P28

A QCIF 145dB Imager For Focal Plane Processor Chips Using a Tone Mapping Technique in Standard 0.35μm CMOS Technology

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 means of an in-pixel tone mapping scheme [1]. The tone mapping curve (TMC) is calculated from a non-linear histogram of the previous image, which serves as a probability indicator of the distribution of illuminations within the present frame. The chip produces 7-bit/pixel images that can map illuminations from 311×10^{-6}lux to 5875 lux in a single frame in a way that each pixel decides when to stop observing photocurrent integration –with extreme values captured at 8s and 20μs respectively. Pixels use a 3x3μm^2 Nwell-Psubstrate photodiode and an autozeroing technique for establishing the reset voltage, which cancels most of the offset contributions created by the analog processing circuitry. 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 top voltage. The chip has been designed 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, non-exclusively, to either high computational costs for the post-processing of the images –in the long bit-words case– or the loss of details and lack of contrast –e.g. log. sensors– due to the fixed compression. In order to avoid these drawbacks, the proposed system produces an adaptive compression of illuminations using only 7-bit per pixel.

II. Tone Mapping Algorithm for HDR Operation

The key point in the operation of the imager 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], [6]. This calculation employs an auxiliar image, named Time Stamp image (TS), which is also generated on chip. TS image information is provided by one out of every four pixels in a 2×2 neighborhood –see Fig. 3(a), and acts as an indicator of the distribution of illuminations in the scene [1]. The generation of the TS image follows the same principle as that of the Tone-Mapped image, as shown in Fig. 1. During photocurrent integration, when the pixel voltage V_{ph} 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]. The exposure time is divided into 16 non-linearly distributed windows, each of them having a different TS value. The duration of every temporal window has been selected so that they are compressed towards the higher illumination bands, mimicking the natural compression ($1/I_{pix}$) of the intersection time expressed in (1), and optimized taken into account the distribution of luminances in public HDR image data bases[1]. In the last temporal window, which lasts for 1.28ms, the analog reference signal evolves linearly from its previous constant value ($V_{bot} = 1V$), to a programmable value V_{top} in 128 steps. Pixels crossing V_{ref} during this interval sample the value 0 in their time stamp information.

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

[1] http://www.cs.ucf.edu/~reinhard/cdrom/hdr.html.

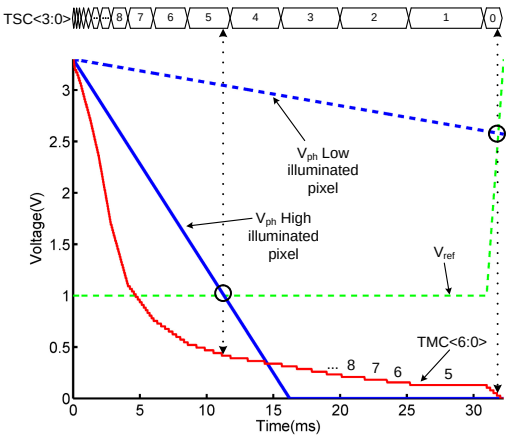


Fig. 1: Signals Involved in HDR Image Acquisition.

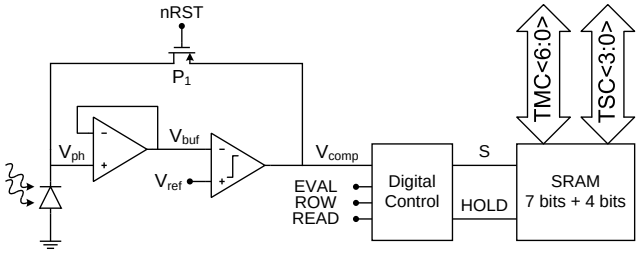


Fig. 2: TS Pixel Block Diagram.

