# A 148dB Focal-Plane Tone-Mapping QCIF Imager

S. Vargas-Sierra, G. Liñán-Cembrano, A. Rodríguez-Vázquez
Instituto de Microelectrónica de Sevilla (IMSE-CNM), CSIC and Universidad de Sevilla
Avda. Américo Vespucio s/n, 41092 Sevilla, Spain
Email: sonia@imse-cnm.csic.es, linan@imse-cnm.csic.es, angel@imse-cnm.csic.es

***Abstract*— This paper presents a QCIF HDR imager where visual information is simultaneously captured and adaptively compressed by an in-pixel tone-mapping scheme [1]. The tone mapping curve (TMC) is calculated from the histogram of an auxiliary previous image, which serves as a probability indicator of the distribution of illuminations within the current frame. The chip maps 148dB scenes onto 7-bit/pixel coding, containing illuminations from 2.2mlux (SNR10) to 55.33klux –with extreme values captured at 8s and 2.34μs, respectively. Pixels use an Nwell-Psubstrate photodiode and autozeroing for establishing the reset voltage. Measured sensitivity is 5.79$\frac{V}{lux \cdot s}$. Dark current effects in the final image are attenuated by an automatic programming of the DAC levels. The chip has been fabricated in the 0.35μm OPTO technology from AMS.**

## I. Introduction

High Dynamic Range (HDR) Imagers usually codify illuminations in the scene non-adaptively, using either long bit-words per pixel –e.g. mantissa exponent [2]– obtained from the combination of images captured at different exposures [3], or a fixed compressive function –e.g. logarithmic approach [4]– among many other possibilities [5]. These non-adaptive approaches usually lead to either high computational costs for the post-processing of the images –in the long bit-words case– or to the loss of details and lack of contrast –e.g. log. sensors– due to the fixed compression. In order to overcome these drawbacks, the proposed system produces an adaptive compression of illuminations using only 7-bit per pixel. Basically, the sensor operates as a Time-to-First-Spike imager [6] and implements the tone mapping compression over temporal information. Typical drawbacks in this type of imagers are non-linear signal compression and lower maximum SNR. Regarding the first, optimized non-linear compression is what it is actually intended in this design. Regarding the latter, it is due to the need of reducing the maximum voltage swing to allocate some range for the operation of the comparator. In our case, the implemented in-pixel ADC has low resolution (7-bit) and the achieved SNR is sufficient for this purpose. This is proven by the measurements, which indicate that the read noise floor is about 0.2 LSBs.

## II. Tone Mapping Algorithm for HDR Operation

The key point in the operation of our sensor is the combination of measuring the crossing time between a reference signal, $V_{ref}$, and the $V_{ph}(I_{pix}, t)$ voltage –for pixels working in the photocurrent integration mode– and ramping up very fast the analog reference at the end of the exposure to allow for poorly illuminated pixels to intersect $V_{ref}$. In either the case, the crossing event makes the pixel to get its value from the current status of a 7-bit globally distributed signal TMC<6:0> whose non-linear temporal evolution is calculated from a tone-mapping algorithm [1][7]. This calculation employs an auxiliar image, named Time Stamp Image (TSI), which is also generated on chip. TSI information is provided by one out of every four pixels in a 2×2 neighborhood –see section III, and acts as an indicator of the distribution of illuminations in the scene [1]. The generation of the TSI follows the same principle as that of the Tone-Mapped Image (TMI) –see an example for a reduced number of bits in Fig. 1. During photocurrent integration, when the pixel voltage $V_{ph}(I_{pix}, t)$ intersects $V_{ref}$, the pixel samples both the status of the 4-bit Time Stamp bus TSC<3:0> and the value of the Tone Mapping bus TMC<6:0>. Exposition time is divided into 16 windows (non-linearly distributed), with TS value just codifying the window number. The duration of temporal windows has been selected so that they are compressed towards the higher illumination bands, mimicking the natural ($1/I_{pix}$) compression of the intersection time (1), and optimized using the distribution of luminances in public HDR image data bases [7]. In the last temporal window, which can be as short as 153.6μs, $V_{ref}$ ramps up from its previous value $V_{bot}$, to a programmable value $V_{top}$ in 128 steps. Pixels crossing $V_{ref}$ during this window store TSC=0.

