A 5.2Mpixel@250Fps 16Gbps CMOS Image Sensor with Embedded Digital Processor for Reconfigurability and on-chip Image Correction

Francisco Jiménez-Garrido [1], *José Fernández-Pérez* [1], *Cayetana Utrera* [1], *José Ma. Muñoz* [1], *Ma. Dolores Pardo* [1], *Rafael Domínguez-Castro* [1], *Fernando Medeiro 1 and Angel Rodríguez-Vázquez* [2,1]

[1] Universidad de Sevilla, Instituto de Microelectrónica de Sevilla (IMSE-CNM)

C/ Americo Vespucio (esquina Leonardo da Vinci), Parque Tecnológico Isla de la Cartuja, 41092-Sevilla (SPAIN)

[2] Consejo Superior de Investigaciones Científicas (CSIC), Instituto de Microelectrónica de Sevilla (IMSE-CNM)

[3] Innovaciones Microelectrónicas S.L. (ANAFOCUS)

Pabellón de Italia-Planta Ático, Parque Tecnológico Isla de la Cartuja, 41092-Sevilla (SPAIN)

arodri-vazquez@us.es, angel@imse-cnm.csic.es

Abstract – High-speed imagers are required for industrial applications, traffic monitoring, robotics and unmanned vehicles, movie-making, etc. Many of these applications call also for large spatial resolution, high sensitivity and the ability to detect images with large intra-frame dynamic range. This paper reports a CIS intelligent digital image sensor with 5.2Mpixels which delivers 12-bit fully-corrected images at 250Fps. The new sensor embeds on-chip digital processing circuitry for a large variety of functions including: windowing; pixel binning; sub-sampling; combined windowing-binning-sub-sampling modes; fixed-pattern noise correction; fine gain and offset control; color processing, etc. These and other CIS functions are programmable through a simple four-wire serial port interface.

I. Introduction

CMOS Image Sensor (CIS) chips and architectures for high-speed and high-resolution are becoming common [1]-[6]. The sensor presented in this paper adds the features of re-configuration and intelligence to those of high-speed and resolution. Thus, for instance, it can be re-configured to implement different binning modes and to download VGA images at 2,030Fps. It can also be programmed to deliver 12bit images either in linear acquisition mode (with 60dB intra-frame DR) or in compressed high dynamic range mode (with 97dB intra-frame DR). Besides supporting re-configuration, on-chip intelligence is used for on-chip image correction (such as suppression of vertical and horizontal fixed-pattern noise) thus allowing fully corrected digital images be delivered at the sensor outputs and largely simplifying the implementation of camera systems.

II. Sensor Architecture and Readout Channel

Fig.1(a) shows the sensor architecture. It consists of an array of 2,560 x 2,048 active global-shutter 5-T, 0.18μm CMOS, 5μm-pitch pixels surrounded by optical black columns and rows employed for image correction. The readout channel incorporated into the sensor combines per-colour (R, G1, G2, B) configurable analog gain and offset, together with 12-bit ADC and fine digital gain and offset adjustment, to deliver best image quality pictures under all illumination conditions.

The exposure and read-out timings are generated by a control-unit that enables a wide variety of operational modes. Trigger management is especially designed for those industrial applications demanding ultra-low trigger-to-exposure latency. Re-configuration and processing functions embedded on-chip include windowing; pixel binning; sub-sampling; combined windowing-binning-sub-sampling modes, fixed-pattern noise (FPN) calibration and correction and defective pixels correction. All these functions can be accessed through a four-wire serial port interface (SPI). Alternatively, the sensor can be fully controlled through command-interpreter software that runs on the on-chip, 12-bit on-chip microprocessor.

The sensor includes 24 LVDS high-speed data ports that transfer 12, 10, or 8-bit images at a speed up to 16.59Gbit/sec. Two additional LVDS channels are provided for clock recovery and synchronization. The number of active LVDS ports can be reduced when either the frame rate or the output word-length is reduced, thus minimizing the complexity of the external components required. Power dissipation scales accordingly. The whole system runs with an external clock of 9.6MHz. All the required timing and reference voltages are internally generated, thus minimizing the need for external components. The sensor includes a power down capability for very low power dissipation.

