

# Reconfigurable Focal-Plane Hardware for Block-Wise Intra-Frame HDR Imaging

Jorge Fernández-Berni, Ricardo Carmona-Galán, Ángel Rodríguez-Vázquez  
Institute of Microelectronics of Seville (CSIC - Universidad de Sevilla)  
C/ Américo Vespucio s/n, 41092, Seville, Spain. Phone: +34954466666  
Contact email: berni@imse-cnm.csic.es

**Abstract**—This paper presents circuitry conceived to be part of a massively parallel image processing array. Its main feature is a high degree of reconfigurability in order to extract different regions of interest (ROIs) across intra-frame high dynamic range (HDR) scenes. To this end, two photodiodes are included at each processing element (PE). One of them senses the pixel value whereas the other one, working collaboratively with the rest of PEs of the ROI, senses the average incident illumination on the region. As a result, the photointegration period for each ROI is asynchronously set during a single image acquisition stage. Another remarkable characteristic of the hardware is its low-power operation. We report some experimental results based on a previous prototype smart image sensor. By exploiting its processing capabilities, we emulate the behavior of the proposed circuitry. Direct applicability is demonstrated for the Viola-Jones vision algorithm.

## I. INTRODUCTION

When it comes to image sensing and processing, CMOS technologies presents a fundamental advantage over CCD technologies: the possibility of incorporating processing circuitry close to the photosensitive devices. This enables the VLSI implementation of smart image sensors [1]–[4] capable of delivering not simple raw images but elaborated scene representations. These representations are attained in a very efficient way, greatly lightening subsequent stages along the vision processing chain. One of the primary aspects that smart image sensors have to deal with is the potential existence of scenes featuring a wide range of intra-frame illuminations. In such cases, assuming the usual photocurrent integration mode, the application of the same exposure time across the pixel array leads to either over-exposed or under-exposed regions. The consequent low contrast and missing details can make the processing of these regions unfeasible. Numerous techniques, each differently compromising the performance parameters of the sensor, have been proposed to address this issue [5]. Most of them are focused on providing high-quality images. However, this does not have necessarily to be a strong requirement in artificial vision. There are vision algorithms that can afford to work with lower quality images provided that a minimum amount of information is still available. And lower quality images can be achieved by simpler and more effective techniques, increasing the performance of the system as a whole. The hardware proposed in this paper fits this scenario, being specially suitable for algorithms demanding ROI extraction [6]–[8]. It is based on the realization of a

massively parallel processing array. The interconnections between its constituting mixed-signal PEs are set block-wise through peripheral circuitry. The exposure period of the pixels within each block is independently determined according to the average local incident illumination. As a result, different regions of interest can concurrently be extracted across an HDR scene through a single image acquisition stage. While certainly reducing the fill factor when compared to other approaches previously reported, our proposal presents significant advantages. First of all, the extremely simple control required. Once the pixel reset is done, the only control operation remaining is to finish the image acquisition at the temporal point demanded by the prescribed frame rate. During that time interval, self-adapting sensing takes place within each preconfigured block. Other strength is precisely the reconfigurable nature of the hardware, enabling ROI extraction for any kind of rectangular shapes. Finally, the circuitry features low-power voltage-mode operation. This is a key point for application scenarios where energy-efficient smart image sensors are required, e.g. robotics, autonomous surveillance or UAVs. In order to show some results expected from a chip — currently under design — including this hardware, we present an experimental emulation of its operation. This emulation is accomplished by making the most of the processing capabilities of one of our previous prototype smart image sensors.

## II. PROPOSED CIRCUITRY

Fig. 1(a) shows the circuitry to be included at the PEs in order to achieve block-wise intra-frame HDR imaging. It is meant to operate conjunctively with a similar implementation of the peripheral hardware for column-wise and row-wise focal-plane division control described in [2]. This hardware has demonstrated to be easily programmable and power-efficient, providing a fast way to establish different interconnection patterns between the PEs of a processing array. Let us now explain how the proposed circuitry works, illustrating its operation by means of the timing diagram in Fig. 1(b). Two photodiodes and two sensing capacitances are used. They are nominally identical, but this is not a major limitation in terms of strict matching requirements, as will be seen later. Their reset to the voltage  $V_{rst}$ , which coincides with the upper limit of the signal range, is controlled through the signals RST\_EN and PI\_EN.

