

# A 33 $\mu W$ 42 GOPS/W 64x64 Pixels Vision Sensor with Dynamic Background Subtraction for Scene Interpretation

Nicola Cottini  
Fondazione Bruno Kessler  
38100 Trento, I  
cottini@fbk.eu

Massimo Gottardi  
Fondazione Bruno Kessler  
38100 Trento, I  
gottardi@fbk.eu

Nicola Massari  
Fondazione Bruno Kessler  
38100 Trento, I  
massari@fbk.eu

Roberto Passerone  
Università di Trento  
38100 Trento, I  
roberto.passerone@unitn.it

Zeev Smilansky  
Emza Visual Sense Ltd.  
44422 Kfar Sava, IL  
zeev@emza-vs.com

## ABSTRACT

A  $64 \times 64$  pixel vision sensor performs adaptive background subtraction and event detection at very low power consumption. The chip is based on a VLSI-oriented vision algorithm, implemented at pixel-level, mimicking the basic process of pre-attentive visual perception. Anomalous pixel behaviors are detected and coded into a 2-bit/pixel. Each pixel integrates two programmable Switched-Capacitors Low-Pass Filters and two clocked comparators, which are fundamental blocks for the execution of the vision algorithm. The 45T square pixel has a pitch of  $26\mu m$  and a fill factor of 12%. The vision sensor consumes  $33\mu W$  at  $13fps$  and  $3.3V$ . This turns into a computing performance of  $42\text{ GOPS/W}$  and  $4\text{ GOPS/mm}^2$ , which are values aligned with the most advanced computational vision sensors.

## Categories and Subject Descriptors

B.7.1 [INTEGRATED CIRCUITS]: Types and Design Styles—VLSI

## Keywords

Vision sensor, low-power sensor, early image processing.

## 1. INTRODUCTION

Sensors are becoming increasingly pervasive in our everyday life. With minimum dimensions and less infrastructures, sensory networks need to embed computing resources and wireless data communication at a minimum energy budget, maximizing its operating lifetime and minimizing environmental footprint. In particular, vision is the sensing technology with the largest information density, which is required to be processed in real-time through high-performance computing platforms. Unlike consumer market (cameras, mobile

phones), asking for high-quality and high-resolution images, there are many other application areas where visual information is used for control purposes, ranging from surveillance and people and traffic monitoring, energy-saving in buildings, elderly care and many other. Vision computation did not make over the years significant progress in energy-autonomous applications, despite big advances in microelectronic energy-efficiency. In fact, standard imagers continuously deliver sequences of images with large redundancy, although only a small amount of information is used to perform any kind of visual task. This overloads the processor, which is required to undertake the high level part of the visual task.

Custom digital processors have been proposed for early visual processing [1],[2],[3], thanks to their high parallelism. They are mainly based on Single-Instruction Multiple-Data (SIMD) architectures, offering massive parallel processing capabilities. However, their performance are limited because they need to be fed by the video signal, which is slow.

Embedding some intelligence at sensor level, making it able to recognize and deliver only those data related to image features in the scene which are of interest, would drastically increase the energy efficiency of the entire system, without losing performance.

Despite research activity on custom vision sensors is fairly mature, a lot of work has still to be done on the ultra-low power side.

In literature, several recent examples of custom vision chip implementations are reported, targeted to low-power applications [4],[5],[6],[7]. Although most recent works report very aggressive figures of merit in pixel power consumption, ranging from  $460pW/frame.pixel$  [4] up to  $84pW/frame.pixel$  [7], it has to be pointed out that these works only refer to imagers where the vision processing is accomplished outside the chip and typically takes a large power consumption. On the contrary, a vision sensor embeds a custom image processing and can tolerate a larger consumption at chip-level, in exchange for reduced external computational resources. This main characteristic turns into a significant power saving of the entire vision system [8],[9],[10]. For this reason, dealing with vision sensors, it is more appropriate to talk about computing performance over power consumption,  $P_E$  (GOPS/W), and/or computing density,  $P_A$  (GOPS/mm<sup>2</sup>), rather than absolute power consumption.

In this paper, the architecture of a low-power computa-

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

ISLPED'12, July 30–August 1, 2012, Redondo Beach, CA, USA.  
Copyright 2012 ACM 978-1-4503-1249-3/12/07 ...\$10.00.

tional vision sensor is reported, performing pixel-level adaptive background subtraction as a basic image processing for pre-attentive visual perception [11]. Under normal operating conditions, the sensor delivers two binary images, detecting temporal gradient with sign. During an event detection in the scene, the obtained output data compression ratio is about 0.1 - 0.2, while under stationary condition, no data is delivered off-chip.