Fig.1(b) shows the readout channel of the sensor – responsible for the acquisition, analog conditioning, and digitization of the image data from the pixel array. The resulting digital words are passed to the serialization block to be distributed to the embedded processors. Reading follows a column-parallel approach with 1 readout channel per sensor column. Each individual channel (dashed red line in Fig.1(b) consists of: i) 4 analog memories separated in two sets (with 2 analog memories each) for pipelining input data; ii) 1 double sampling circuit for Correlated Double Sampling (CDS); iii) 1 Programmable Gain Amplifier (PGA), and iv) 1 Analog-to-Digital Converter (ADC).

Fig.1(c) shows the data flow through the analog and digital paths. Note that the read-out channel incorporates also a Digital-to-Analog Converter (DAC) for addition/subtraction of an offset to/from the readout channel input.At the analog side, the purpose of this offset is to make the input signal fit into the range of the ADC. However, it is not a fine offset adjustment,

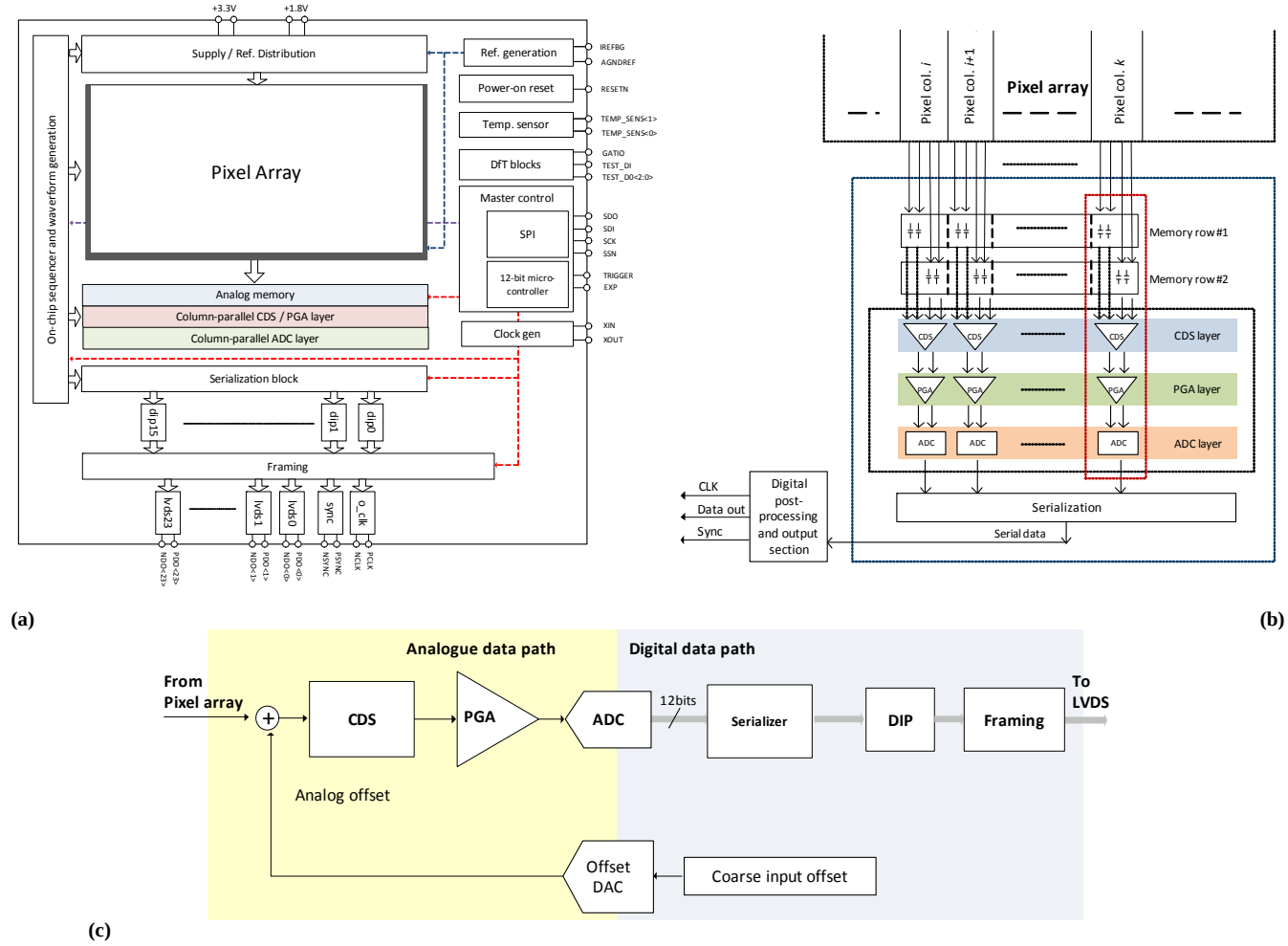

Figure 1. (a) CIS architecture; (b) Readout channel architecture; (c) Data flow through the analog and digital paths. Note that the read-out channel incorporates also a Digital-to-Analog Converter (DAC) for addition/subtraction of an offset to/from the readout channel input.

but a coarse one, as necessary to avoid loss of codes (due to having negative values at the input of the ADC) or significant loss of DR (in case a large dark signal from pixel exhausts a significant portion of the ADC full-scale range). Once the coarse analog input offset has been added, the signal is passed to the CDS block where the difference between the pixel reset and signal level is computed, and then to the PGA where it is amplified before converting it to digital at the ADC. The digitized data is serialized out to a set of Digital Image Processors (DIPs) where additional arithmetic operations, as well as data correction are carried out. The digitally processed data are passed to a Framing Block and outputted through either LVDS or CMOS ports.