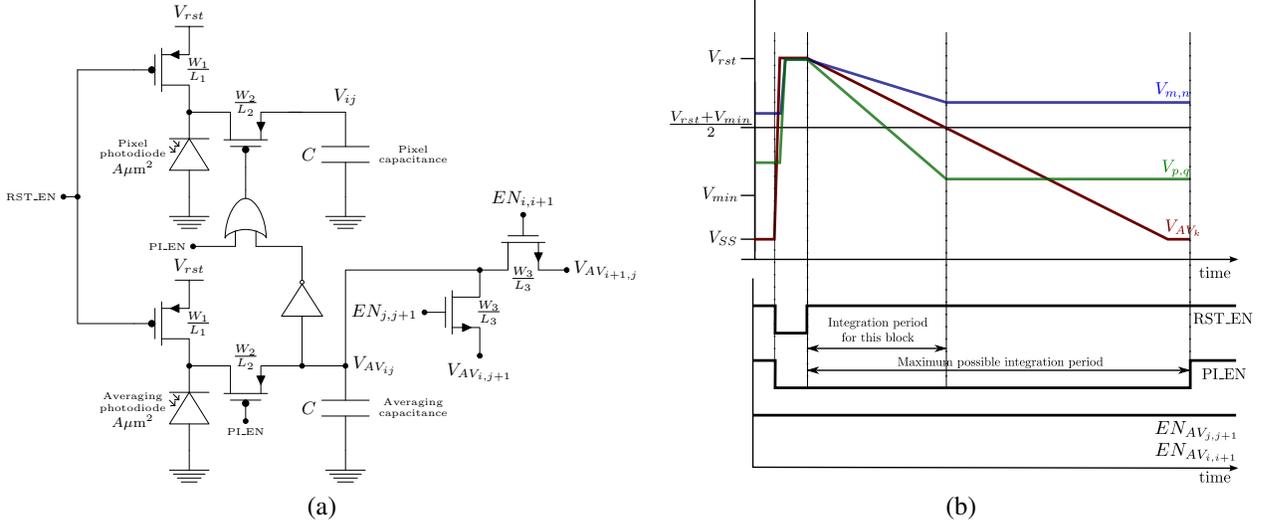


Fig. 1. Proposed circuitry to be included at the PEs of a smart image sensor (a) and an illustrative timing diagram (b).

When both of them are set to logic ‘0’, the reset of the so-called pixel photodiode, averaging photodiode and averaging capacitance is immediately enabled. If the voltage  $V_{AV_{ij}}$  starts from a value above the input threshold voltage of the inverter,  $V_{th_{inv}}$ , the reset of the pixel capacitance is concurrently initiated. Otherwise, as is the case in Fig. 1(b), this reset will be delayed slightly until  $V_{AV_{ij}}$  reaches that threshold. A key aspect of the circuit is that the inverter must be designed in such a way that  $V_{th_{inv}}$  is located just at the middle point of the signal range, that is:

$$V_{th_{inv}} = \frac{V_{rst} + V_{min}}{2} \quad (1)$$

where  $V_{min}$  is the lower limit of the signal range. Thus, the integration interval for each block will be determined asynchronously according to its incident illumination, once RST\_EN is set back to ‘1’. At that time instant, photointegration begins concurrently in both the pixel capacitance and the averaging capacitance. However, while in the former it is carried out in an isolated way, charge redistribution takes place in the latter among the averaging capacitances interconnected through the switches controlled by  $EN_{i,i+1}$  and  $EN_{j,j+1}$ . Keep in mind that these signals come from the peripheral circuitry previously commented. As a result, the voltage excursion<sup>1</sup> due to photointegration for each pixel within a certain block  $k$  is given by:

$$\Delta V_{ij} = \frac{I_{ph_{ij}}}{C} T_k \quad (2)$$

where  $I_{ph_{ij}}$  denotes the sum of the photogenerated current plus the dark current and  $T_k$  is the photointegration period, the same for all the pixels composing the block. In order to obtain an expression for this period, it must

be observed that, since the averaging capacitances are interconnected, the following equation holds:

$$\sum_{\forall i,j \in k} I_{ph_{ij}} = -WHC \frac{dV_{AV_k}}{dt} \quad (3)$$

where  $W \times H$  are the dimensions of the block, in pixels, and  $V_{AV_k}$  is the voltage at the averaging capacitances, the same at each of them due to the charge redistribution taking place constantly. The photointegration at the pixel capacitances will finish when  $V_{AV_k}$  reaches  $V_{th_{inv}}$ , that is, from Eq. 3:

$$T_k = WHC \frac{V_{rst} - V_{th_{inv}}}{\sum_{\forall i,j \in k} I_{ph_{ij}}} \quad (4)$$