## 2. SENSOR OPERATING PRINCIPLE

The proposed sensor extracts temporal contrast at pixel level, embedding an algorithm for dynamic background subtraction. In the presented implementation, the pixel is the main building block, executing in parallel the most power consuming operations of the vision algorithm. Its basic operating principle is shown in Fig.1, plotted over 20 frames. The activity over time of the current signal  $V_P$ , acquired during the integration time, is monitored over a certain number of frames by means of two voltage signals,  $V_{Min}$  and  $V_{Max}$ . These two voltages slowly keep track of the low values and the high values of  $V_P$ , respectively. The basic function of the two signals,  $V_{Min}$  and  $V_{Max}$ , is to define a voltage gap inside which  $V_P$  is to be considered normal, *cold-pixel* ( $V_{Mini} < V_{Pi} < V_{Maxi}$ ). This means that until the pixel behavior is in line with its past history, no anomalous events are detected. Outside the voltage boundary, the signal is recognized as anomalous (*hot-pixel*), detecting a potential alert situation ( $V_{Pi} > V_{Maxi}$  or  $V_{Pi} < V_{Mini}$ ). Depending on the binary status of the pixel, *hot-pixel* or *cold-pixel*, the voltage thresholds  $V_{Max}$  and  $V_{Min}$  will be updated as follows:

*Max Value:*

$$Q_{Max} = 1 \rightarrow V_{Max_{i+1}} = \alpha_H \cdot V_{Pi} + (1 - \alpha_H) \cdot V_{Max_i}, \quad (1)$$

$$Q_{Max} = 0 \rightarrow V_{Max_{i+1}} = \alpha_C \cdot V_{Pi} + (1 - \alpha_C) \cdot V_{Max_i}; \quad (2)$$

*Min Value:*

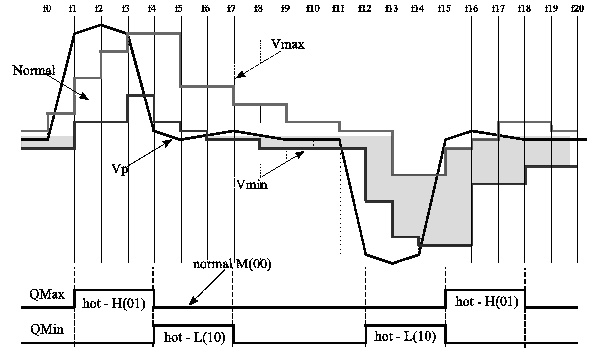
$$Q_{Min} = 1 \rightarrow V_{Min_{i+1}} = \alpha_H \cdot V_{Pi} + (1 - \alpha_H) \cdot V_{Min_i}, \quad (3)$$

$$Q_{Min} = 0 \rightarrow V_{Min_{i+1}} = \alpha_C \cdot V_{Pi} + (1 - \alpha_C) \cdot V_{Min_i}; \quad (4)$$

with  $\alpha_H > \alpha_C$ . According to the four equations, when the current value of the pixel exceeds one of the two thresholds, ( $V_P > V_{Max}$  or  $V_P < V_{Min}$ ), an *hot-pixel* is detected and the corresponding threshold is rapidly updated ( $\alpha_H$ ), to absorb the unusual signal variation within few frames. In case of a *cold-pixel*, the threshold slowly ( $\alpha_C$ ) converges toward  $V_P$ . The main advantage of this approach, over the classic temporal difference [12], consists of its adaptive capability, which makes it very reliable in detecting true alert situations.

## 3. PIXEL ARCHITECTURE

The basic pixel schematic is shown in Fig. 2. The photodiode (PD), working in storage mode, is buffered by a source follower (M1,M2,M3), which is turned ON by  $V_{p-clk}$  only when capacitors  $C_{1M}$  and  $C_{1m}$  have to be pre-charged. The two Switched-Capacitors Low-Pass Filters (SC-LPF1/SC-LPF2) are fed by  $V_P$  and compute  $V_{Max}$  and  $V_{Min}$  respectively, clocked at the sensor frame rate. The two filters share the