During reading, data in line memory stays for two row times (row time is defined as the time required for a row of pixels readout) thus indicating that during one conversion time it is read from the pixel and, during the following one, data is fed to the rest of the signal path. This dual analog memory (or "ping-pong") structure in the readout channel permits increasing the available time to read out a row; this significantly decreases the speed and the power consumption requirement of the readout channels.

III. On Chip Intelligence

Fig.2 highlights the contents of the DIP within the analog and digital data paths. The DIPS perform the image correction and processing functions embedded on-chip.

A. On-chip image processing functions

Analog offset and analog gain. The sensor allows addition/subtraction of a coarse analog offset to/from the pixel output. The analog gain applied after CDS is selectable among x1, x2, x4, x8, and x16. In fact, a single readout channel can work with two different analog gains so that these are exchanged in alternating pixel rows. This makes it possible to configure analog gain per colour in devices with Colour Filter Array (CFA). Changing analog gain implies modification of the waveforms controlling the readout blocks. Thus, there is no register for directly specifying the required analog gain, but a set of registers configuring such waveforms.

Digital offset and digital fain. After digitization, global digital

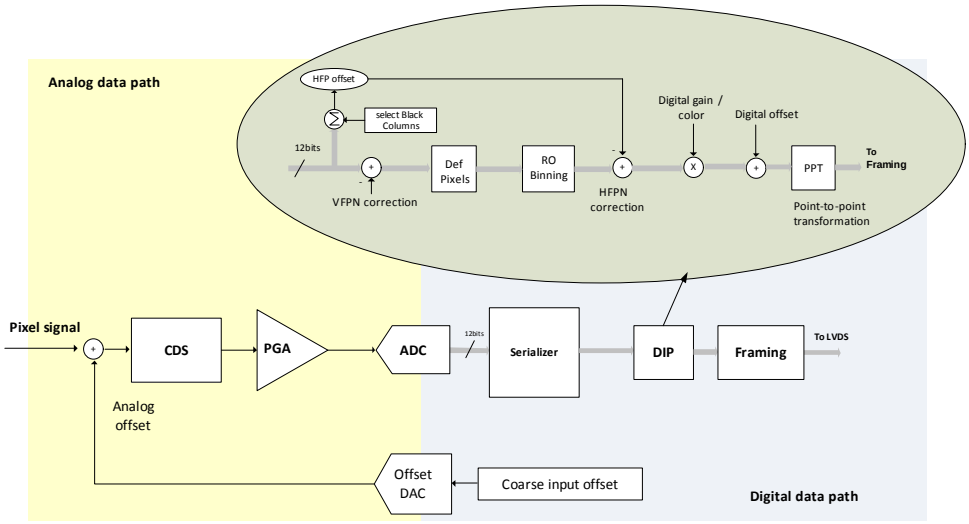

Figure 2. Data path scheme showing the location and contents of the DIP

offset and gain can also be adjusted through registers programmed by SPI. For colour sensing, global gain can be defined separately for green, red, and blue pixels, so that it can be used to perform on-chip white balance. For monochrome devices, the same gain is applied to all the pixels. Global offset has a range of ±2048. Global gain can be expressed as:

$$G_{GLOB} = 2^{G_{COARSE}} \times G_{FINE} \quad (1)$$

where G_{COARSE} = 0, 1, 2 .. 7 and G_{FINE} goes from 0 to 2.

