Tri-Mode Smart Vision Sensor With 11-Transistors/Pixel for Wireless Sensor Networks

Dongsoo Kim, *Member, IEEE*, and Eugenio Culurciello, *Member, IEEE*

Abstract—We present the *T-Sensor*, a smart vision sensor capable of providing intensity, spatial contrast, and temporal difference images through the pixel level processing. The T-Sensor smart pixel is composed of only 11 transistors, allowing tight integration of different functionalities in a 16×21 μm^2 pixel area. Focal-plan processing for temporal difference and spatial contour is implemented with maximum and minimum comparing analog circuits, allowing T-Sensor to provide high-speed imaging with low-power consumption. The sensor array is 128×128 pixels, with a fill factor of 42%, and operates at 800 fps and 13M events/s with a power consumption of 1.02 mW.

Index Terms—Address-event representation (AER), CMOS image sensor (CIS), edge detection, image sensor, smart vision sensor, smart image sensor, spatial contour, temporal difference, winner-takes-all (WTA), wireless sensor networks (WSN).

I. INTRODUCTION

WIRELESS Sensor Networks (WSNs) have a significant impact on advanced sensing technologies and a wide range of applications ranging from military, to scientific, to industrial, to health-care, to home. A group of wirelessly-connected sensing devices collaborate and collect raw local data, producing globally meaningful information [1], [2]. However, WSNs have severely limiting resource bottlenecks in communication bandwidth and power, especially when used with commercial-off-the-shelf image sensors. Therefore, many custom image sensors have been published in the literature to reduce the information data size and the power consumption by performing focal-plane image processing as temporal difference [2]–[8] and spatial contrast [2], [9]–[11] and intensity provided in a digital binary format [12]–[14]. Conventionally image processing is performed off-chip but at the expense of much larger power consumption. Temporal-difference (motion) and contour detection can be performed by a GPU, digital image processing on a FPGA, or a micro-processor. However, the power of these devices is much larger

than the sensor presented here (25 W for a mobile GPU, a few watts for FPGAs, and at least 100 mW for low-power micro-controllers), and the frame rate is limited by the necessary analog-to-digital conversion (ADC) and digital signal processing time which also adds a few 100 mW to the power budget, and is thus beyond what WSNs can provide.

Several recent smart vision sensors with focal-plane processing have been proposed to reduce the sensor data and at the same time provide low-power processing of frames. Many of these sensor, a few of which are reported in Table I, implement analog image processing to provide output binary images in the form of temporal-differences or spatial contours. Smart image sensor with focal-plane processing have the drawback of requiring analog circuits that are more complex than standard image sensor pixels in order to perform signal processing. These circuits limit the compactness of the smart pixel and require larger silicon area per pixel, ultimately reducing the pixel array densities. Lichtsteiner, *et al.* [6] presented an asynchronous temporal difference sensor with wide dynamic range feature. The smart pixel is composed of 26 transistors and 3 capacitors and the fill factor, the ratio of the photodiode area to the pixel area, is 8.1%. Costas, *et al.* [9] presented a smart vision sensor for the spatial contour detection and 30 transistors were used to implement a smart pixel which has a fill factor of 3%. Recently Gottardi, *et al.* [3] showed a smart vision sensor, which can provide the temporal difference and spatial contour together with low power consumption. The smart pixel is composed of 45 transistors and its fill factor is 20%.

In this paper, we present a custom image sensor capable of reporting intensity, spatial contour, and temporal difference images using only 11 transistors per pixel, which also has the feature of low power consumption. The Section II describes the architecture of the proposed sensor. A detailed description of the circuit implementation is presented in Section III. Measurement results from the fabricated prototype sensor are reported in Section IV.

II. ARCHITECTURE OF THE PROPOSED SENSOR

The most challenging design requirements of a smart vision sensor that can provide at the same time intensity, spatial contour (edge), and temporal difference images is detecting spatial contours with as little transistors as possible, in order to keep the pixel fill factor as high as possible with small pixel area. The intensity image can be easily obtained using a source follower from the photo-integrated signal of the photo-diode [15]. Temporal-difference information can be obtained

|	Gottardi [3]	Lichtsteiner [6]	Costas [9]	Ruedi [10]	This work
Smart function	Temporal difference, Spatial contour	Temporal difference, wide dynamic range	Spatial contour	Spatial contour, wide dynamic range	Temporal difference, Spatial contour
Min gate length in pixel	$0.35 \ \mu m$	$0.35 \ \mu m$	$0.35 \ \mu m$	$0.5 \ \mu m$	$0.35 \ \mu m$
Array size	128×64	128×128	32×32	128×128	128×128
Number of TRs/pixel	45T	26T+3Cap	30T	more than 80T	11T
Pixel size	$26 \times 26.5 \ \mu m^2$	$40 \times 40 \ \mu m^2$	$58 \times 56 \ \mu m^2$	$69 \times 69 \ \mu m^2$	$16 \times 21 \mu m^2$
Fill factor	20 %	8.1 %	3 %	9 %	42 %
Chip size	$4.5 \times 2.5 \ mm^2$	$6 \times 6.3 \ mm^2$	$3.1 \times 3.1 \ mm^2$	$99 \ mm^2$	$2.8 \times 2.8 \ mm^2$
Power consumption	100 μW @ 3.3V	30 mW @ 3V	9.9 mW @ 3.3V	300 mW @ 3.3V	1.02 mW @ 3V
Event rate	32M events/s	2M events/s	1.6M events/s	8M events/s	13M events/s
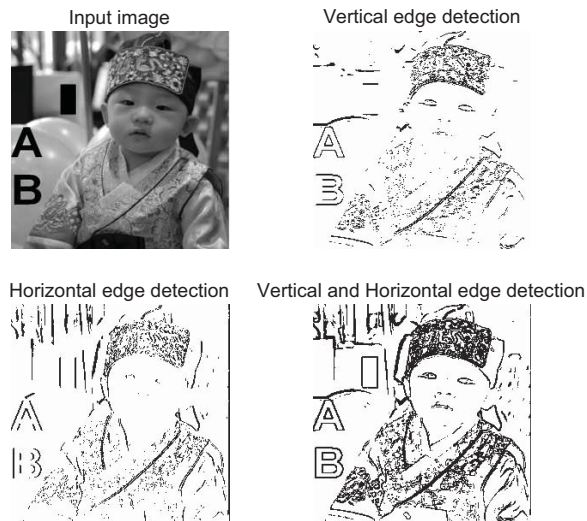
Fig. 1. Edge detection algorithm using WTA and LTA operations, vertical edge detection is the result of WTA/LTA operations for the vertically located pixels, and horizontal edge detection is the result of WTA/LTA operations for the horizontally located pixels.

