A Bio-Inspired APS for Selective Visual Attention

Nicola Cottini, Massimo Gottardi, *Member, IEEE*, Nicola Massari, *Member, IEEE*,
and Roberto Passerone, *Member, IEEE*

***Abstract*— This letter presents an active pixel sensor architecture that mimics the eye adaptive behavior through dynamic background subtraction as a basis for image interpretation tasks. The pixel extracts temporal contrast as a low-level visual feature for successive high-level processing targeted to monitoring applications. To do this, each pixel compares the acquired value with a voltage range computed over the past pixel behavior; values outside this range signal a potential alert. Two switched-capacitor programmable low-pass filters are embedded. The 26 μm, 12% fill-factor, fully functional 45 T pixel has been fabricated in a 0.35 μm CMOS technology.**

***Index Terms*— Vision sensors, active pixel sensors, visual attention, early visual features, background subtraction.**

I. INTRODUCTION

THIS letter presents a prototype of APS that detects temporal changes in light intensity [1] as a primitive feature for pre-attentive visual perception. The proposed architecture takes inspiration from bottom-up visual attention, where the first processing stage involves a massive computation of early visual features that are then mapped onto a higher representation [2]. Motion is one of the visual features playing a fundamental role in the process of visual attention. Motion can be detected by computing the changes $(F_i - F_{i-1})$ between two successive frames, and comparing them against a threshold *Th*. Despite its straightforward implementation [3]–[6], this technique is very sensitive to the threshold *Th* and works properly only in particular conditions of object speed and frame rate. A more reliable technique makes use of the running average, so that the background model is based on the pixel recent history.

II. OPERATING PRINCIPLE

In the proposed pixel architecture, the background B is modeled as an exponential moving average:

$$B_{i+1} = \alpha \cdot F_i + (1 - \alpha) \cdot B_i,$$

where F_i is the current frame and $\alpha < 1$ is the smoothing factor. Based on this model, the presented algorithm defines two signal thresholds V_{Max} and V_{Min} inside which the current value V_P of the pixel is considered normal ("cold pixel"), carrying no information. In case $V_P > V_{\mathrm{Max}}$ or $V_P < V_{\mathrm{Min}}$ the pixel is labeled as "hot pixel", indicating a potential alert condition. These values change over time, adapting to the signal dynamics as an exponential moving average, where the smoothing factor α depends on the current status of the pixel:

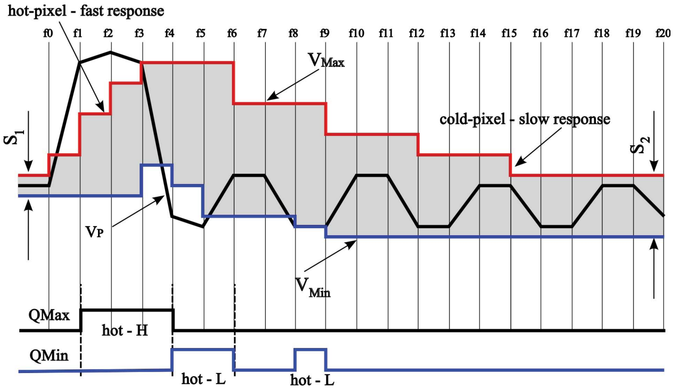

Fig. 1. Simplified representation of the evolution of the gray region based on adaptive background-subtraction, implemented at the pixel-level.

$$V_{P,i} > V_{\mathrm{Max},i} \rightarrow V_{\mathrm{Max},i+1} = \alpha_H \cdot V_{P,i} + (1 - \alpha_H) \cdot V_{\mathrm{Max},i},$$
$$V_{P,i} < V_{\mathrm{Max},i} \rightarrow V_{\mathrm{Max},i+1} = \alpha_C \cdot V_{P,i} + (1 - \alpha_C) \cdot V_{\mathrm{Max},i},$$