The temporal evolution of TMC[6:0] is created from the histogram of the TS image in the previous frame (this is why we consider the information in TS as an indicator of probability rather than as an exact evaluation, and so it may fail when the exposure 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 proportional[2] to the weight of this temporal window in the histogram of the TS image. Just for illustration purposes, if this histogram shows that half of the pixels crossed V_{ref} during temporal window #3, the TMC[6:0] curve will 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 non-linear in time as well (or piece-wise linear to be more precise). More details about the generation of these curves and the duration of temporal windows is provided in [1].

III. Pixels

As mentioned previously, pixels have been arranged in two categories:

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

A block level schematic of a TS pixel is shown in Fig. 2. The sensor, a $3 \times 3\mu m^2$ Nwell/Psubs diode[3], works in photocurrent integration mode. It uses an auto-zeroing technique 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 read and write 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 –which are obviously 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.

[2]This is the most straightforward way of assigning codes to temporal windows. In our system implementation we can choose among different options including minimum threshold to activate a temporal window, logarithmic distribution of codes depending on the histogram, enhancement of different bands (priority to dark or priority to bright), etc.

[3]Aperture in metal structures over the diode is $9.75 \times 7.30\mu m^2$. 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.

Pixels are physically arranged as shown in Fig. 3(a). Notice that each TS pixel takes, conceptually, some area from its 3 BP neighbors which is used to allocate the 4 SRAM modules for TSC storage. Indeed, all pixels have 8(7TMC+1TSC) SRAM blocks. TSC modules 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 we have grouped the SRAM modules in the central vertical region, sharing global control, digital power and ground lines. This increases the attainable pitch and reduces the noise from digital switching in the analog blocks.

A. Auto-zeroing Technique

A crucial issue in the operation of the imager is the usage of an auto-zeroing technique to cancel out most offset contributions from the two amplifiers in the pixel. During 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 created by the two amplifiers and the reset switch. 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:

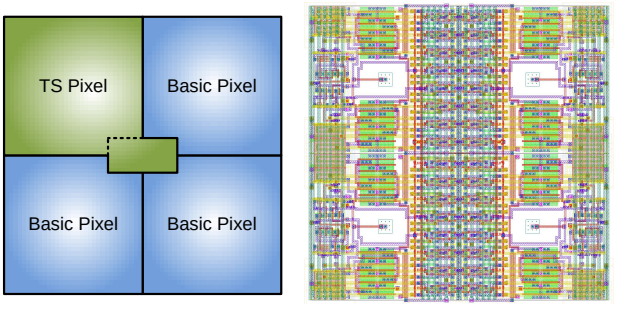
$$V_{ph_{rst}} \cong (1+\epsilon_C)^{-1} \cdot [(1+\epsilon_B) \cdot (V_{rst} + V_{OC}) - V_{OB}] \quad (2)$$

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:

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

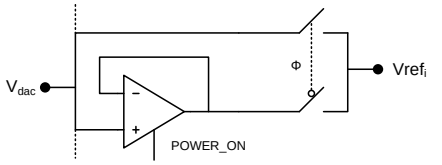
$$\epsilon = \epsilon_C V_{rst} - \epsilon_B \frac{I_{pix}}{C_{pix}}\Delta t + \epsilon_C(1-\epsilon_C)[V_{OC} - (1-\epsilon_B)V_{OB}] \quad (4)$$

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. The pixel design has been made under the 3 sigma constraint for all added non-idealities.

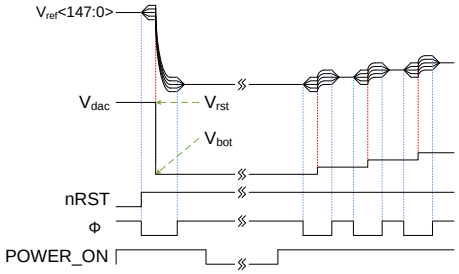


(a) Pixel Group Organization. (b) Pixel Group Layout.

Fig. 3: Pixels Group.



(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 (in Fig. 4) has been developed in order to transmit the V_{ref} signal to the array. As shown in Fig. 4(b), V_{ref} must drop very quickly from V_{rst} to its constant value during most of the exposure ($V_{bot} = 1V$), and, in the last window, move from $V_{bot} = 1V$ to V_{top} in 128 steps to perform a kind of single ramp A-to-D conversion of 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 corresponding 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).