Auto-black level control. Black level can be self-adjusted on-chip by using the OB pixels present in all sensor rows. The black level of a row is measured and subtracted from all pixels in that row before data is outputted.

Point-to-point transformation The Point-to-Point Transformation (PPT) block allows the user to program a pixel value-based transformation law. The sensor DIP implement a piece-wise linear approximation of the desired PPT, based on a Look-Up-Table (LUT). The on-chip PPT LUT can store 33 points defining 32 intervals in the overall response curve. The PPT procedure implements the following interpolation:

$$y = \frac{Y_{j+1} - Y_j}{X_{j+1} - X_j} \cdot (x - X_j) + Y_j$$

$$y = \frac{Y_{j+1} - Y_j}{X_{j+1} - X_j} \cdot (x - X_j) + Y_j \quad (2)$$

where x represents the incoming pixel code and X_j, Y_j are the coordinates of the interval containing such a code. X_j values are prefixed to 0, 128, 256, … 4096; whereas Y_j's can be programmed by the user, except the last one which is pre-fixed to Y_{33}=4096. This enables implementation of any non-linear characteristic, such as the ones required for gamma-correction or histogram equalization.

For colour devices, the 32-point programmable values can be split into two sets with 16-point each, so that 2 different LUTs can be programmed. Since on-chip digital processing is implemented so that each DIP processes pixels of two colours only, this is useful to perform PPT per colour. This involves point-topoint transformations.

Readout channel binning. When enabled via configuration register, CMOS pixel binning with enhanced image quality is performed. It is accomplished through averaging of the response of the binned readout channels. This function is also performed at the DIP level.

A. On-chip image corrections

Vertical FPN (VFPN). The readout path includes a digital calibration mechanism to remove column-to-column VFPN associated to the use of multiple readout channels. The DIP incorporates a memory to store correction data for each readout column. In order to perform column-to-column VFPN correction, it is necessary to update one correction register per internal readout channel. The calibration can be done using either the test rows or the OB rows of the pixel array. Both options are available to calculate the zero signal image level. The process of VFPN calibration is not performed during normal acquisition. This is a calibration step and must be performed during the camera startup phase and occasionally in the field, especially if operational conditions change significantly. Acquiring calibration data is performed several times in each calibration cycle. This is done by reading several times the test row or by reading several OB lines. Multiple readings reduce the effect of the temporal noise in the signal path (by averaging). If the selected VFPN calibration mode is external, data will be outputted through the LVDSs like a regular image. In this case, the user must process this data externally to compute the correction coefficients and put them in at the internal memories via

SPI.

Horizontal FPN (HFPN). The sensor can calibrate and correct row-by-row (horizontal) fixed-pattern noise (HFPN). Such noise is normally due to a random, row-to-row variation in the readout channel response, which generates a horizontal pattern superimposed to the signal. In comparison to VFPN, this error varies randomly from frame to frame as it is not associated to the readout channel mismatch, but to random temporal noise (from supplies, reference voltages, and bias currents) that is "frozen" (sampled and held) when a pixel row is read out. Consequently, the error is not the same for all pixels in a row, but changes from row to row and from frame to frame. In order to remove the HFPN, the sensor uses the information from OB columns. When a row is read out, the first pixels that are digitized correspond to the OB columns. Via a configuration register, some of these OB columns can be selected to enter an averaging block to compute a per-row value of the error. It is important to note that since the sampled OB columns belong to the same row, the row-to-row HFPN sensed will be the same as in the remaining (active) pixels in the row. Such correlation allows cancelling this noise out by simply subtracting the averaged value from the row regular pixels. Note that this function inherently provides auto-black level control. In fact, row-by-row black level compensation is more effective than performing a global calibration because with the latter large dark currents (when the temperature is high) may produce wrong black level correction for pixels that are processed at the beginning or at the end of the readout time.

