6: Comparison table and sample images.

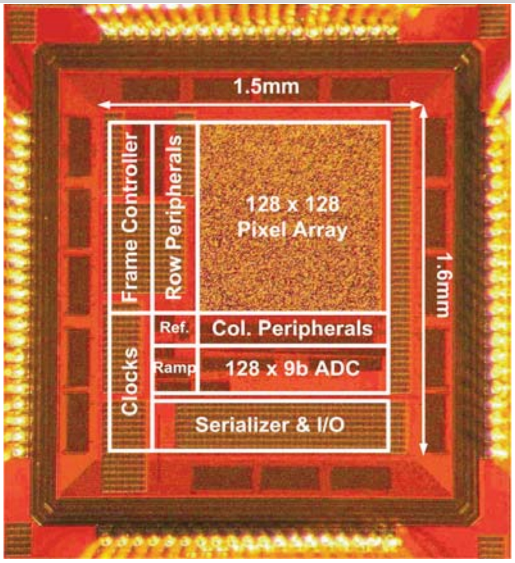

Figure 27.2.7: Die micrograph.

978-1-4673-4516-3/13/$31.00 ©2013 IEEE