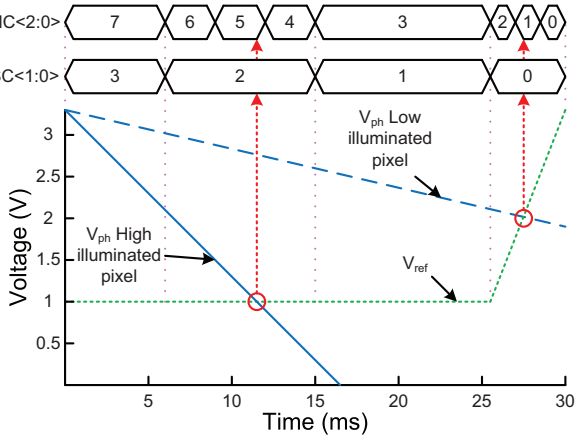



Fig. 1: Signals Involved in HDR Image Acquisition

$$T_{cross}(I_{pix}, V_{ref}) = \frac{C_{pix}}{I_{pix}}(V_{rst} - V_{ref}) \quad (1)$$

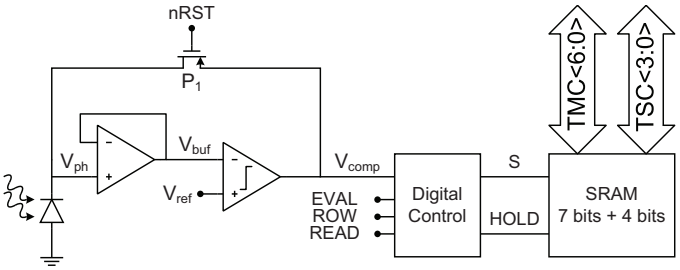
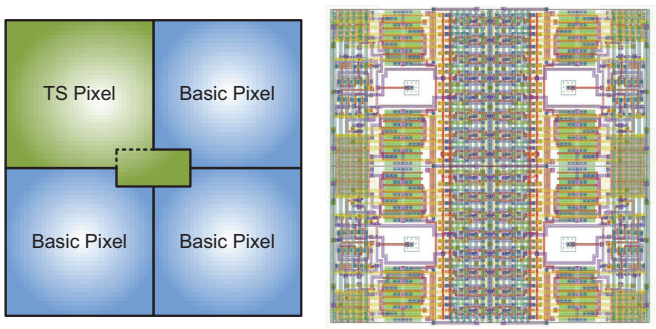



Fig. 2: TS Pixel Block Diagram




(a) Pixel Group Organization. (b) Pixel Group Layout.

Fig. 3: Pixels Group.

The temporal evolution of TMC<6:0> for the current frame is created from the histogram of the TSI in the previous frame. Thus, we consider the TS information as an indicator of probability, and so it may fail when the exposition time is too long as compared to the rate of changes in the image. TMC<6:0> varies linearly within each temporal window, spanning over a number of LSBs which is a function of the relevance of this window in the histogram of TSI. For instance, if this histogram shows that half of the pixels crossed $V_{ref}$ during temporal window with TSC=3, the TMC<6:0> curve could span over 64 codes during this temporal window. Finally, since the duration of temporal windows is non-linearly distributed in time, the obtained profile for the TM curve is piece-wise linear in time. More details about the generation of these curves and the duration of temporal windows is provided in [1].

## III. Pixels

Pixels have been arranged in two categories:

- TS Pixels: including both TMC and TSC circuitry.
- Basic Pixels: including only TMC circuitry.