**Figure 1: Basic operating principle of the APS with thresholds adapting capabilities.**

first clock  $SETV_P$ , which is used to store the current value  $V_P$  on the two capacitors  $C_{1M}$  and  $C_{1m}$ . The SC-LPF1 (SC-LPF2) operation is managed by a digital control block (UPDATE REGISTER, Fig. 4), which is placed outside the imager and properly activates the phase MM (Mm), according with the high-level algorithm for scene interpretation. The output of the filter is stored onto the PMOS capacitor  $C_{2M}$  ( $C_{2m}$ ). The switches MSH1 (MSH2) and MSW1 (MSW2) are PMOS transistors with the bulk tied to  $V_{blk}$ , generated by the structure shown in the blow-up of Fig. 2 and placed in a dummy pixel. Here, the photodiode is shorted to  $V_{RES}$  and the output buffer is always on, generating  $V_{blk} = V_{RES} - V_{gsM1}$ . This makes all the switches of the filters to always turn off into the accumulation region, minimizing the the leakage current through the transistor channel [?].

The SC-LPF transfer function is:

$$H(s) = \frac{1}{1 + s\tau_n} = \frac{1}{1 + s \cdot \left( \frac{C_{2M}}{C_{1M}} \cdot \frac{n}{f_0} \right)}; \quad (5)$$

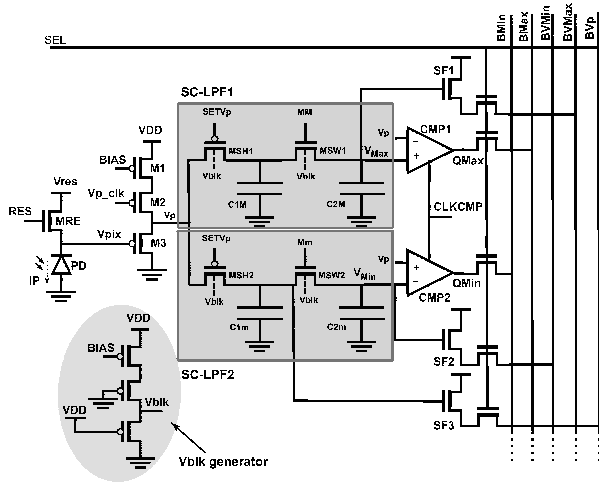
where  $C_{1M}$  and  $C_{2M}$  are two NMOS filter capacitors, with  $C_{2M}/C_{1M} = 220fF/150fF \sim 1.5$ . The integer  $n$  makes the filter to be clocked once over  $n$  frames, thus changing the SC-LPF response. With  $n = 1$ , the filter exhibits the fastest response, being clocked at each frame. Referring to the equations in Section II, the value of  $n$  directly affects the values of  $\alpha_H$  and  $\alpha_C$ , hence the behavior of the filters. The larger the  $n$ , the larger is the time constant of the filter ( $\tau_n$ ). The value of  $n$  can be digitally changed by properly clocking the MM (Mm) phase of the filter.

After the exposure time, a new value of  $V_P$  is acquired and the analog outputs of the two filters ( $V_{Max}$  and  $V_{Min}$ ) are simultaneously compared with respect to  $V_P$ , by clocking the comparators CMP1, CMP2 with CLKCMP. The pixel provides two binary outputs ( $Q_{Max}$ ,  $Q_{Min}$ ), which are directly available on the two bit-lines of Fig. 2 ( $B_{Min}$ ,  $B_{Max}$ ):

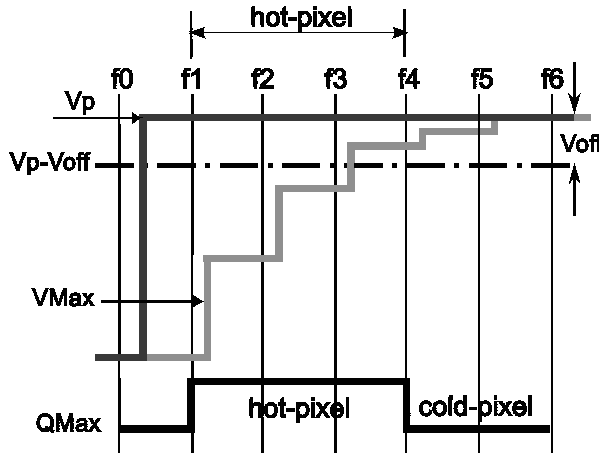
- $Q_{Max} \otimes Q_{Min} = L \rightarrow$  cold pixel;
- $Q_{Max} \oplus Q_{Min} = H \rightarrow$  hot pixel.