which can be expressed, taking into account Eq. 1, as:

$$T_k = \frac{C}{2} \frac{\Delta V_{ij_{MAX}}}{\bar{I}_{ph_k}} \quad (5)$$

where  $\Delta V_{ij_{MAX}} = V_{rst} - V_{min}$  represents the maximum pixel excursion, and  $\bar{I}_{ph_k}$  is the block average current generated during the photointegration period. If the effect of the dark currents can be neglected, this current is directly proportional to the average incident illumination on the block. By substituting Eq. 5 in Eq. 2, the following voltage excursion for each pixel is obtained:

$$\Delta V_{ij} = \frac{\Delta V_{ij_{MAX}}}{2} \frac{I_{ph_{ij}}}{\bar{I}_{ph_k}} \quad (6)$$

where we can see that the maximum pixel illumination to be detected without saturation is double of the average illumination of the block. It is this property, together with the possibility of confining its application to any particular rectangular-shaped block, what really endows our hardware with the capability of extracting ROIs across HDR frames. As an example, the dynamics of two pixels belonging to the same block along with the corresponding local averaging voltage is depicted in Fig. 1(b). In blue, a pixel whose illumination is below the

<sup>1</sup>For the sake of simplicity, we define the voltage excursion as the difference between the initial voltage,  $V_{rst}$ , and the final voltage. This permits to get rid of the minus sign.

average block illumination is represented. In green, we show a pixel receiving illumination above that average. Notice that the integration period for all the pixels of the block ends at the same time instant without requiring any external control. That instant is set when the sensing of the average incident illumination reaches the middle point of the signal range. Nevertheless, the signal PI\_EN permits to finish the photointegration for those regions whose illumination is so low that it does not reach that middle point within the limit established by the prescribed frame rate. Notice also that most of the energy  $E$  required to complete the operation is injected during the reset stage:

$$E = MN(C + C_{ph})V_{rst}^2 \quad (7)$$

where  $M \times N$  is the resolution of the array and  $C_{ph}$  represents the capacitance of the photodiodes. Considering, for instance, a VGA array with  $C = 100\text{fF}$ ,  $C_{ph} = 20\text{fF}$ ,  $V_{rst} = 1.5\text{V}$  and 30fps, the power consumption would only amount to  $\sim 2.5\mu\text{W}$ . There is also some consumption associated with the switching of the digital circuitry at the periphery and the two digital gates at each PE, but it should definitely be negligible at typical frame rates. The worst scenario that could happen is for those blocks featuring low illumination, and consequently slow photointegration discharge. In such a case, the averaging voltage  $V_{AV_k}$  could remain close to  $V_{th_{inv}}$  during a non-negligible fraction of the photointegration period, drawing current from the inverters forced to work in their transition region. This highlights again the importance of a careful design of this logic gate, not only by adjusting its input threshold but also by making its voltage transfer characteristic as abrupt as possible.

To finish this section, we must clarify the implicit assumption in Eqs. 2 and 3 that the current generated during the integration period coincides at both the pixel photodiode and the averaging photodiode. There are two major sources of non-ideality. First of all, the unavoidable mismatch of any physical implementation, which is translated into a slight deviation with respect to the nominal integration period. Anyway, we will demonstrate in the next section that a very accurate adjustment of that period is not so crucial, at least for the vision algorithm analyzed. And second, the existence of textures featuring a spatial resolution smaller than the maximum spatial resolution observable with the pixel pitch available. In this case, the deviation between the sensed illuminations could certainly be significant, depending on the characteristics of the texture. As a concluding remark, we must say that the idea of using two photodiodes per PE to capture HDR images is not new. It was already reported in [9], but the approach to the problem is very different. In that work, the authors posed a time-to-saturation technique where a photodiode is used to detect saturation and the other holds the adequate value of the pixel. The control scheme proposed is more complex than the one described in this paper and no ROI adaptation is provided.