where $\alpha_H > \alpha_C$. The formulas indicate that when the value of the pixel exceeds a threshold, the corresponding value is updated rapidly (α_H), to try to absorb the unusual signal variation, while the other threshold follows the new value more slowly (α_C). When the pixel is cold ($V_{P,i} < V_{\mathrm{Max},i}$), both thresholds slowly converge toward V_P. A graphical representation of the proposed algorithm is shown in Fig. 1, where the gray area denotes the interval between V_{Min} and V_{Max} inside which the pixel V_P can change without generating an alert. The wider the interval, the less sensitive is the pixel. If the light intensity is constant, V_{Max} and V_{Min} converge toward V_P maintaining a minimum distance, as shown in the leftmost part of Fig. 1 (S1). This guarantees the highest sensitivity, as well as immunity to the electrical and to the image noise. If V_P rapidly changes between two values, the pixel decreases its sensitivity to compensate for the anomalous light changes by separating V_{Max} and V_{Min}, as shown in the rightmost part of Fig. 1 (S2). It slowly resumes its maximum sensitivity after the suspicious event is absorbed. The pixel state is encoded by two digital signals Q_{Max} and Q_{Min} updated at the end of each frame to reflect the current situation (Fig. 1 - bottom).

III. PIXEL SCHEMATIC

The photodiode (PD) of Fig. 2 works in storage mode, buffered by a PMOS source follower, which is turned on by Vp_clk only when necessary, reducing the pixel DC power consumption. V_{Max} and V_{Min} are stored in two analog memories. The two SC-LPF1/2, fed by V_P, compute V_{Max} and V_{Min}

Manuscript received July 13, 2012; accepted May 11, 2013. Date of publication May 17, 2013; date of current version August 6, 2013. This work was supported by the Project "A Battery Operated Vision System for Wireless Sensor Network Applications," within the Italy-Israel Coop. under Program 2009. The associate editor coordinating the review of this letter and approving it for publication was Prof. Julian C. C. Chan.

N. Cottini, M. Gottardi, and N. Massari are with the Fondazione Bruno Kessler, Povo, Trento 38050, Italy (e-mail: cottini@fbk.eu; gottardi@fbk.eu; massari@fbk.eu).

R. Passerone is with the Dipartimento di Ingegneria e Scienza dell'Informazione, University of Trento, Italy (e-mail: roberto.passerone@unitn.it).

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSEN.2013.2263962

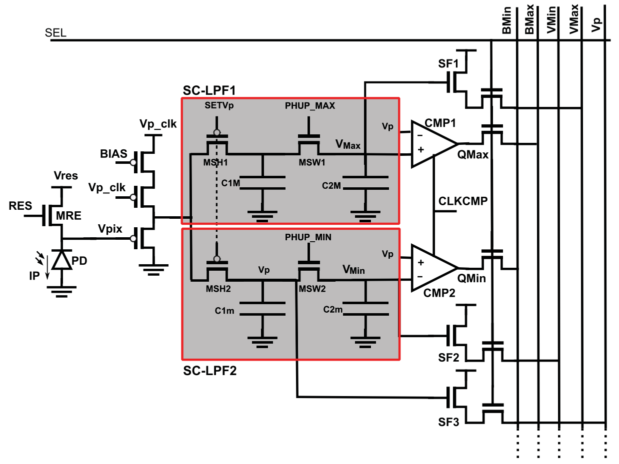

Fig. 2. Pixel schematic, embedding two switched-capacitor low-pass filters (SCLPF1, SCLPF2) and two clocked comparators (CMP1, CMP2). The two digital bitlines (BMax, BMin) detect the "hot/cold" status of the pixel.

TABLE I
MAIN PIXEL CHARACTERISTICS

Parameter	This Letter	Kim [3]	Mallik [4]
CMOS Process (μm)	0.35	0.5	0.5
Pixel size (μm)	26×26	29×28	25×25
Fill factor	12%	23%	17%
FPN	2.5%	0.4%	0.5%
Dynamic Range	52dB	-	51dB
Supply voltage	3.3 V	2.5V - 3V	3V
Power consumption	580pW/frame	8nW/frame	17nW/frame

respectively, with a clock running at the frame rate. The output of the filter is stored onto the PMOS capacitor C2M (C2m). At the end of the exposure time, two clocked comparators (CMP1, CMP2) [7] compare V_P with V_{Max} and V_{Min} respectively, providing the two bits Q_{Max} and Q_{Min}, which represent the state of the pixel. The SC-LPF transfer function is:

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

where C_{1M} and C_{2M} are the filter capacitors, with $C_{2M}/C_{1M} \sim 1.5$, f_0 is the frame rate, and the integer $n > 0$ indicates that the filter is activated (clocked) once every n frames. Thus, the filter response can be adjusted by changing its activation rate n, as described in the next section.