Depending on the pixel status, the analog memories  $C_{2M}$  and  $C_{2m}$  are updated through the activation of MM and/or Mm respectively by means of a proper control block (UPDATE REGISTER in Fig. 4), placed outside the pixel array.

In order to prevent false *hot-pixel* detection, the two comparators have a symmetric built-in offset ( $V_{off}$ ) (Fig. 3).



**Figure 2: Schematic of the Active Pixel Sensor.** In the blow-up, the reference pixel generates the bias voltage for the bulk (Vblk) of the PMOS switches of the filters (MSH1, MSW1, MSH2, MSW2). Here , the photodiode is always reset to Vres and the buffer is always ON (Vp.clk=0).



**Figure 3: Comparator offset defining a safe margin.** For  $V_{Max} > V_p - V_{off}$ , the pixel is considered a *cold-pixel*. In this example, signal  $V_p$  changes abruptly before  $f_1$ , setting an *hot-pixel* and forcing the pixel to adjust the  $V_{Max}$  threshold in order to compensate for the alert status.

The grey-zone represents the range of values inside which the pixel is detected as a *cold-pixel*. Under an *hot-pixel* detection ( $Q_{Max}$ ), the related memory is updated, changing the threshold voltage (e.g.  $V_{Max}$ ) toward the current value  $V_P$ . The *hot-pixel* status starts from frame  $f_1$  until  $f_4$ , being  $V_{Max} > V_P - V_{off}$ . Moreover, the pixel is equipped with three NMOS source-followers (SF1, SF2, SF3) buffering the three analog voltages ( $V_{Max}$ ,  $V_{Min}$ ,  $V_P$ ). The analog bit-lines,  $BV_{Max}$ ,  $BV_{Min}$ ,  $BV_P$ , deliver the analog signals outside the imager and feed three operational amplifiers, which are directly connected to the output pads. This last functionality was implemented

for debug and calibration purposes, while during the sensor operating condition, it is completely turned off.

## 4. SENSOR ARCHITECTURE

The proposed sensor consists of a  $64 \times 64$  pixels array with peripheral circuitry for array scanning, data readout and pixel control for the algorithm implementation. The ROW DECODER progressively selects the rows of the imager. The COLUMN DECODER delivers the binary output of the pixels in asynchronous way. The 64-stages UPDATE REGISTER can deliver both digital and analog signals coming from the bit-lines. Each stage consists of two D-type flip-flops, storing the binary values of the two bit-lines ( $BM_j$ ,  $Bmj$ ), and three analog switches, connecting  $V_{Max_j}$ ,  $V_{Min_j}$  and  $V_{P_j}$  to the three opamps. The temporal filter, implemented into the sensor, can work in two different modes: *learning mode* and *adaptive mode*. The learning mode is typically used during the system setup. The background of a static scene is stored into the two memories and compared with the current signal. It takes several frames to settle  $V_{Max}$  and  $V_{Min}$  by continuously updating the filters. In the *adaptive mode*, the memories are quickly updated under *hot-pixel* and slowly under *cold-pixel*, mimicking the eye adaptive behavior.

After the image acquisition phase and after comparing  $V_P$  with the two analog memories,  $V_{Max}$  and  $V_{Min}$  need to be updated before the binary data to be delivered off-chip. This operation takes place at each row selection. In this case, the bit-line pairs ( $B_{Max}$ ,  $B_{Min}$ ) of each column are used as masks for the memories update phase.

There are two ways to read out the binary informations from the sensor: asynchronous and mixed sync/asynchronous. In the asynchronous mode, after pulsing START, the first row of the array is selected by the ROW DECODER. As soon as the bit-lines are settled and the update phase has been accomplished, the COLUMN DECODER sequentially looks fro bit-line pairs disparities and for each pixel delivers signals signals, MAX synchronized on WRN, at a data rate of about 100MHz. One row will therefore take about  $650ns$  to be read out. If the current pixel is hot,  $V_P > V_{Max}$ , the corresponding output will be (MIN=0, MAX=1). After 64 WRN pulses, the End Of Row (EOR) is provided and the readout process stops, waiting for the acknowledge (ACK) to be provided off-chip, selecting next row. At the end of the readout process, an End Of Frame (EOF) is provided.