with an additional in-pixel storage capacitor to store previous frames [16]. However, obtaining spatial contour from an image requires a comparison of the signals in the neighboring pixels in order to compute edge information. This comparison increases the number of required transistors and interconnections, causing an increase of the pixel size which limits the integration of bigger imaging array in the limited silicon area. The main contribution of the work presented in this paper is the use of winner-takes-all (WTA) and loser-takes-all (LTA) circuit blocks to obtain the contour information. WTA (LTA) circuit is a non-linear functional block that identifies the largest (least) input out of multiple inputs and copies the winner (looser) value to the output. These WTA and LTA can be used to detect the temporal and spatial difference. These circuit blocks use few interconnections between neighboring pixels and can be implemented with a small number of the transistors. This in turn reduces the smart pixel size and allows for high densities smart imaging arrays.

Fig. 1 shows a system level simulation using WTA and LTA operations to detect image edges/contours. The maximum pixel signal (darker signal) and the minimum pixel signal (brighter signal) between the reference pixel and its neigh-

boring pixels are calculated by the WTA and LTA operations. If the difference between the maximum and minimum values is higher than a threshold, the reference pixel can be marked as an edge. Vertical edges or horizontal edges can be detected by WTA and LTA operations between the pixels located on the vertical and horizontal sides of the evaluated pixel. When the WTA and LTA operations are performed for the evaluated pixel and its three neighbors (right, bottom, and right-bottom pixels), the vertical and horizontal edges can be computed at the same time. Therefore, if there is any gradient (difference) in the light intensity horizontally or vertically or even diagonally (any one of three directions), a contour is detected. When the kernel size of the image processing is increased, it can detect spatial difference which changes smoothly over the bigger image area. However there is the limitation of the physical connection between the pixels. The kernel size of WTA/LTA operations can be adjusted to detect the edges in a larger number of neighboring pixels. However the bigger kernel size increases the complexity of the circuit implementation, and for this reason we consider a 2×2 pixel neighborhood in our sensor.

The proposed tri-mode vision sensor (*T-Sensor*) is composed of a 128×128 smart pixel array, a row control circuit for a rolling shutter operation, column readout circuits, and a digital event generator as shown in Fig. 2. The smart pixel contains WTA and LTA circuits to find the maximum and the minimum input values. There are two vertical column line (B and $\bar{B}$) in each column array. The maximum and minimum values that are calculated by in-pixel WTA/LTA circuits are transferred to the column circuits through these column lines. In the intensity mode (*I-mode*), each smart pixel transfers the reset signal and the photo-integrated signal of the photodiode (PD) to the column readout circuit through the two column lines. Since the reset signal of the pixel is higher than the photo-integrated signal, the reset signal is transferred to the column line B. Line B is connected with the WTA circuits and the photo-integrated signal is on column line $\bar{B}$ connected with the LTA circuits. The column readout circuit computes the difference between the signals on the two column lines and generates the intensity image through delta-difference sampling (DDS) [17]. The DDS technique provides a high quality light intensity image by removing low frequency noise and offsets of the readout circuits. In the temporal difference mode (*T-mode*), the previous frame signal,
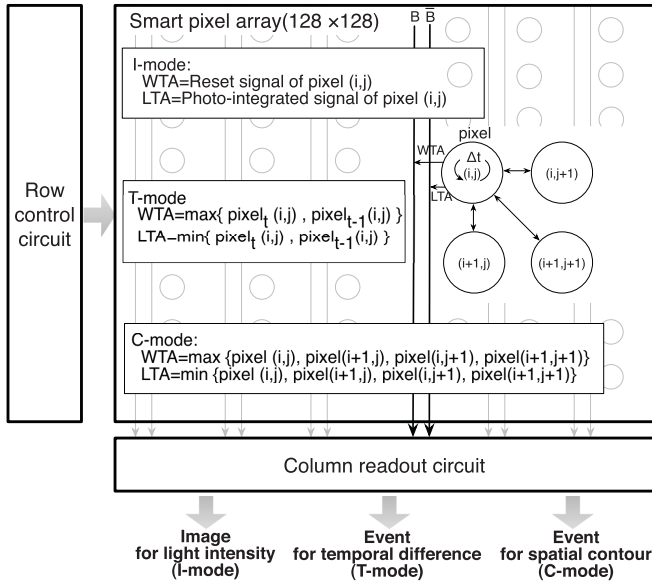
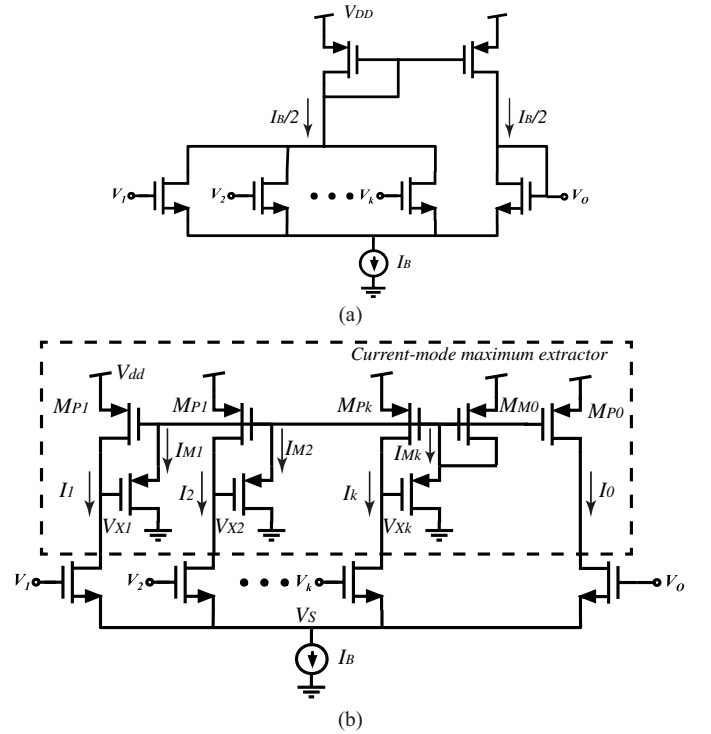Fig. 2. Block diagram of the proposed *T-Sensor* system.

