27.1 A 3.4μW CMOS Image Sensor with Embedded Feature-Extraction Algorithm for Motion-Triggered Object-of-Interest Imaging

Jaehyuk Choi, Seokjun Park, Jihyun Cho, Euisik Yoon

University of Michigan, Ann Arbor, MI

Distributed sensor nodes typically operate under the constraint of limited energy source, and power consumption is an important factor to extend the lifetime of sensor systems. Several low-power imagers have been reported for application to wireless sensor networks [1,2]. However, the biggest power consumption comes from wireless signal transmission due to the large bandwidth of image signals [3]. One way to reduce the bandwidth is to generate signals only when an event happens, by monitoring temporal changes [3-5]. However, this event-based imaging has extraneous redundancy because the sensor may also respond to environmental conditions, such as change of illumination or background movement in addition to actual target objects.

In this paper, we report motion-triggered object-of-interest (OOI) imaging to suppress the redundancy in imaging as well as transmission of signals. Most of the time the sensor is in sleep mode, where it remains until it is triggered by motion. When it wakes up, it generates and transmits 8b features for object detection. The signal-processing unit, which resides either in the host or in the sensor node, performs detection of objects and feeds back a 1b request signal to initiate further imaging operation if the OOI is identified. Among many feature-extraction algorithms, we incorporate the histogram-of-oriented-gradients (HOG) because it gives a high detection rate of objects in simple operation [6]. The HOG feature output requires only 3.5% of the bandwidth of conventional 8b image capturing. In this work, we implement the HOG feature-extraction algorithm using mixed-signal circuitry to save both power and area.

Figure 27.1.1 shows the simplified block diagram of three different modes of operation in the proposed sensor chip. In motion-sensing mode, the sensor generates a 128×128 1b motion map for motion triggering. The column-parallel ADCs operate as a 1b motion comparator. Once motion is detected, the sensor wakes up and turns into feature-extraction mode. In this mode, the sensor generates and transmits 8b feature signals that are extracted from full 8b 256×256 images. We have SRAM blocks that temporarily store intermediate signals during HOG feature calculation. One part of the SRAM blocks is allocated for line buffers to store 3 rows of images, and the rest is allocated for feature signal accumulation. In imaging and storing mode, the sensor generates and transmits 8b images upon a request from the host. The SRAM operates as a frame memory that can store region-of-interest images containing the OOI.

The overall architecture of the sensor is shown in Fig. 27.1.2. The column-parallel 8b single-slope ADCs operate as a 1b motion comparator in the motion-sensing mode and generate a 1b motion map with frame-difference signals from the pixel. The pixel array is divided into m×m sub-blocks, where the size of sub-blocks is fully programmable. A set of histograms is generated in each block to map the magnitude of spatial gradients in each angle. The processing of feature signal extraction is as follows: (1) three rows of images are read out and temporarily stored in SRAM; (2) spatial gradients (G_X and G_Y) are calculated with masking [-1 0 1] and $[-1\ 0\ 1]^T$, respectively; (3) a gradient-to-angle converter (GAC) generates 9 bin numbers that map 9 different angles from 0° to 180°; and (4) a histogram generator accumulates the magnitude of gradients ($|G_X+G_Y|$) corresponding to its angle ($\theta = \tan^{-1}(G_Y/G_X)$) calculated from GAC, and stores them back to SRAM. The accumulated magnitudes are normalized by K in order to deliver full-scale 8b feature signals.