The second operating mode of the sensor is mixed, synchronous/asynchronous. At the START, the first row is selected and the binary bit-lines ( $BM_j$ ,  $Bmj$ ) of the current row are uploaded into the  $2 \times 64$  Flip-Flops stages UPDATE REGISTER, ready to be read out through DOUT, by providing an external clock (CLK). Next row is selected at the rising edge of ACK. EOR is provided by the chip at the end of each row readout. In this case, two binary images are delivered for each frame: one for the MAX hot-pixels and one for the MIN hot-pixels. An analog multiplexer is embedded into the UPDATE REGISTER, multiplexing the three analog voltage signals ( $V_{Max}$ ,  $V_{Min}$ ,  $V_P$ ) of the pixels of the selected row. This capability is only available under the synchronous/asynchronous operating mode, and needs a relatively slow clock rate ( $f_{CLK} < 2MHz$  typ). Although adding 6 NMOS and 3 bit-lines to each pixel will turn into a significant electronic overhead, this additional functionality is worth to be embedded, allowing sensor's self-test and diagnostic capabilities.

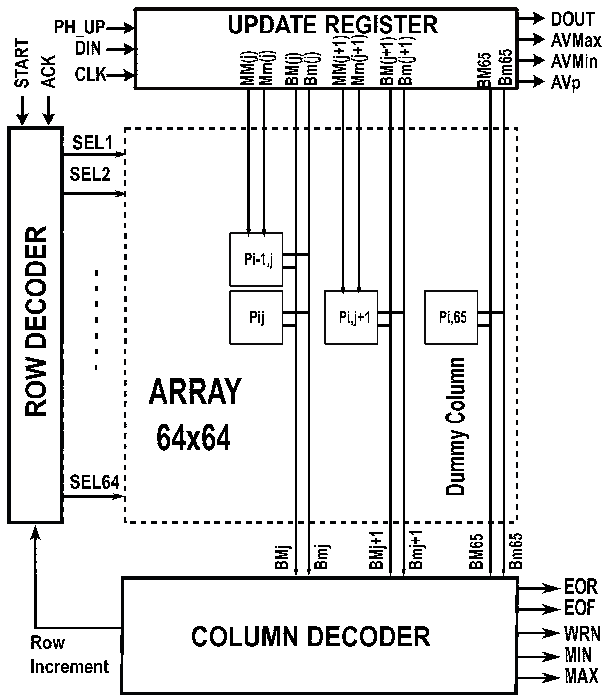


Figure 4: Block diagram of the vision sensor architecture.

## 5. EXPERIMENTAL RESULTS

Electro-optical tests have been carried out on the sensor, measuring the most important pixel characteristics: filter's transfer function and memory retention time, power consumption and computational performance.

### 5.1 Memories Update

The filter transfer function has been measured through an electrical test, applying at first a positive voltage step on the photodiode, thus simulating an abrupt light transition, from light-to-dark. Right after the two filters has been settled, a negative voltage step has been applied on the photodiode, mimicking a dark-to-light transition (Fig. 5). In the first case, an *hot-pixel* on  $V_{Max}$  ( $Q_{Max} = 1$ ) has been forced. Hence, the memory C2M is updated at every frame ( $n = 1$ ), by clocking MM, until the pixel returns into the cold-pixel status ( $Q_{max} = 0$ ), which takes about 15 frames. During this phase, the memory C2m, storing  $V_{Min} < V_P$ , is updated once every 4 frames ( $n = 4$ ), slowing down its time response. During the second phase, an *hot-pixel* is detected on  $V_{Min}$  ( $Q_{Min} = 1$ ) and the memories update process is complementary, as shown in Fig. 5. It is possible to notice a difference in the time response between the  $V_{Max}$  and  $V_{Min}$  memories. In the first case, the *hot-pixel* compensation takes about 15 frames, while in the second case, it takes more than 20 frames. This is due to some offset between the two filters. One reason could be that, despite both channels are identical, with similar leakage currents and coupling effects, the signals they have to deal with are complementary. However, this is not a serious problem, because each pixel is intended to work in a closed control loop, adjusting its time response according with the output of the high-level algorithm.