Fig. 3. Schematic diagram of winner-take-all (WTA). (a) Simple schematic diagram of WTA circuit based on the common-source configuration. (b) Common-source WTA circuit combined with current-mode maximum extractor.

$pixel_{t-1}(i, j)$ which was stored in the storage capacitor at the time of the previously readout and the current frame signal, $pixel_t(i, j)$ are compared by the WTA/LTA circuits. The smart pixel outputs the previous frame signal and the current frame signal to the column readout circuit after WTA/LTA operations (WTA $= \max(pixel_{t-1}(i, j), pixel_t(i, j))$, LTA $= \min(pixel_{t-1}(i, j), pixel_t(i, j))$). The column readout circuit generates a digital output event when the difference between the previous frame signal and the current frame signal is greater than a certain threshold. In the spatial contour mode (*C-mode*), the reference pixel, $pixel_{(i,j)}$ finds the maximum and the minimum photo-integrated signals in the 4 neighboring pixels, $pixel_{\{(i,j),(i+1,j),(i,j+1),(i+1,j+1)\}}$ and transfers the maximum and minimum signals to the column readout circuit. If the spatial contrast in the four pixels is high, it means that a contour (edge) was found. The column readout circuit calculates the difference between the maximum and the minimum signal and generates an edge digital output event by comparing the difference with a threshold.

III. Circuit Implementations of *T-Sensor*

The Winner-Take-All (WTA) circuit implementation derives from a common-source configuration as shown in Fig. 3(a) [18]–[20] . If the PMOS current mirror is ideal then two branches should have the same current, $I_B/2$. Assuming that the transconductance of the input transistors are large enough so that the input transistors operate as ideal switches, only the transistor with the highest gate voltage (winner) is turned on because all the input and output transistors share the common source node, V_S and the same amount of current should flow through both of input and output branches. Because of this, the gate-to-source voltage of the output transistor is the same as that of the winning transistor.

In reality, the bias current, $I_B/2$ is divided into several input branches due to the finite transconductance of the transistors if the maximum input voltage is not big enough to turn off

all other transistors. If all the input voltages are equal, as the worst case, then $I_B/2$ splits equally into k-input branches and the output voltage is deviated from the correct value. This error is called as *corner error* and can be obtained as follows [18]:

$$\Delta V_O = \sqrt{\frac{I_B}{\beta} \cdot (1 - \frac{1}{\sqrt{k}})} \tag{1}$$

where $\beta = \frac{K \cdot W}{2 \cdot L}$, K is the transconductance parameter, and W/L is the aspects ratio of the input transistor. The corner error can be improved by combining the common source WTA circuit with a current-mode maximum extractor as shown in Fig. 3(b) [18]. The corner error in this configuration is given by

$$\Delta V_O = \frac{(\lambda_n + \lambda_p) \cdot I_B}{2\sqrt{\beta_n \cdot \beta_p}} \tag{2}$$

where λ_n and λ_p are the channel length modulation coefficient of NMOS and PMOS transistors, respectively. Since all input branches are off except the winning input branch, V_{Xi} are kept at the positive supply voltage Vdd except the winning input branch. Therefore, V_{Xi} can be considered as a logic output that indicates the winning input. In addition to the corner error, the more critical error in a practical WTA implementation is the offset error caused by the threshold voltage and the geometric device mismatch. Loser-Takes-All (LTA) circuit can be implemented by the same WTA circuits, just by changing NMOS transistors with PMOS transistors.