Defective pixels correction. The sensor can correct on-chip up to 512 defective pixels. A dedicated internal memory stores the position of defective pixels. Changes in sub-sampling or binning will require a re-configuration of the defective pixels memory. Defective pixels in monochrome and colour devices are handled differently. A defective pixel is tagged and replaced by one of its closest active neighbours in the same row and within the same group of 16 pixels. Defective pixel correction by closest neighbor replacement and defective pixel correction in monochrome 2x2 sub-sampling case and in colour show examples of defective pixel replacements. Note that the neighbouring location changes when either sub-sampling or binning are applied or the device is a colour sensor.

IV. Illustrative Results

Fig.3 shows a summary of the sensor performance and a representative image capture. The sensor qualifies for a large variety of machine vision applications for "-D cameras as well as for 3-D cameras. This sensor sets a milestone in the evolution towards CIS with embedded intelligence. The trend towards the design of *systems* with reduced SWaP (*Size*, *Weight* and *Power*) call for sensors capable of making content analysis and of prompting decisions at maximum speed.

V. Acknowledgement

This sensor development has been partially supported by the Spanish Ministry of Science and Innovation through Project IPT-2011-1625-430000.

VI. References

[1] www.anafocus.com.
[2] Takayanagi and J. Nakamura, "High-Resolution CMOS Video Image Sensors". *Proceedings of the IEEE*, to appear.
[3] S. Huang et al., "Design of a PTC-Inspired Segmented ADC for High-Speed Column-Parallel CMOS Image Sensor". 2011 IISW, pp. 328-331, June 2011.
[4] X. Wang et al., "A 2.2M CMOS Image Sensor for High-Speed Machine Vision Applications". Proc. SPIE, vol. 7536, January 2010.
[5] Jan Bogaerts et al., "High Speed 36Gbps 12Mpixel Global Pipelined Shutter CMOS Image Sensor with CDS". 2011 IISW, pp. 335-338, June 2011.
[6] A. Rodríguez-Vázquez et al., "A CMOS Vision System On-Chip with Multi-Core, Cellular Sensory-Processing Front-End". Chapter 6 in Cellular Nanoscale Sensory Wave Computers (edited by C. Baatar, W. Porod and T. Roska). Springer 2010.

Performance Specs	Gain x 1	Gain x 2	Gain x 4	
DR (temporal)	64,46	61,47	56,29	dB
SNR_max	44,37	40,36	37,54	dB
DR (total)	58,57	54,86	49,25	dB
DSNU	0,10	0,16	0,31	%FS
VFPN (dark)	0,02	0,03	0,05	%FS
HFPN (dark)	0,01	0,02	0,04	%FS
PRNU Bidimensional	0,47	0,66	1,11	%
Bright VFPN	0,10	0,11	0,15	%
Bright HFPN	0,15	0,18	0,23	%
Non-linearity	+/-0.9	+/-2.4	+/-4.6	%FS

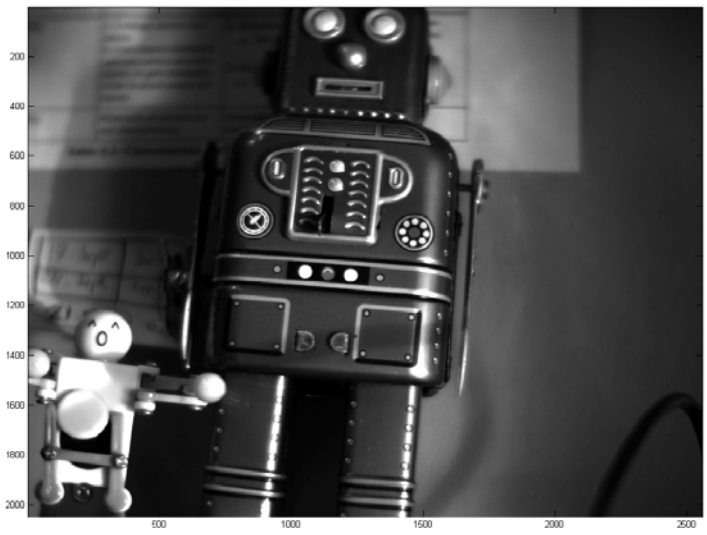

Figure 3. Illustrative performance measurements and image acquisition