A block level schematic of a TS pixel is shown in Fig. 2. The sensor, a $3\times3\mu m^2$ Nwell/Psubs diode[1], works in photocurrent integration mode. It uses auto-zeroing to establish the reset voltage through the combined action of a buffer, which in operation isolates the photodiode capacitor from comparator's kickback noise, an analog comparator (where $V_{ref}$=$V_{rst}$ during reset phase) and a PMOS feedback switch $P_1$. Additionally, digital circuitry is included to control R/W operations of the SRAM cells. Signal ROW controls the external write of data row by row –for evaluation and initialization purposes, signal EVAL activates internal write operation and signal READ enables external readouts synchronized with the ROW signal. TS pixels contain 7(TMC)+4(TSC)=11 bits of SRAM, whereas Basic Pixels (BP) do only include 7(TMC) SRAM modules.

Pixels are physically arranged as shown in Fig. 3(a). Notice that each TS pixel takes some area from its 3 BP neighbors to allocate the 4 SRAM modules for TSC storage. TSC SRAMs are grouped in the middle of the 2×2 arrangement –as shown in Fig. 3(a)– and controlled by signals produced in the TS pixel only. The layout of a group of 2×2 pixels is shown in Fig. 3(b). Observe that SRAM modules are grouped in the central vertical region, sharing global control, digital power and ground lines. This increases the attainable pitch and reduces the noise coupling from digital switching in the analog blocks.

[1]Aperture in metal structures over the diode is $9.75\times7.30\mu m^2$. Due to this, carriers created within this area can also contribute to the photogenerated current by reaching the photodiode through diffusion, increasing the effective fill-factor.

### A. Auto-zeroing Technique

Cornerstone in the operation of the imager is the auto-zeroing technique to cancel out most offset contributions from the two amplifiers in the pixel. During the reset phase, the voltage $V_{rst}$ is applied to the $V_{ref}$ input in Fig. 2, and transmitted to the photodiode's integrating capacitor through the negative feedback loop. If we consider that amplifiers can be efficiently modeled to this purpose by their input-referred offset voltage $V_{Ox}$ and a finite DC gain $A_x$ –where $x = B$ for the Buffer and $C$ for the Comparator, one finds after simple calculations, which include Taylor's series expansion and neglecting second order error terms, that the reset value is approximately established to:

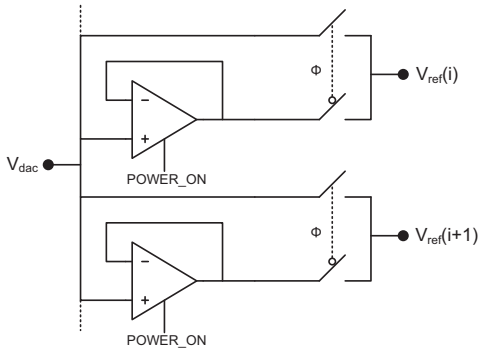
$$V_{ph_{rst}} \cong (1+\epsilon_C)^{-1}\cdot[(1+\epsilon_B)\cdot(V_{rst}+V_{OC})-V_{OB}] \quad (2)$$

where $\epsilon_C = 1/A_C$, and $\epsilon_B = 1/A_B$.
Thus, the effective differential voltage applied at comparator's input during operation –including the feedthrough contribution $V_{FT}$ introduced by the reset switch– is:

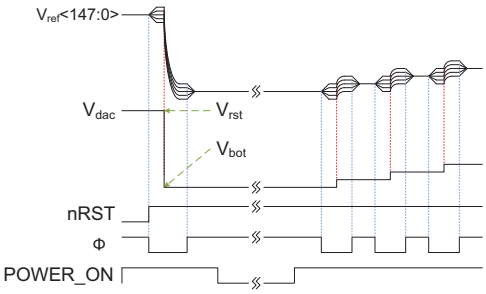
$$V_{eff} \approx V_{ref} - (V_{rst} - \frac{I_{pix}}{C_{pix}}\Delta t) - (1-\epsilon_B)V_{FT} + \epsilon \quad (3)$$

$$\epsilon = \epsilon_C V_{rst} - \epsilon_B \frac{I_{pix}}{C_{pix}}\Delta t + \epsilon_C(1-\epsilon_C)[V_{OC} - (1-\epsilon_B)V_{OB}] \quad (4)$$

Clearly, most of errors –except the feedthrough, which is the main error contribution at the end– vanish as the amplifiers gain is sufficiently high. This, in practice, is translated into a small residual contribution due to the impossibility of designing very large gain low-power amplifiers (each amplifier consumes 50nA) within such small area. Pixel design has been made under the 3 sigma constraint for all added non-idealities.