The schematic diagram of the proposed smart pixel is presented in Fig. 4 (a). Only 11-transistors/pixel provide intensity,
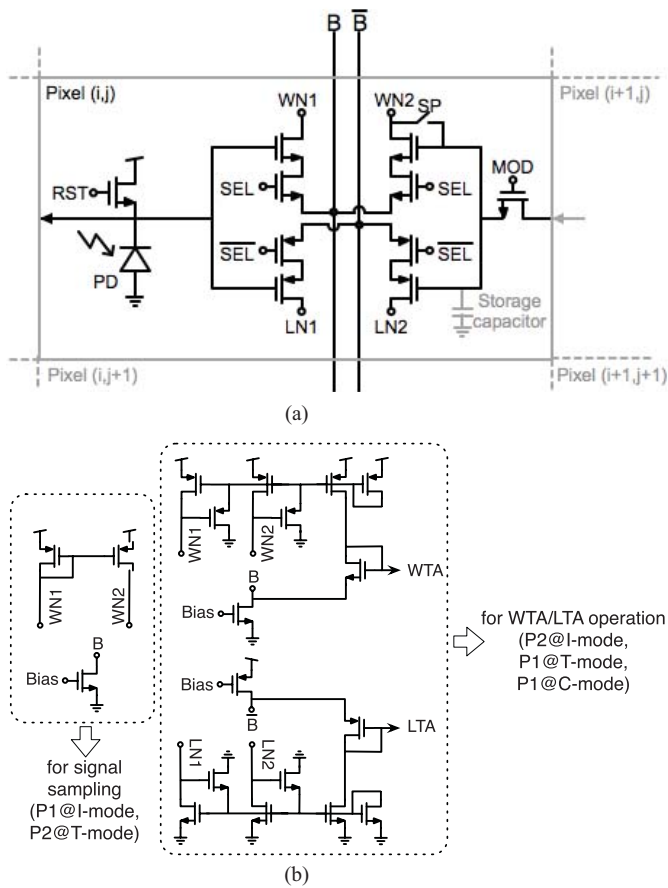
Fig. 4. Schematic diagram of (a) the proposed smart pixel and (b) the WTA and LTA operations of both the proposed smar pixel and the column readout circuit for the sampling.
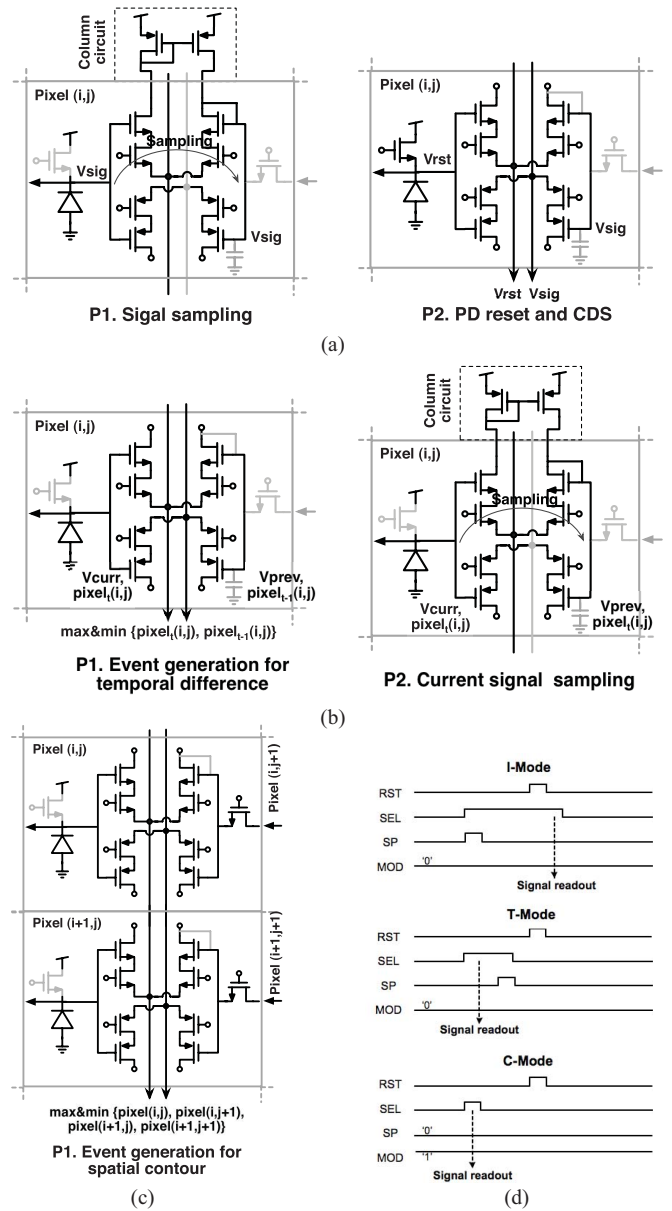

Fig. 5. Operation principals of *T-Sensor*. (a) Operation phases of I-mode. (b) Operation phases of T-mode. (c) Operation phases of C-mode. (d) Timing diagram.

temporal difference, and spatial contrast images together and enables greater pixel array integration in the same die size. 11 transistors is also much less than 45 transistors/pixel for temporal difference and spatial contrast images in [2], and 51 transistors/pixel for spatial contrast and wide dynamic intensity images in [10]. In addition to a reset transistor, RST and PD, four NMOS transistors compose the WTA input circuit and four PMOS transistors form the LTA input circuit from WTA/LTA circuits as shown in Fig. 3(b). The current mirror of WTA circuits in Fig. 3(b) will be included in the column circuits. SP is turned on to transfer the signal from the PD to the parasitic storage capacitor. MOD is turned on in C-mode to connect the signal of the PD in the right side as an input of WTA/LTA circuit. Fig. 4(b) shows the column readout circuit that configures the pixels to perform a photodiode sampling operation, or the WTA/LTA operations, by means of the SP control signal.