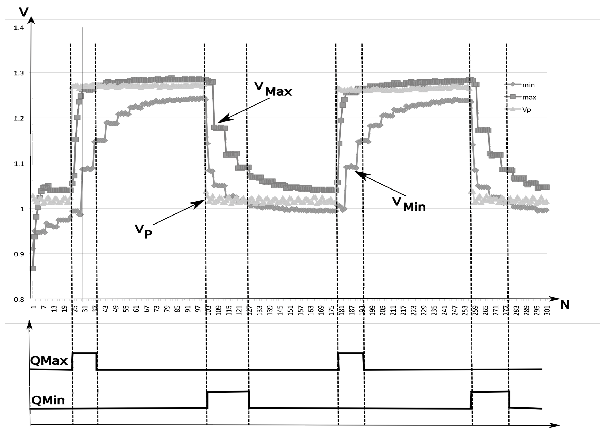


Figure 5: Example of the measured pixel response to a light step stimulus. *Hot-pixels* are updated with  $n = 1$ ; *cold-pixels* with  $n = 4$ . In this way, an asymmetric functionality is implemented, mimicking the eye behavior in presence of sudden light stimuli.

### 5.2 Memory Retention Time

Another important pixel parameter is the analog memory retention time. The two memories C2M and C2m are expected to store the two voltage thresholds,  $V_{Max}$  and  $V_{min}$ , for number of frames that can range from 1 up to 50, depending on the application. Making the sensor to operate at a frame rate of 10Hz, it means that the memory must retain information, without being updated, for about 5s. The pixel analog memories have PMOS switches with leakage current which makes the signal to drift toward increasing voltage values. Therefore, the worst case is given by the memory C2M pre-charged at the lowest value  $V_{Max}(t_0) = V_{sat}$ , with no update applied. In this case,  $V_{Max}$  drifts toward larger values. We measured the retention time by forcing  $V_P$  at its highest voltage ( $V_P = V_{dark}$ ), making the pixel to be hot ( $V_{Max}(t_0) < V_{dark}$ ). The comparator is clocked at 10Hz, comparing  $V_P$  with  $V_{Max}$ .  $Q_{Max} = 1$  for the time taken by  $V_{Max}$  to cover the voltage range of 750mV, from  $V_{sat}$  to  $V_{dark}$ . This time takes about 37 s (Fig. 6). It is obtained with the sensor placed under a light source making the photodiode to saturate in 100 $\mu$ s, which is very close to the worst case conditions. If we consider 5s to be the useful memory retention time, we can see that after 5s the memory loses 1LSB on a 4-bit scale.

### 5.3 Power Consumption

The minimum power consumption of the chip has been measured under mixed synchronous/asynchronous operating mode, with a 2MHz readout clock (CLK) and a frame rate of 13fps. The three operational amplifiers dedicated to the analog imager readout, have been turned off, as well as all the three analog source-followers inside each pixels of the array (SF1, SF2, SF3), by forcing the analog bit-lines to VDD (BVp, BVMax, BVMin). The 1 $\mu$ A pixel buffer is turned-on for 10 $\mu$ s only after the integration time, needed to charge  $V_P$  on the two capacitors C1M and C1m. In this operating condition, the sensor exhibits a total power consumption of 66 $\mu$ W. However, it has to be considered that the BIAS block, shown on the top-left side of Fig. 8, providing the temperature compensated bias voltage to the pixels (BIAS in Fig. 2),

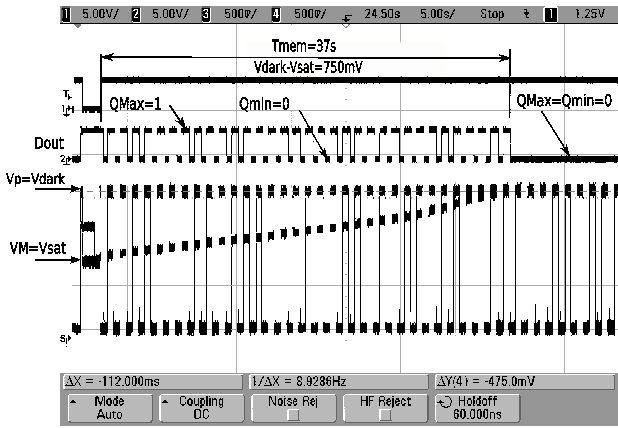


Figure 6: Retention time of the analog memory C2M (C2m) measured under an impinging light forcing the photodiode PD to saturate in  $100\mu s$ . The memory covers the entire signal range (750mV) in about 37s.