### III. EXPERIMENTAL RESULTS

In order to emulate the behavior of the circuitry just presented, we have exploited the focal-plane processing capabilities of the FLIP-Q prototype [2] as well as the external memory and processing resources of the Wi-FLIP system [10]. The FLIP-Q prototype already implements block-wise focal-plane reconfigurability and block averaging. Unfortunately, this averaging is destructive with respect to the pixel values, what forces us to first find the integration time  $T_k$  for each block. It is accomplished by setting different integration periods. For each period, and for each block, we check the corresponding average value, registering the period that leads to the closest averaging to the middle point of the signal range. A matrix is eventually obtained in which a certain integration period is assigned to each block. This subset of integration periods is then swept, disabling previously the averaging operation in order not to destruct the pixel values. As an example, consider Fig. 2. The first column shows a scene uniformly illuminated on which we apply the Viola-Jones algorithm for face detection<sup>2</sup> [6]. All the faces are detected, together with three false positives. In the second column, the same scene presents now a noticeable non-uniform illumination, causing a number of saturated pixels at the upper left area. Both images — uniformly and non-uniformly illuminated — were captured by Wi-FLIP considering the whole frame as a single block. The same exposure period, 992ms and 11ms respectively, was therefore applied to all the pixels according to the average global illumination. In the first case, this approach is valid due to the uniform illumination. In the second case, it does not work. Face detection is not possible in the saturated area since most of the information has been lost. Note on the other hand the robustness of the algorithm for detection in dark areas. This robustness arises from the cascade of classifiers supporting its operation. These classifiers are normalized in order to handle different lighting conditions. As a result, even when feeding the algorithm with low-quality data, it still performs well. This is confirmed by the results shown in the third column<sup>3</sup>. Here, in order to retrieve the missing details at the over-exposed upper left area, we have established a regular  $2 \times 2$  block division. The integration period for the pixels within each block is independently determined according to Eq. 5. Since the sweeping of exposure periods cannot be carried out with infinite temporal resolution, we set search steps of  $100\mu\text{s}$ . The real hardware should provide a much finer resolution. The artifacts generated by this scheme are quite noticeable, but from the perspective of the algorithm, they do not disturb: all the faces are detected without false positives. This absence of false positives deserves further consideration. The proposed block division is capable of providing the algorithm with

<sup>2</sup>Specifically, we apply the implementation of this algorithm provided by the OpenCV library, available at [opencv.willowgarage.com](http://opencv.willowgarage.com).

<sup>3</sup>All the images shown in Fig. 2 together with others demonstrating the capability of adaptation to regions of different sizes can be found at [www.imse-cnm.csic.es/mondego/IISW2013/](http://www.imse-cnm.csic.es/mondego/IISW2013/)