Fig. 5 describes the operation principals for I-mode, T-mode, and C-mode. During P1 signal sampling phase of I-mode, the column circuit for signal sampling in the left of Fig. 4 (b) is connected with the pixel and the pixel circuit operates as a signal buffer. Therefore, the photo-integrated signal during an exposure time can be sampled into the storage capacitor. During P2 phase after sampling the photo-integrated

signal, the column circuit for WTA/LTA in the right schematic of Fig. 4(b) is connected with the pixel. PD is reset and the reset signal, $Vrst$ at the PD and the photo-integrated signal, $Vsig$ at the storage node are readout through WTA and LTA circuits, respectively. Since the reset signal is greater than the photo-integrated signal, the reset signal is transferred through column line B and the photo-integrated signal is transferred through column line $\bar{B}$. The column circuit generates the light intensity image by subtracting $Vsig$ from $Vrst$ for DDS function.

During P1 phase of T-mode, the column circuit for WTA/LTA operations in the right schematic of Fig. 4(b) is connected with the pixel as same with P2 phase of I-mode. The previous frame signal, $pixel_{t-1}(i, j)$ in the storage capacitor (which was stored at the previous frame) and the current
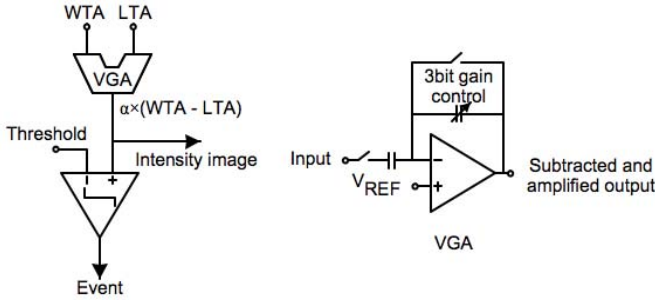
Fig. 6.   Block diagram of the event generator and the variable gain amplifier.

frame signal, $pixel_t(i, j)$ in the PD are readout simultaneously through the WTA/LTA circuits. The column circuit generates digital events based on the difference between the previous frame signal and the current frame signal. During phase P2, the column circuit for sampling the current frame signal is sampled and stored in the parasitic storage capacitor as the previous signal for the next frame which is the same operation of P1 phase of I-mode, and the PD is reset for a new integration. In C-mode, the row selection signals for two rows are turned on together and MOD transistor is turned on as shown in Fig. 5(c). By turning on the SEL row controls for two rows, a 2×2 pixel neighborhood is connected together through the two column lines B and $\bar{B}$.

In C-mode, the storage node is not used and is thus connected to the neighboring pixels by turning on MOD transistors. These allow the pixel circuits to be reconfigured into a 4-inputs ($pixel_{(i,j),(i+1,j),(i,j+1),(i+1,j+1)}$) WTA/LTA circuit used for detecting spatial contrast in the 4 neighboring pixels without complicated signal connections between the neighboring pixels. Since in C-mode the photo-integrated signals in two row are compared, the exposure time for the pixels in different rows should be same. However, in C-mode, the integration time is different if one uses the same rolling shutter employed in the I-mode and T-mode. This can place a limitation on the minimum detectable contrast difference. In C-mode, we therefore employ a different row control method that provides the reset signals of the PD for each pair of rows at the same time. The only issue this this row control method, is that it will only provide edge information every other row. The problem of missing horizontal edges by skipping rows can be solved by changing the skip row location in a frame-by-frame fashion. For example, we skip the even row for one first frame and then skip the odd row during the next frame, and so forth.

The event generator including a variable gain amplifier (VGA) and a comparator is presented in Fig. 6. The VGA is implemented using a switch capacitor circuit and the gain α is adjustable from 1 to 8 using a digitally-controlled capacitor bank. As shown in Fig. 6, the same column circuit can be used for three different modes. The VGA provides the subtraction between two column lines which have outputs from WTA/LTA. The subtracted difference, (WTA-LTA) is the light intensity signals after DDS for I-mode, temporal difference between the previous and current frames for T-mode, and the spatial contrast for the 2×2 neighboring pixels for

C-mode, respectively. The difference amplified by the gain, α can be digitized by A/D conversion for the light intensity image, or can be input for the comparator to generate events with a threshold value. The generation of Address Event Representation (AER) is also possible using two 7-bit counters which have synchronized clocks with the row control circuits and the column readout circuits.