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 exposure times and operating temperatures. These measurements allow us to diminish the visual effect of dark current in pixels crossing V_{ref} during the last temporal window simply by lowering V_{top} as shown in Fig. 5, where $I_{dark} = \overline{I_{DC}} + 3\sigma(I_{DC})$. It is worth mentioning that the optimum V_{top} level is automatically generated by the FPGA controlling the chip using exposure time, DC measurements and the input from an on-chip PTAP sensor.

V. Chip Architecture

The architecture of the chip is shown in Fig. 6(a), with its core array of 148 × 180 pixels (QCIF + 2 dummy rows and columns on each side). Pixels functionality is supported by additional periphery blocks. An 8-bit DAC generates the reset voltage V_{rst} during reset, the fixed voltage V_{bot} during the exposition time and finally the 128 levels ramp signal from V_{bot} to V_{top} during the last temporal window. 148 buffers (one per row) enhance the dynamics of distributing V_{ref} 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. This coding reduces switching at the pixel level to only one SRAM module at a time (instead of 7) for Basic Pixels and 2 (instead of 11) for 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 –eq. to 43MBytes/s). Fig. 6(b) shows a microscope's capture.

VI. Experimental Results

Due to paper size limits, we only present here[4] a comparison of images captured from 3 commercial systems and our chip (see Fig. 7). The Sony Cybershot DSC-W80 [7] –which includes an enhanced sensitivity CCD sensor ($Super\ HAD^{TM}\ CCD$), the Iphone4 camera –which allows HDR Mode [8] (since iOS 4.2) by using a combination of 3 pictures, and the Photonfocus MV-D752E-40-U2-12 [9], 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. Fig. 8 shows the results for the cumulative application of four Sobel filters (-45, 0, 45, 90) –using matlab– over the images in Fig. 7 in order to illustrate how information is kept when capturing a HDR scene. Clearly, images in Fig. 8(a), (b), exhibit some lack of details near –or within– the bulb, whereas Fig. 8(c) losses the information in the low contrast printed characters. The results from our chip show little lower details in the bulb area than those in (c) but much better performance in the low-contrast areas. Besides our sensor also exhibit lower noise than those in Fig. 8(b), and (c), being only slightly outperformed –in terms of noise– by the results in Fig. 8(a). Table I summarizes the most important characteristics of the chip.

VII. Conclusions and Future Work

We have presented an imager that automatically adapts to compress the HDR scene in a 7-bit format using a Tone-Mapping algorithm with information from the previous frame. Pixels include auto-zeroing and in-pixel SRAM storage which allows for long exposure shots.

[4]More results will be presented at the conference.

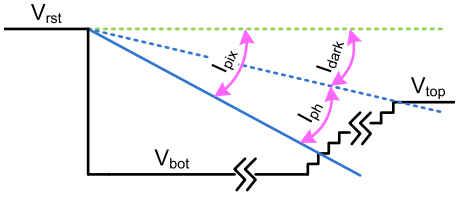
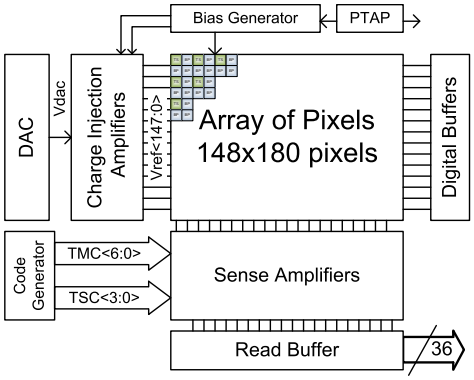
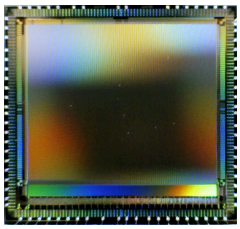


Fig. 5: Dark Signal Contribution Mitigation Scheme.