Figure 27.1.3 shows the pixel architecture. The pixel employs a reconfigurable differential topology that enables both source-follower readout and differential common-source amplification without any additional transistors [1]. Two vertical neighboring pixels share the COM line and connect signals through SIG0 and SIG1. An in-pixel capacitor $C_{E/O}$ is used as frame-buffer memory for motion sensing as well as for low-power imaging. The in-pixel capacitor is implemented using an MIM capacitor and placed on top of pixel circuits. Therefore, the capacitor does not induce any fill-factor loss. Two vertically neighboring pixels are connected with a merging switch (M) for pixel merging. In the motion-sensing mode, the operation sequence is as follows: (1) Two pixels are merged and one in-pixel capacitor samples the previous frame signal (V_1); (2) the other in-pixel capacitor samples the current frame signal (V_2); (3) a short pulse is applied to $V_{COMP(E)}$ and V_1 is compared with V_2; (4) likewise, another pulse is applied to $V_{COMP(O)}$ and two voltages are compared again. When the frame difference $|V_1-V_2|$ is higher than the amplitude of V_{COMP} swing, a 1b motion output is generated. For low-power imaging in the feature-extraction mode, the pixel circuit operates as a pre-amplifier for an 8b single-slope ADC. The integrated signal in a photodiode is transferred to an in-pixel capacitor in the even pixel, whereas the odd pixel keeps the reset voltage. A ramp signal is applied to $V_{COMP(E)}$, and the latching occurs around the reset voltage. In this scheme, the V_{GS} drop in the amplification transistors does not induce the loss of signal swing as in the conventional source-follower readout. We used 1.3V for the pixel power supply, which significantly saves power consumption.

In the feature-extraction mode, the calculation of angle ($\theta=\tan^{-1}(G_Y/G_X)$) involves complex operations such as division and trigonometric functions. In this work, we choose a simple mixed-signal approach to avoid a digital implementation that would consume huge area/power. Figure 27.1.4 shows the gradient-to-angle converter (GAC). The GAC generates 9-angle information from 0° to 180° in 4b bin numbers. The 8b gradient signals (G_X and G_Y) are converted to analog signals using two binary capacitive DACs. When the switch S1 is off, the binary DAC output is divided by two. One capacitor ($C_{0,1}$) operates as the part of DAC that generates the trigonometric function, while the other ($C_{0s,1s}$) operates as a hold capacitor that contains the gradient. The gradient is multiplied by $\tan\theta$ and $1/\tan\theta$ in each channel, respectively, according to capacitance values of C_{20}, C_{40}, C_{60}, and C_{80}. After 2-step comparisons during t3 and t4, the angle information is encoded into a 4b bin number.

A prototype chip is fabricated using a 0.18μm CMOS process. Figure 27.1.5 shows the captured images from the fabricated device including a 128×128 image motion map and a 256×256 image from the in-pixel single-slope ADC. A 256×256 angle map shows the angle calculated from the GAC. An 8b feature from 8×8 blocks is shown with the vector. In this figure, only the two angles that have the biggest magnitude are shown for simplicity. The performance of the sensor is summarized in Fig. 27.1.6. To verify the performance of the integrated feature-extraction unit, we test the object detection from the extracted features using 200 pedestrian images from DaimlerChrysler dataset [7]. For testing, we input test images serially into the 8b latch in the column-parallel ADC and generate features. The test result shows a 94.5 % detection rate. We achieve a normalized power of 13.46pW/frame·pixel in motion sensing, and 51.94pW/frame·pixel in feature extraction. A chip micrograph is shown in Fig. 27.1.7.

References:

[1] J. Choi, S. Park, J. Cho, and E. Yoon, "A 1.36μW Adaptive CMOS Image Sensor with Reconfigurable Modes of Operation from Available Energy/Illumination for Distributed Wireless Sensor Network," *ISSCC Dig Tech. Papers*, pp. 112-113, Feb. 2012.

[2] M.-T. Chung and C.-C. Hsieh, "A 0.5V 4.95μW 11.8fps PWM CMOS Imager with 82dB Dynamic Range and 0.055% Fixed-Pattern Noise," *ISSCC Dig Tech. Papers*, pp. 114-114, Feb. 2012.

[3] S. Chen, W. Tang, X. Zhang, and E. Culurciello, "A 64×64 Pixels UWB Wireless Temporal-Difference Digital Image Sensor," *IEEE Trans. VLSI*, vol. 20, no. 12, pp. 2232-2240, Dec. 2012.

[4] U. Mallik, M. Clapp, G. Cauwenberghs, and R. Etienne-Cummings, "Temporal Change Threshold Detection Imager," *ISSCC Dig Tech. Papers*, pp. 362-363, Feb. 2005.

[5] N. Massari, M. Gottardi, and S. Jawed, "A 100μW 64×128-Pixel Contrast-Based Asynchronous Binary Vision Sensor for Wireless Sensor Networks," *ISSCC Dig Tech. Papers*, pp. 588-589, 638, Feb. 2008.