The threshold is the important reference to decide the event generation and can be ideally decided as follows

$$V_{TH} = L_{DET} \times S \times T_{int} \times \alpha \qquad (3)$$

where L_{DEC} is the minimum light intensity difference to generate the events in $[(\mu W/cm^2)]$, S is the sensitivity of the sensor in $[V/(s \cdot (\mu W/cm^2)]$, T_{int} is the integration time of the sensor which is the reciprocal of the frame rate, and α is the gain of the readout circuit before the event generation. However the minimum V_{TH} is limited by the readout noise from the pixel to the comparators as follows to avoid noisy event generations.

$$V_n = 2 \cdot (V_{n,SHOT} + V_{n,WTA} + \frac{V_{n,VGA}}{\alpha})$$
$$+ \frac{V_{n,COMP}}{\alpha} \ [V^2/Hz] \qquad (4)$$

where $V_{n,SHOT}$ is the shot noise of the photodiode and $V_{n,WTA}$ and $V_{n,VGA}$ are the thermal noise of input transistors in WTA/LTA and VGA since the flicker noise can be reduced effectively by the auto zeroing of VGA. The factor of 2 is caused by the thermal noise summing due to the DDS function. The noise $V_{n,WTA}$, $V_{n,VGA}$ and $V_{n,COMP}$ can be represented by

$$V_{n,WTA} = V_{n,VGA} = \frac{16KT}{3g_m} \ [V^2/Hz] \qquad (5)$$

$$V_{n,comp} = \frac{16KT}{3g_m} + \frac{K_F}{C_{OX}^2 WL}\frac{1}{f} \ [V^2/Hz] \qquad (6)$$

where K is Boltzmann's constant, T is the absolute temperature in kelvins, g_m is the transconductance of the transistors, K_F is coefficients of the flicker noise, C_{OX} is the gate oxide capacitance, and W and L are the width and length of the transistors, respectively [21].

IV. MEASUREMENT RESULTS

The proposed *T-Sensor* was fabricated with a 0.18-μm SiGe BiCMOS 7-metals process and the microphotograph is presented in Fig. 7(a). We used this process because it was available to us at no cost from a grant from MOSIS. Since the process supports minimum gate length of 0.18-μm with a thin oxide, the dynamic power consumption can be reduced by smaller power supply voltage using the thin transistor in the digital blocks (row and column decoders). However the dynamic power consumption in the digital circuit is small and the dual power supplying in the wireless sensor nodes makes the sensor node system complicated. Therefore, only thick oxide transistors which have the minimum gate length of 0.35-μm are used for the design of the proposed sensor. A CMOS image sensor process will deliver better sensitivity to light and improved pixel noise performance. But the processed
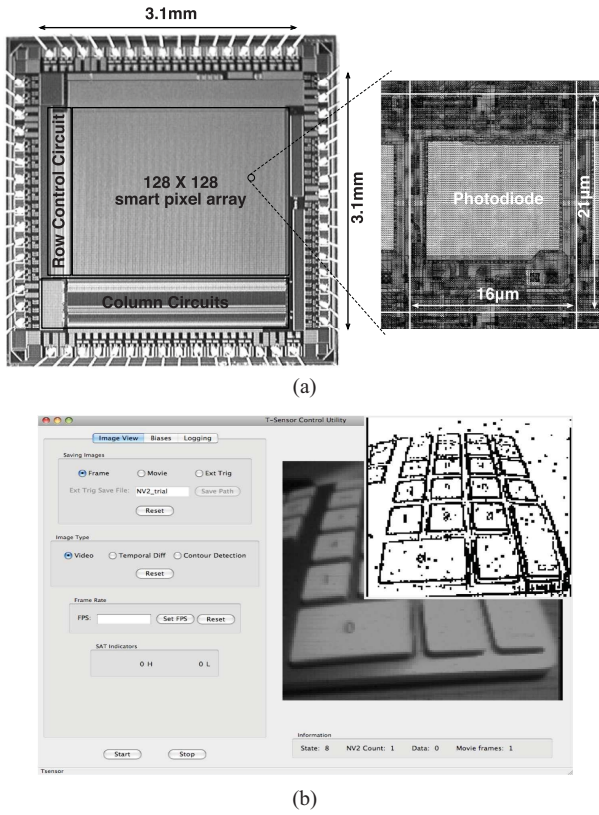
Fig. 7. Microphotograph of the fabricated *T-Sensor* die and graphic user interface(GUI).

used is sufficient to present the innovative design ideas of the T-Sensor prototype reported in this paper.

The fabricated *T-Sensor* has 128×128 smart pixel array and the core area is 3.1×3.1 mm^2. Each smart pixel has a pitch of 16 μm × 21 μm with a fill factor of 42%. The photodiode is implemented with n$^+$/p-sub diode. The measured pixel sensitivity is 2.14 V/s·(μW/cm^2) at 550-nm light wavelength and 0.31 V/s·(μW/cm^2) at 850-nm light wavelength. The measured dynamic range of the sensor is 58 dB in 1.7×10^6 e^- full well capacity with the peak SNR of 43 dB. The measured pixel FPN and column FPN were 0.18% and 0.10%, respectively. Since the T-Sensor was fabricated with a silicon germanium process that has better sensitivity in the longer wavelengths, the sensor can be used in dark lighting conditions with infra-red light sources. Fig. 7(b) shows a graphic user interface (GUI) to verify the performance and the sample light intensity and spatial contour images for a keyboard.