Table 1: Main chip characteristics

Parameter	Value
Technology	0.35 $\mu m$ 2P-4M
Sensor resolution	64 $\times$ 64 pixels
Chip Size (pads included)	6.6 mm <sup>2</sup>
Pixel Pitch	26 $\mu m$ $\times$ 26 $\mu m$
Number of transistors/pixel	45
Fill Factor	12%
Memory retention time	0.2 LSB/s 4-bits scale
Supply Voltage	3.3 V
Pow. Cons./pixel	620pW/frame $\cdot$ pixel
Computing performance - $P_E$	42 GOPS/W
Computing density - $P_A$	4 GOPS/mm <sup>2</sup>

consumes 33 $\mu W$  (10 $\mu A$ ). This means that, the vision sensor alone burns 33 $\mu W$ . Based on this value, the derived power consumption per pixel per frame is 620pW/pixel  $\cdot$  frame. This value is larger than that of other low power vision sensors [6],[7]. However, the absolute value of power consumption is not to be consider as an appropriate figure of merit for this class of sensors. More interestingly is the computing performance related to the power consumption. Considering that each pixel executes 26 operations per frame on a gray-scale, the total estimated computing performance of the sensor is 42 GOPS/W and about 4 GOPS/mm<sup>2</sup>. The obtained values are competitive with respect to most advanced computational vision sensors [2], [14], [15], [16].

## 5.4 Operating Functions

The sensor has been tested on moving objects to validate the vision algorithm. In Fig. 7 the sensor looks at a hand moving left and right periodically. At the beginning, this movement activates a large number of hot-pixels in the sensor. After a certain number of frames, the sensor starts to progressively absorb these pixels until the image returns to be black.

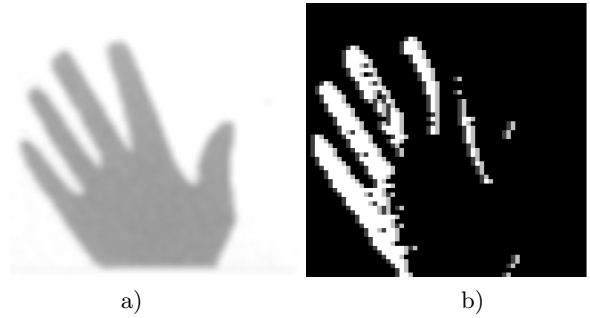


Figure 7: Example of hot-pixel detection. a) periodically left and right moving hand acquired by means of one of the three analog channels of the sensor; b) binary image resulting from the sensor image processing. The hot-pixels are white.

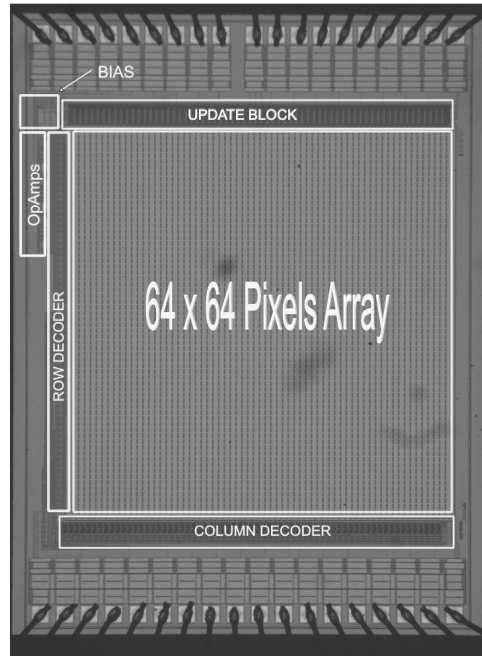


Figure 8: Photograph of the vision sensor chip.

## 6. CONCLUSION

A 64  $\times$  64 pixels low-power vision sensor has been presented, performing 26 $\mu m$  pixel pitch and 12% fill-factor in a 0.35 $\mu m$  CMOS technology. The sensor executes an algorithm for adaptive background subtraction, aimed at detecting anomalous pixel behaviors, which could be part of a potential event, occurring in the scene. The algorithm is mainly implemented at pixel-level, mimicking the eye adaptation behavior. Each pixel embeds two switched-capacitors low-pass filters (SC-LPF1/2), which can be individually controlled from outside the imager, thanks to a custom control logic (UPDATE REGISTER), which can properly mask each filter's clock signal. The sensor power consumption has been minimized, reducing the operating duty-cycle of the voltage buffer inside each pixel, which is the main source of dc power consumption. Thanks to the pixel-level advanced image processing, the output of the sensor can be directly binarized, allowing the successive image processing to be

fully binary and much easier to be implemented at a very low supply voltage and reduced energy budget. The computing performance of the sensor over power consumption is 42 GOPS/W.