IV. EXPERIMENTAL RESULTS

Table I summarizes the main pixel characteristics, while Fig. 3 shows the measured response to light variations.

When V_P crosses V_{Max}, we update V_{Max} with the fastest time constant, by activating the filter at every frame ($n = 1$), while V_{Min} slowly follows V_P with $n = 4$. The behavior is dual when V_P crosses V_{Min}. Thus, if the higher level algorithm does not recognize an alert, the pixel quickly adjusts the thresholds to the new situation aimed to suppress the event. Under a "cold pixel", both filters are activated every 4 frames ($n = 4$), to let the outputs slowly converge toward V_P, retrieving high sensitivity, thus mimicking the eye adaptation process. This techniques allows the pixel to quickly absorb sudden changes in light intensity, while keeping the memory of past signal variations, preventing potential false alerts. The binary signals (Q_{Max}, Q_{Min}) generated by each pixel are collected at the array level and can be processed outside the chip by the higher-layer algorithm, implementing complex vision tasks. The value is compliant with typical higher level algorithm specifications. The power consumption of the pixel is 7.5nW at 3.3V, assuming a frame rate of 13 fps, measured by pulsing the 1μA buffer (Vp_clk) after the integration time, when C_{1M} and C_{1m} are pre-charged, thus reaching a duty-cycle of 10 μs/76ms = 1/7600. This turns into a 580 $pW/frame$. This compares favorably to the state of the art at the system level, considering that our pixel directly performs adaptive computation and binarization, therefore drastically reducing the data to be delivered and processed outside the sensor. During the 10μs ON-state, the buffer drains a current of 1μA, which is sufficient to charge the two 100fF capacitors, C_{1M} and C_{1m}. Although such large values of capacitors significantly increase the pixel pitch, they are needed to mitigate charge feedthrough and leakage effects, caused by the impinging light.

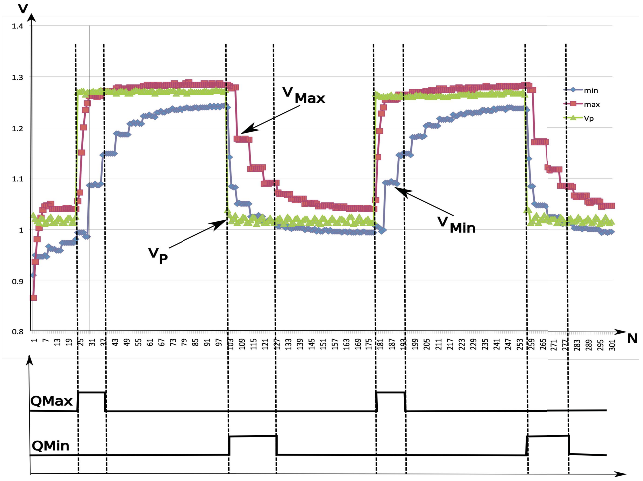

Fig. 3. Example of the measured pixel response to a light step stimulus. "Hot pixels" are updated with $n = 1$; "cold pixels" with $n = 4$.

REFERENCES

[1] Z. Smilansky, "Miniature autonomous agents for scene interpretation," U.S. Patent 7 489 802 B2, Feb. 10, 2009.

[2] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 11, pp. 1254–1259, Nov. 1998.

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

[4] U. Mallik, M. Clapp, E. Choi, G. Chauwenberghs, and R. Etienne-Cummings, "Temporal change threshold detection imager," in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, Feb. 2005, pp. 362–363.

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

[6] M. Zhang and A. Bermak, "Quadrant-based online spatial and temporal compressive acquisition for CMOS image sensor," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 19, no. 9, pp. 1525–1534, Sep. 2011.

[7] Y.-T. Wang and B. Razavi, "An 8-bit 150-MHz CMOS A/D converter," *IEEE J. Solid State Circuits*, vol. 35, no. 3, pp. 308–317, Mar. 2000.