(a) $V_{ref}$ Distribution Circuitry




(b) $V_{ref}$ Distribution Signals

Fig. 4: $V_{ref}$ Distribution Scheme

## IV. Chip-Level Additional Functionalities

### A. Analog Reference Generation

A dynamic biasing mechanism has been developed in order to transmit the $V_{ref}$ signal to the array. As shown in Fig. 4(b), $V_{ref}$ drops very quickly from $V_{rst}$ to its value during most of the exposure $V_{bot}$, and, in the last window, moves from $V_{bot}$ to $V_{top}$ in 128 steps to perform a kind of single slope A-to-D conversion of the pixels not crossing $V_{ref}$ previously. Every row is provided with an analog buffer that receives $V_{ref}$, from an on-chip DAC, and drives all the nodes in its row. Clearly, there will be slight differences in the final voltage reached by each row due to offset, and other non-idealities. The next step is to switch-off the amplifiers and short-circuit all $V_{ref_i}$ nodes –Fig. 4(a)– to the DAC's output. This forces all nodes to reach the same final voltage in a shorter time than only using one driver at the output of the DAC (due to RC effects in wires driving the signal to the different points in the array) [8].

### B. Dark Signal Contribution Attenuation

Dark current effects are specially noticeable in dark pixels, that may look very noisy in long-exposure shots. In order to attenuate the visual degradation produced by this undesired contribution, we have experimentally measured average dark signal contribution $\overline{I_{DC}}$ and standard deviation $\sigma(I_{DC})$ for different exposition times and operating temperatures (from an on-chip temperature sensor). The visual effects of dark current in pixels crossing $V_{ref}$ during the last temporal window can be attenuated by automatically adapting the voltage levels of the ADC. Pixels with a photogenerated current smaller than

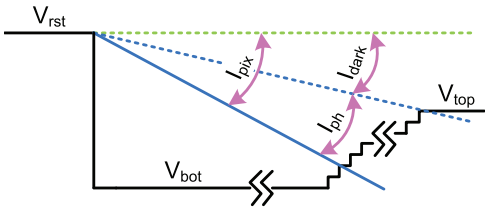
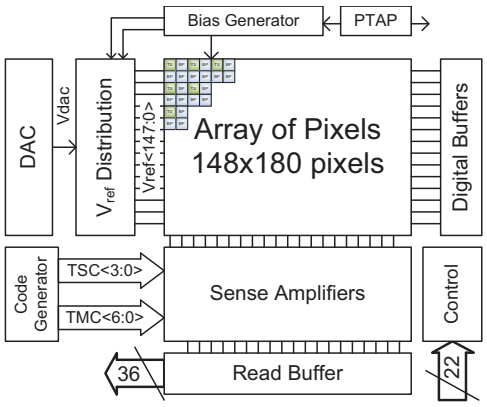



Fig. 5: Dark Signal Contribution Mitigation Scheme

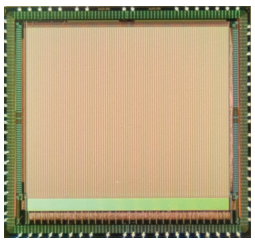
$I_{dark} = \overline{I_{DC}} + 3\sigma(I_{DC})$ are out of the ADC operating range and hence assigned to the value 0 –see Fig. 5.