## 7. ACKNOWLEDGMENTS

This work was partially supported by the Project BOViS ("A Battery Operated Vision System for Wireless Sensor Network Applications"), within the Italy-Israel Cooperation Program 2009.

## 8. REFERENCES

- [1] F. Gregoretti, R. Passerone, L. M. Reyneri, and C. Sansoe, "A High Speed VLSI Architecture for Handwriting Recognition," *Journal of VLSI Signal Processing* 28, 259-278, 2001.
- [2] T. Komuro, I. Ishii, M. Ishikawa, and A. Yoshida, "A digital vision chip specialized for high-speed target tracking," *IEEE Trans. Electron Devices*, vol. 50, no. 1, pp. 191-199, 2003.
- [3] Komuro, T.; Kagami, S.; Ishikawa, M.; , "A dynamically reconfigurable SIMD processor for a vision chip," *IEEE Journal of Solid-State Circuits*, vol.39, no.1, pp. 265- 268, Jan. 2004.
- [4] K. Kagawa, S. Shishido, M. Nunoshita, and J. Ohta, "A 3.6 pw/frameá pixel 1.35 v pwm cmos imager with dynamic pixel readout and no static bias current," in *Solid-State Circuits Conference, 2008. ISSCC 2008. Digest of Technical Papers, IEEE, 2008*, pp. 54-55.
- [5] K. Cho, D. Lee, J. Lee, G. Han, "Sub-1-V CMOS image sensor using time-based readout circuit," *IEEE Transactions on Electron Devices* vol. 57, No.1, Jan. 2010, pp. 222-227.
- [6] S. Hanson, D. Sylvester, "A 0.45-0.7V Sub-Microwatt CMOS Image Sensor for Ultra-Low Power Applications", *Symp. on VLSI Circuits*, 2009, pp. 176-177, 2009.
- [7] F. Tang, Y. Chao, A. Bermak, "An Ultra-Low Power Current-Mode CMOS Image Sensor with Energy Harvesting Capability," *Proc. of European Solid State Circuits Conf.*, pp. 126-129, Sept. 2010.
- [8] Dongsoo Kim, Zhengming Fu, Joon Hyuk Park, and Eugenio Culurciello, "A 1-mW CMOS Temporal-Difference AER Sensor for Wireless Sensor Networks," *IEEE Transactions on Electron Devices*, vol. 56, no. 11, pp. 2586-2593, November 2009.
- [9] M. Gottardi, N. Massari, S.A. Jawed, "A 100uW 128x64 Pixels Contrast-Based Asynchronous Binary Sensor for Sensor Network Applications," *IEEE Journal of Solid State Circuits*, vol. 44, no. 5, pp. 1582-1592, May 2009.
- [10] Suat U. Ay, "A 1.32pW/frame-pixel 1.2V CMOS Energy-Harvesting and Imaging (EHI) APS Imager," *Digest of ISSCC 2011, February 2011, San Francisco*.
- [11] Miniature Autonomous Agents For Scene Interpretation, US Patent 7,489,802 B2, Filing Date Feb. 10, 2009.
- [12] V. Gruev, R. Etienne-Cummings, "A pipelined temporal difference imager," *IEEE Journal of Solid-State Circuits* vol.39, N. 3, pp. 538-543, March 2004).
- [13] M. O'Halloran, and R. Sarpeshkar, "A 10-nW 12-bit Accurate Analog Storage Cell With 10-aA Leakage," *IEEE Journal of Solid-State Circuits*, Vol. 39, No. 11, pp. 1985-1996, Nov. 2004.
- [14] A. Lopich, P. Dudek, "A SIMD Cellular Processor Array Vision Chip With Asynchronous Processing Capabilities," *IEEE Trans. on Circuits and Systems*, vol. 58, no. 10, pp. 2420-2431, Oct 2011.
- [15] G. C. Linan, A. Rodriguez-Vazquez, and R. C. Galan et al., "A 1000 FPS at 128 x 128 vision processor with 8-bit digitized I/O," *IEEE J. Solid-State Circuits*, vol. 39, no. 7, pp. 1044-1055, 2004.
- [16] M. Wei, L. Qingyu, Z. Wancheng, and W. Nan-Jian, "A programmable SIMD vision chip for real-time vision applications," *IEEE J. Solid-State Circuits*, vol. 43, no. 6, pp. 1470-1479, 2008.