(a) Chip Block Diagram



(b) Microscope Capture

Fig. 6: Chip Block Diagram and Microscope Capture.

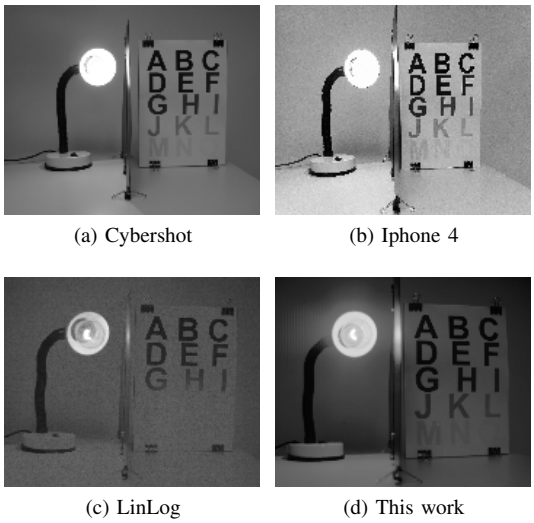
An automatic dark signal contribution mitigation scheme has been implemented to enhance the visual quality in dark areas. Global analog reference to the pixels is dynamically distributed to allow for low-power, fast, and precise operation. We are currently working in a megapixel resolution imager using a 130nm 3D technology with pitch estimations below 7µm and fill factor near 100% using BSI approach.

Acknowledgment

This work is partially funded by TEC2009-11812, CENIT ADAPTA, ONR Grant N000141110312, and FEDER 2007-2013.

References

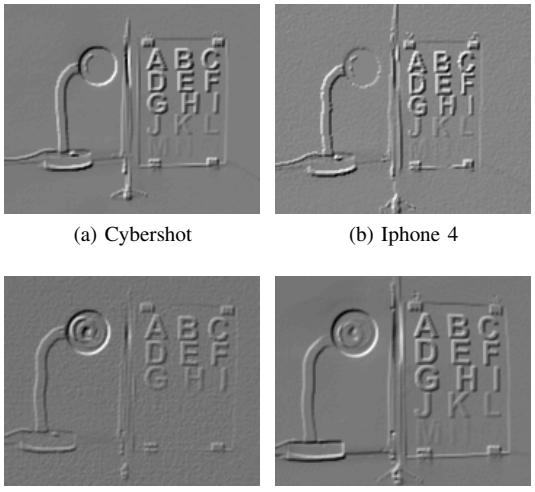
[1] S. Vargas-Sierra, G. Liñán Cembrano, and A. Rodríguez-Vázquez, "High-dynamic range tone-mapping algorithm for focal plane processors," in *SPIE Microtechnologies, 2011.*, April 2011.



(a) Cybershot (b) Iphone 4

(c) LinLog (d) This work

Fig. 7: Comparison with commercial cameras.



(a) Cybershot (b) Iphone 4

(c) LinLog (d) This work

Fig. 8: Processed Images Comparison.

TABLE I: Chip Characteristics

Characteristic	Value
Technology	3.3V 0.35µm 2P4M AMS OPTO
Image Size	180(H)x148(V) (QCIF+dummies)
Pitch	33µm
Photodiode	NW-Psub, Aperture 9.75×7.3µm^2
Fill Factor	0.8% (Diode) 6.5% (Aperture)
Full Well Capacity	267 ke$^-$
Exposure Time	20µs to 8s
Image coding	7 bits
Sensitivity	5.79V·lux^{-1}·s^{-1}
Average Dark Signal	10.8mV·s^{-1}
Maximum Dynamic Range	145dB (from 311µlx to 5875lx)
Fastest Image Download Time	666µs
Fastest Operation Power Consumption	562mW@511fps

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

[7] [Online]. Available: http://news.sel.sony.com/assets/Cybershot_2007/specs/DSC-W80.pdf

[8] [Online]. Available: http://manuals.info.apple.com/en_US/iphone_user_guide.pdf

[9] [Online]. Available: http://www.photonfocus.com/html/eng/products/products.php?prodId=55