[6] N. Dalal and B. Triggs, "Histogram of Oriented Gradients for Human Detection," *CVPR Dig Tech. Papers*, pp. 886-893, June 2005.

[7] Daimler Pedestrian Benchmark Data Sets, [online] http://www.gavrila.net/Research/Pedestrian_Detection/ pedestrian_detection.html, accessed Nov. 17, 2012.

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

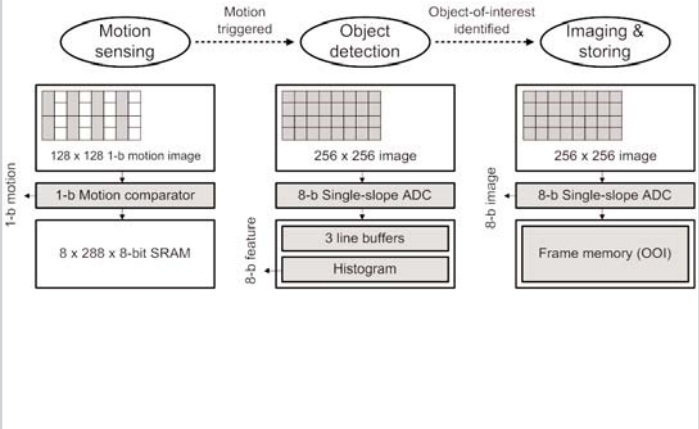


Figure 27.1.1: Objective-of-interest imaging operation.

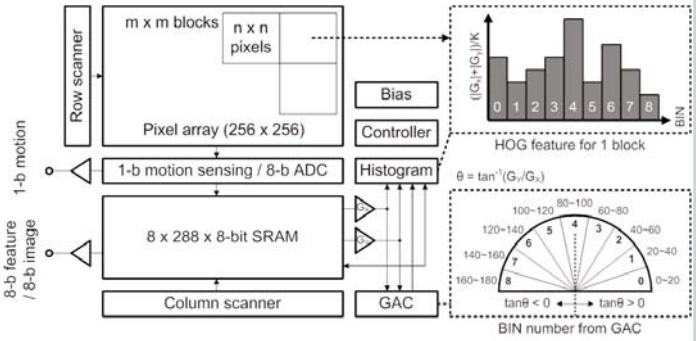


Figure 27.1.2: Sensor architecture.

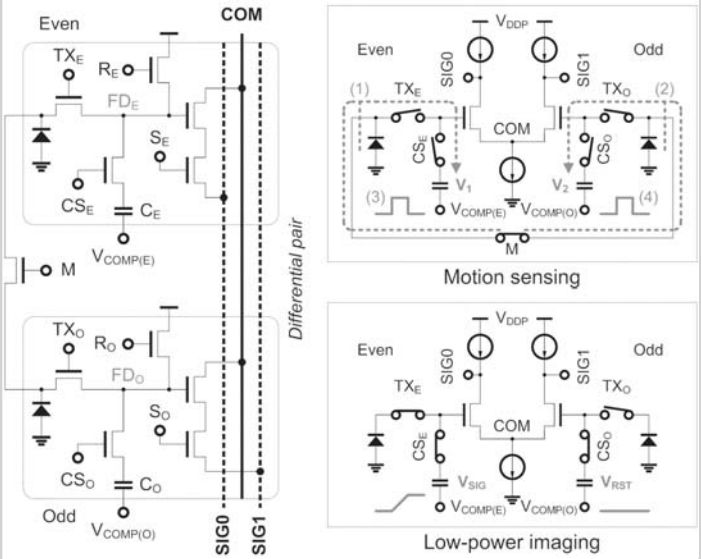
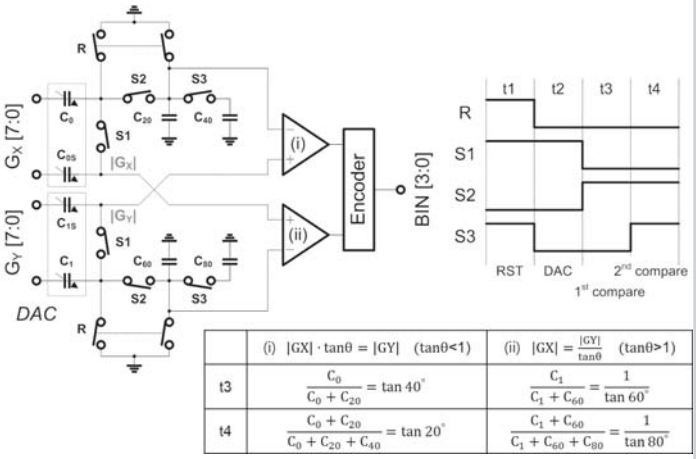


Figure 27.1.3: Pixel architecture and equivalent circuits.