Fig. 8(a) shows the test board that includes the fabricated *T-Sensor*, CS-mount lens (f=12 mm, 1:1.4) and USB interface components. The *T-Sensor* can operate with power supplies of 2V to 3.3V and the power consumption is 1.23 mW (I-mode) and 1.02 mW (T- and C-modes) at 3 V, 500 μW (I-mode) and 440 μW (T- and C-modes) at 2V, respectively. Fig. 8(b), (c), and (d) show the sample images taken from a rotating pattern on the resolution chart. The sample images show that the proposed *T-Sensor* can provide for the light intensity images, the temporal difference image, and the spatial contour images effectively. Fig. 8(e) shows the number of
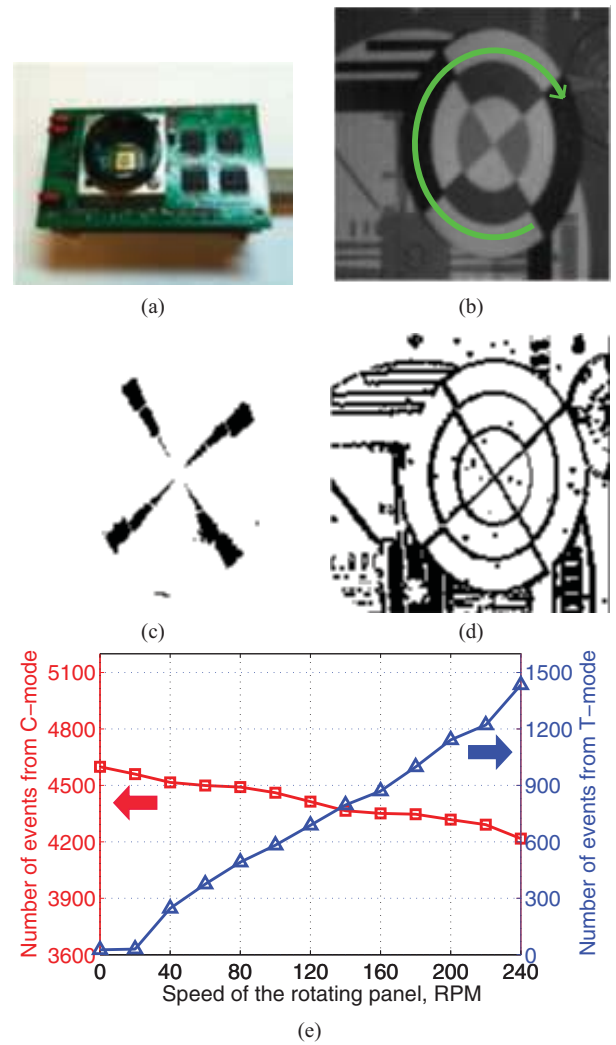

Fig. 8. Test PCB board and sample images taken from a rotating circle panel on the resolution chart. (a) Test board. (b) Intensity image. (c) Temporal difference image from the rotating circle panel. (d) Spatial contour image. Measurement results with CS-mount lens (f=12 mm, 1:1.4).

TABLE II

PERFORMANCE SUMMARY

Process	0.18-μm SiGe BiCMOS 7-metal
Power supply & power consumption	2V to 3.3V
I-mode	500 μW at 2 V, 1.23 mW at 3 V
C- and T-mode	440 μW at 2 V, 1.02 mW at 3 V
Chip size	3.1 × 3.1 mm^2
Array size	128 ×128
Pixel size	16×21 μm^2
Fill factor	42 %
Conversion gain	1.17 μV/e^-
Dynamic Range	58 dB
Full well capacity	1.7×10^6 e^-
Sensitivity	2.14 V/s·(μW/cm^2) @ 550 nm 0.31 V/s·(μW/cm^2) @ 850 nm
Frame(event) rate	200 fps for I-mode 800 fps (13M events/s) for C- and T-mode

events change for T-mode and C-mode when the speed of the rotating panel is increased linearly with revolutions-per-minute (RPM). The maximum frame rate is 200 frames/s (I-mode),

800 frames/s (T- and C-modes) and the maximum event rate is 13M event/s with 3-V power supply. Additional test videos can be found here: https://engineering.purdue.edu/elab/blog/research/synthetic-vision/bio-inspired and on YouTube: http://www.youtube.com/watch?v=VYpfYAE1IPE, on the Culurciello's channel.

V. CONCLUSION

The proposed smart vision sensor applied for wireless sensor networks can provide intensity, spatial contrast, and temporal difference images with low power and high speed of 800 fps and 13M events/s. The compact size of the pixel composed of only 11 transistors enables the integration of a large pixel array in a limited sensor size and high light sensitivity given its large fill factor. The *T-Sensor* can operate more than 120 days with two alkaline AA batteries and these features of *T-Sensor* promises the commercialization of wireless sensor networks application with the integration of the low power wireless communication system like ultra wide band (UWB) system.

VI. ACKNOWLEDGMENT

REFERENCES

[1] C. Y. Chong and S. P. Kumar, "Sensor networks: Evolution, opportunities, and challenges," *Proc. IEEE*, vol. 91, no. 8, pp. 1247–1256, Aug. 2003.

[2] N. Massari, M. Gottardi, and S. Jawed, "A 100μW 64 × 128 pixel contrast-based asynchronous binary vision sensor for wireless sensor networks," in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, Feb. 2008, pp. 588–638.

[3] N. Gottardi, M. Massari, and S. A. Jawed, "A 100μW 128 × 64 Pixels contrast-based asynchronous binary vision sensor for for sensor networks applications," *IEEE J. Solid-State Circuits*, vol. 44, no. 5, pp. 1582–1592, May 2009.

[4] V. Gruev and R. Etienne-Cummings, "A pipelined temporal difference imager," *IEEE J. Solid-State Circuits*, vol. 39, no. 3, pp. 538–543, Mar. 2004.