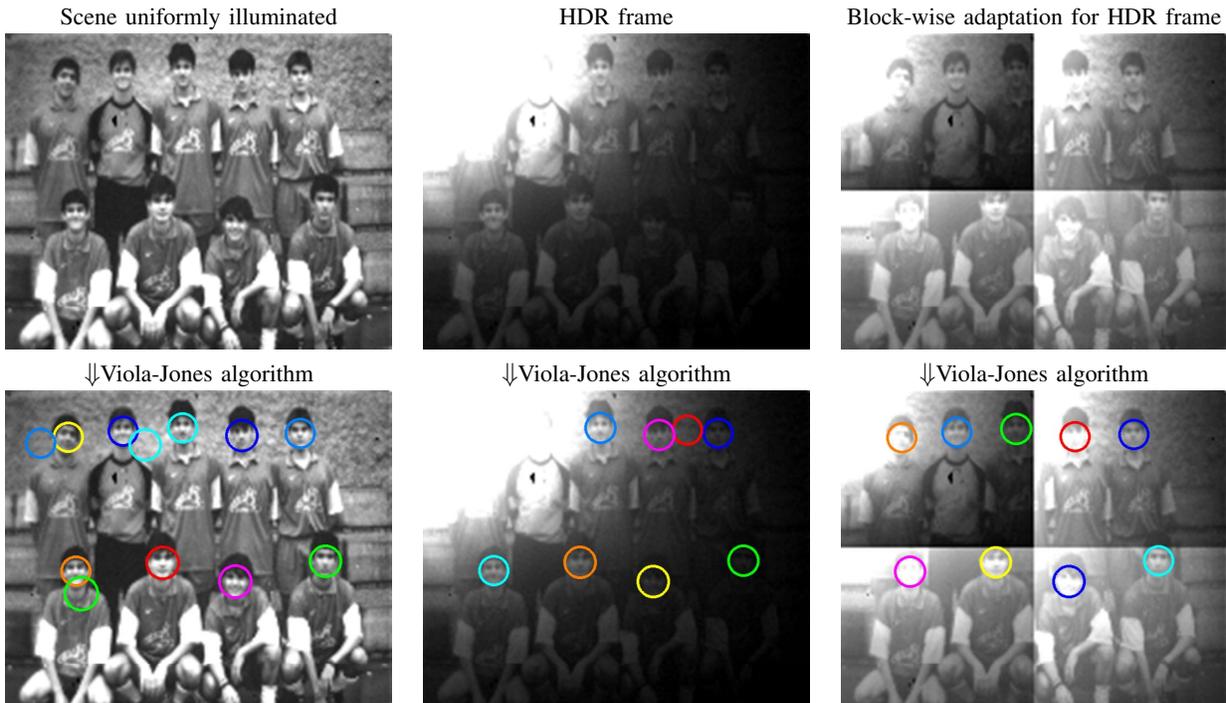


Fig. 2. Experimental results demonstrating that block-wise HDR imaging can be very effective to provide a vision algorithm with just the information required to succeed. In the third column, the integration periods for each block are, from left to right, up to down, 3.9ms, 24.2ms, 21ms and 45.1ms respectively.

the required information to succeed. But at the same time, it filters out spurious details that, as demonstrated for the uniformly illuminated case, trigger false positives. This corroborates the idea stated in the introduction. When it comes to artificial vision, understood as the process of extracting relevant information from a visual stimulus, the design of smart image sensors should not be driven exclusively by the capture of high-quality images and the subsequent implementation of typical low-level processing. The characteristics of the underlying vision algorithm must be taken into account since the very beginning. A thorough analysis of such characteristics will enable the possibility of conveying to the sensing stage just those operations where the smart sensor can really become efficient.

#### IV. CONCLUSIONS

In artificial vision frameworks, smart image sensing must intimately be related to the vision algorithm to be implemented. In this paper, we propose reconfigurable low-power circuitry for block-wise HDR imaging. The objective is to provide vision algorithms with the flexibility to deal with a wide intra-frame range of illuminations according to their specific needs.

#### ACKNOWLEDGMENT

This work is funded by MEyC (Spain) through projects TEC2012-38921-C02-01, co-funded by the European Regional Development Fund, and IPT-2011-1625-430000, by the Office of Naval Research (USA) through grant N000141110312, and by CDTI (Spain),

co-funded by the European Regional Development Fund, through Project IPC-20111009.

#### REFERENCES

- [1] M. El-Desouki, O. Marinov, M. Deen, and Q. Fang, "CMOS active-pixel sensor with in-situ memory for ultrahigh-speed imaging," *IEEE Sensors J.*, vol. 11, no. 6, pp. 1375–1379, 2011.
- [2] J. Fernández-Berni, R. Carmona-Galán, and L. Carranza-González, "FLIP-Q: A QCIF resolution focal-plane array for low-power image processing," *IEEE J. Solid-State Circuits*, vol. 46, no. 3, pp. 669–680, 2011.
- [3] A. Lopich and P. Dudek, "A SIMD cellular processor array vision chip with asynchronous processing capabilities," *IEEE Trans. Circuits Syst. I*, vol. 58, no. 10, pp. 2420–2431, 2011.
- [4] B. Zhao, Z. Xiangyu, C. Shoushun, L. Kay-Soon, and Z. Hualiang, "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, 2012.
- [5] A. Spivak, A. Belenky, A. Fish, and O. Yadid-Pecht, "Wide-dynamic-range CMOS image sensors — comparative performance analysis," *IEEE Trans. Electron Devices*, vol. 56, no. 11, pp. 2446–2461, 2009.
- [6] P. Viola and M. Jones, "Robust real-time face detection," *Int. J. of Computer Vision*, vol. 57, no. 2, pp. 137–154, 2004.
- [7] S. Boltz, E. Debreuve, and M. Barlaud, "High-dimensional statistical measure for region-of-interest tracking," *IEEE Trans. Image Process.*, vol. 18, no. 6, pp. 1266–1283, 2009.
- [8] J. Chiverton, X. Xie, and M. Mirmehdi, "Automatic bootstrapping and tracking of object contours," *IEEE Trans. Image Process.*, vol. 21, no. 3, pp. 1231–1245, 2012.
- [9] V. Brajovic, R. Miyagawa, and T. Kanade, "Temporal photoreception for adaptive dynamic range image sensing and encoding," *Neural Networks*, vol. 11, no. 7-8, pp. 1149–1158, 1998.
- [10] J. Fernández-Berni, R. Carmona-Galán, G. Liñán Cembrano, A. Zarándy, and A. Rodríguez-Vázquez, "Wi-FLIP: A wireless smart camera based on a focal-plane low-power image processor," in *IEEE/ACM Int. Conf. on Distributed Smart Cameras*, 2011.