	(i) $\lvert GX\rvert \cdot \tan\theta = \lvert GY\rvert$ $(\tan\theta<1)$	(ii) $\lvert GX\rvert = \frac{\lvert GY\rvert}{\tan\theta}$ $(\tan\theta>1)$
t3	$\frac{C_0}{C_0+C_{20}} = \tan 40^\circ$	$\frac{C_1}{C_1+C_{60}} = \frac{1}{\tan 60^\circ}$
t4	$\frac{C_0+C_{20}}{C_0+C_{20}+C_{40}} = \tan 20^\circ$	$\frac{C_1+C_{60}}{C_1+C_{60}+C_{80}} = \frac{1}{\tan 80^\circ}$

Figure 27.1.4: Gradient-to-angle converter.

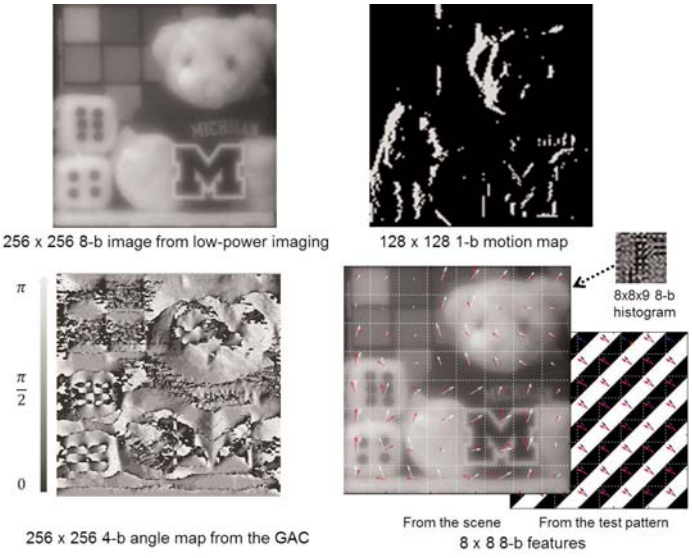


Figure 27.1.5: Sample images.

Process	0.18 µm 1P4M CMOS			
Core size	2.35 x 3.18 mm²			
Pixel size	5.9 x 5.9 µm²			
Pixel array	256 x 256			
Fill factor	30 %			
FPN	0.05 %			
Dynamic range	54.8 dB			
Power supply	Pixel		1.3 V	
	Digital		0.8 V	
Power consumption & Power FOM	Mode	Power (@ 15 fps)	Power / frame	FOM (power/pixel•frame)
	Motion sensing	3.31 µW	0.22 µW	13.46 pW
	Imaging with feature extraction	51.06 µW	3.4 µW	51.94 pW
Detection rate of object detection	94.5 %			

Figure 27.1.6: Chip characteristics.

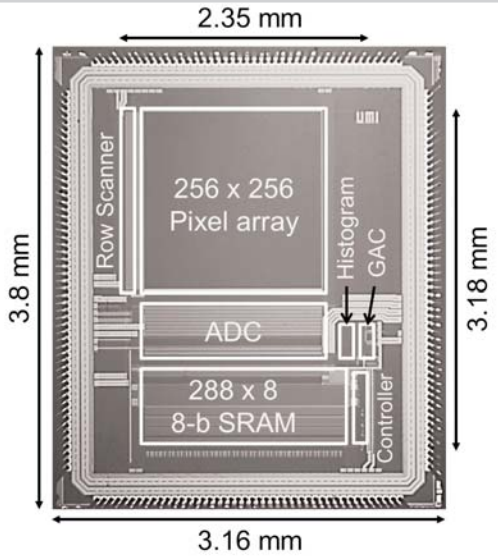


Figure 27.1.7: Chip micrograph.

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