[5] S. Mehta and R. Etienne-Cummings, "A simplified normal optical flow measurement cmos camera," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 53, no. 6, pp. 1223–1234, Jun. 2006.

[6] P. Lichtsteiner, C. Posch, and T. Delbruck, "A 128 × 128 120 dB 30 mW asynchronous vision sensor that responds to relative intensity change," in *Int. Solid State Circuits Conf. Dig. Tech. Papers*, Feb. 2006, pp. 2060–2069.

[7] P. Lichtsteiner, C. Posch, and T. Delbruck, "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.

[8] B. Zhao, X. Zhang, S. Chen, K.-S. Low, and H. Zhuang, "A 64 × 64 CMOS image sensor with on-chip moving object detection and localization," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 4, pp. 581–588, Apr. 2012.

[9] J. Costas-Santos, T. Serrano-Gotarredona, R. Serrano-Gotarredona, and B. Linares-Barranco, "A spatial contrast retina with on-chip calibration for neuromorphic spike-based aer vision systems," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 54, no. 7, pp. 1444–1458, Jul. 2007.

[10] P. Ruedi, P. Heim, F. Kaess, E. Grenet, F. Heitger, P. Burgi, S. Gyger, and P. Nussbaum, "A 128 × 128 pixel 120-dB dynamic-range vision-sensor chip for image contrast and orientation extraction," *IEEE J. Solid-State Circuits*, vol. 38, no. 12, pp. 2325–2333, Dec. 2003.

[11] M. Barbaro, P.-Y. Burgi, A. Mortara, P. Nussbaum, and F. Heitger, "A 100 × 100 pixel silicon retina for gradient extraction with steering filter capabilities and temporal output coding," *IEEE J. Solid-State Circuits*, vol. 37, no. 2, pp. 160–172, Feb. 2002.

[12] E. Culurciello, R. Etienne-Cummings, and K. A. Boahen, "A biomorphic digital image sensor," *IEEE J. Solid-State Circuits*, vol. 38, no. 2, pp. 281–294, Feb. 2003.

[13] N. Takahashi, K. Fujita, and T. Shibata, "A pixel-parallel self-similitude processing for multiple-resolution edge-filtering analog image sensors," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 56, no. 11, pp. 2384–2392, Nov. 2009.

[14] J. M. Margarit, L. Teres, and F. Serra-Graells, "A Sub-μW fully tunable CMOS DPS for uncooled infrared fast imaging," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 56, no. 5, pp. 987–996, May 2009.

[15] E. R. Fossum, "CMOS image sensors: Electronic camera-on-a-chip," *IEEE Trans. Electron Dev.*, vol. 44, no. 10, pp. 1689–1698, Oct. 1997.

[16] D. Kim, Z. Fu, J. H. Park, and E. Culurciello, "A 1-mW CMOS temporal-difference AER sensor for wireless sensor networks," *IEEE Trans. Electron Dev.*, vol. 56, no. 11, pp. 2586–2593, Nov. 2009.

[17] S. K. Mendis, S. E. Kemeny, R. C. Gee, B. Pain, C. O. Staller, K. Quiesup, and E. R. Fossum, "CMOS active pixel image sensors for highly integrated imaging systems," *IEEE J. Solid-State Circuits*, vol. 32, no. 2, pp. 187–197, Feb. 1997.

[18] I. E. Opris, "Analog rank extractors," *IEEE Trans. Circuits Syst I, Fundam. Theory Appl. Reg. Papers*, vol. 44, no. 12, pp. 1114–1121, Dec. 1997.

[19] B. Choi and J. Sheu, "A high-precision VLSI winner-take-all circuit for self-organizing neural networks," *IEEE J. Solid-State Circuits*, vol. 28, no. 5, pp. 576–583, May 1993.

[20] D. Kim, J. Cheon, and G. Han, "An offset cancelled winner-take-all circuit," *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.*, vol. 92, no. 2, pp. 430–435, 2009.

[21] *Spectre Circuit Simulator Reference*, Cadence Design Systems, Inc., San Jose, CA, USA, Jun. 2003.

Dongsoo Kim (M'02) received the M.S. and Ph.D. degrees in electrical and electronics engineering from Yonsei University, Seoul, Korea, in 2004 and 2008, respectively.

He is currently a Staff Analog Design Engineer with Aptina Imaging, San Jose, CA, USA. From 2008 to 2010, he was a Post-Doctoral Associate with the Department of Electrical Engineering, Yale University, New Haven, CT, USA. His current research interests include CMOS image sensors, smart sensors, low-noise circuit design, and biomedical instrumentation.

Eugenio Culurciello (S'97–M'99) received the Ph.D. degree in electrical and computer engineering from Johns Hopkins University, Baltimore, MD, USA, in 2004.

He is an Associate Professor with the Weldon School of Biomedical Engineering, Purdue University, West Lafayette, IN, USA, where he directs the 'e-Lab' Laboratory. He has authored the book *Silicon-on-Sapphire Circuits and Systems, Sensor and Biosensor interfaces* (McGraw Hill, 2009). His research interests include analog and mixed-mode integrated circuits for biomedical instrumentation, synthetic vision, bio-inspired sensory systems and networks, biological sensors, and silicon-on-insulator design.

Dr. Culurciello was the recipient of The Presidential Early Career Award for Scientists and Engineers and the Young Investigator Program from ONR, and he is a Distinguished Lecturer of the IEEE (CASS).