## V. Chip Architecture

The architecture of the chip is shown in Fig. 6(a), with its core array of QCIF resolution (+ 2 dummy rows and columns on each side). Pixels functionality is supported by additional periphery blocks. An 8-bit DAC generates $V_{ref}(t)$ and 148 buffers (one per row) enhance the dynamics of its distribution to the array. Digital control signals also employ per-row distributed digital buffers (including clock-tree generation). TSC<3:0> and TMC<6:0> are generated by a Code Generator in gray format in order to reduce switching at the pixel level to only one SRAM module at a time (instead of 7) in Basic Pixels and 2 (instead of 11) in TS Pixels. Read and write operations from the array are accomplished by a bank of sense amplifiers. Image is retrieved row by row and stored in a read buffer (1 row), which outputs images through a high-speed 36-bit bus (4 TMC codes + 2 TSC codes at a time –equivalent to 43MBytes/s). Fig. 6(b) shows a microscope capture.

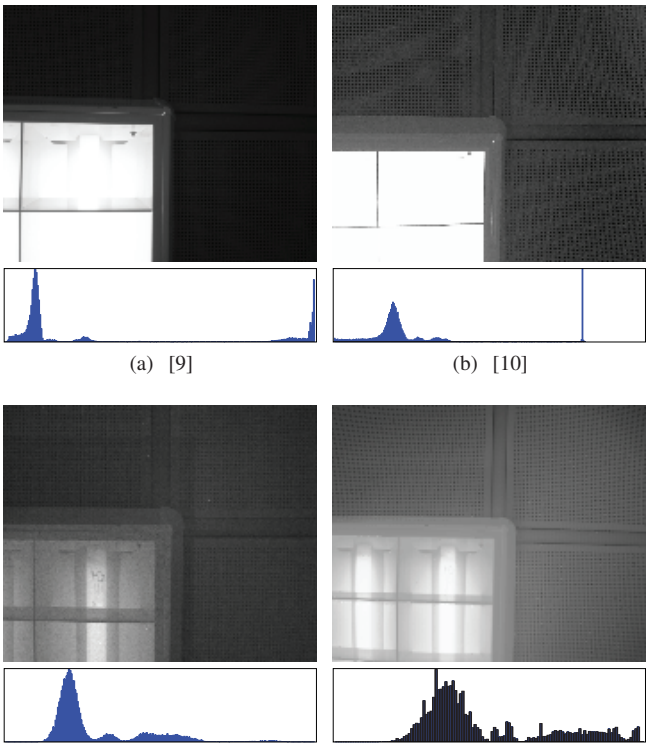



(a) Chip Block Diagram



(b) Microscope Capture

Fig. 6: Chip Block Diagram and Microscope Capture



(a) [9] (b) [10]

(c) [11] (d) This work

Fig. 7: Comparison with commercial cameras.

## VI. Experimental Results

We present here a comparison of images captured from 3 commercial systems and our chip (see Fig. 7). The Sony Cybershot DSC-W80 [9] –which includes an enhanced sensitivity CCD sensor ($Super\,HAD^{TM}\,CCD$), the iPhone4 camera –which allows HDR Mode [10] (since iOS 4.2) by using a combination of 3 pictures, and the Photonfocus MV-D752E-40-U2-12 [11], which employs the Lin-Log technology. Noticeably, despite using only half of the codes (128 vs. 256) for image representation, our approach produces an image which is –visually– competitive with the other approaches. The Lin-Log sensor shows little more details within the fluorescent lamp area at the expense of a higher noise. The DSC-W80 is less noisy but it shows both over and under exposed areas. Finally, the HDR mode in the iPhone4 shows some similar performance in the darker areas, but fails to produce details in the brighter ones. Table I summarizes the most important electro-optical characteristics of the chip.

## VII. Conclusions and Future Work

We have presented a 148dB (SNR10) imager that automatically adapts to compress the HDR scene in a 7-bit format by a Tone-Mapping algorithm using information from the previous frame. Pixels include auto-zeroing and SRAM storage which allows for long exposure shots. A dark signal contribution mitigation scheme has been implemented to enhance the visual quality in dark areas. Global analog reference to the pixels is dynamically distributed to allow for low-power, fast, and precise operation. We are currently working in a megapixel resolution imager using a 130nm 3D (vertical integration) technology with pitch estimations below 7µm and fill factor near 100% using a BSI approach.

TABLE I: Chip Characteristics

| Characteristic | Value |
|---|---|
| Technology | 3.3V 0.35µm 2P4M AMS OPTO |
| Image Size | QCIF + dummies |
| Pixels Size | 33×33µm$^2$ |
| Photodiode | 3×3µm$^2$ NW-Psub |
| Metals Aperture | 9.75×7.3µm$^2$ |
| Fill Factor | 0.8%(Photodiode), 6.5%(Aperture) |
| Read Noise | 25e$^-$(0.19DN) |
| PRNU | 1.8% |
| Conversion Factor | 128.7(e$^-$/DN) |
| Full Well Capacity | 12.2ke$^-$ |
| Exposition Time | 2.34µs to 8s |
| Image coding | 7 bits |
| Sensitivity | 5.79V·lux$^{-1}$·s$^{-1}$ |
| Average Dark Signal | 10.8mV·s$^{-1}$ |
| Dynamic Range (SNR10) | 148dB, 2.2mlux (SNR10) - 55.3klux |
| Fastest Image Download Time | 666µs |
| Maximum Power Consumption | 111.2mW@1205fps |


## Acknowledgment

This work is partially funded by TEC2009-11812, CENIT ADAPTA, ONR Grant N000141110312, FEDER 2007-2013, WiVisNet and IMPACTO. The authors also thank the useful comments from the reviewers.


## References


[1] S. Vargas-Sierra, G. Liñán Cembrano, E. Roca, and A. Rodríguez-Vázquez, "High-dynamic range tone-mapping algorithm for focal plane processors," vol. 8068, no. 1. SPIE, 2011, p. 806807.

[2] A. Belenky, A. Fish, A. Spivak, and O. Yadid-Pecht, "Global shutter cmos image sensor with wide dynamic range," *Circuits and Systems II: Express Briefs, IEEE Transactions on*, vol. 54, no. 12, pp. 1032 –1036, dec. 2007.

[3] M. Mase, S. Kawahito, M. Sasaki, Y. Wakamori, and M. Furuta, "A wide dynamic range cmos image sensor with multiple exposure-time signal outputs and 12-bit column-parallel cyclic a/d converters," *Solid-State Circuits, IEEE Journal of*, vol. 40, no. 12, pp. 2787 – 2795, dec. 2005.

[4] H.-Y. Cheng, B. Choubey, and S. Collins, "An integrating wide dynamic-range image sensor with a logarithmic response," *Electron Devices, IEEE Transactions on*, vol. 56, no. 11, pp. 2423 –2428, nov. 2009.

[5] A. Spivak, A. Belenky, A. Fish, and O. Yadid-Pecht, "Wide-dynamic-range cmos image sensors - comparative performance analysis," *Electron Devices, IEEE Transactions on*, vol. 56, no. 11, pp. 2446 –2461, nov. 2009.

[6] X. Guo, X. Qi, and J. Harris, "A time-to-first-spike cmos image sensor," *Sensors Journal, IEEE*, vol. 7, no. 8, pp. 1165 –1175, aug. 2007.

[7] E. Reinhard, G. Ward, S. Pattanaik, and P. Debevec, *High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting*. Elsevier / Morgan Kaufmann, 2006.

[8] S. Vargas-Sierra, "Proposal of architecture and circuits for dynamic range enhancement of vision systems on chip designed in deep sub-micron technologies," Ph.D. dissertation, University of Seville, 2012.

[9] Sony Cybershot DSC-W80 specifications: http://news.sel.sony.com/assets/Cyber-shot_2007/specs/DSC-W80.pdf.

[10] Apple iPhone 4 User Guide: http://manuals.info.apple.com/en_US/iphone_user_guide.pdf.

[11] Photonfocus Products: http://www.photonfocus.com/html/eng/products/products.php?prodId=55.